BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 8, Number 2, March 1983

STRICTLY PSEUDOCONVEX DOMAINS IN C^n

BY MICHAEL BEALS, CHARLES FEFFERMAN AND ROBERT GROSSMAN

CONTENTS

0. Foreword
Chapter 1. Brief Introduction
Chapter 2. Mechanics
1. Newton's equations and canonical transformations
2. Generating functions
3. The Hamilton-Jacobi equation
Chapter 3. Elliptic partial differential equations
1. SIOs and ψDOs
2. Boundedness of SIOs on $\Lambda(\alpha)$
3. Boundedness of ψDOs on L^2
4. Stationary phase and the calculus of ψDOs
5. Elliptic regularity
6. Schauder theory
7. Boundary value problems
Chapter 4. The wave equation
1. A parametrix for $\Box$
2. Hörmander's theorem
3. Fourier integral operators and stationary phase
4. Egorov's theorem
5. The composition formula
6. Proof of Hörmander's Theorem
7. Introduction to global Fourier integral operators
8. Introduction to Fourier integral operators with complex phase
Chapter 5. Elementary differential geometry and the heat equation
1. Riemannian manifolds
2. Complex manifolds
3. Bundles over Riemannian manifolds
4. The Laplace operator
5. The Euler characteristic
6. Asymptotics of the heat kernel
7. The Chern-Gauss-Bonnet Theorem
8. Hermann Weyl's invariant theory
9. Gilkey's invariant theory
Chapter 6. An overview of topics in several complex variables
1. Introduction
2. Linear analysis: $\bar\partial$, $\Box$, $\bar\partial_b$, $\Box_b$
3. The unit ball and the Siegel domain

Received by the editors June 10, 1982.
1980 *Mathematics Subject Classification*. Primary 32F15.

0273-0979/82/0000-0716/$91.75

4. Geometry of strictly pseudoconvex domains: Moser normal form
5. The Poincaré metric and the complex Monge-Ampère equation
6. Refined linear analysis
7. Reflections

Chapter 7. Analysis on the Siegel domain and its boundary, the Heisenberg group
1. Solution to $\bar{\partial}_b$, $\Box_b$ on the Heisenberg Group (Folland-Stein)
2. $\Box$ on the Siegel domain with $\bar{\partial}$-Neumann boundary conditions
3. Phong's theorem
4. Approximate solution to $\Box_b$ on strictly pseudoconvex domains

Chapter 8. Ellipsoids and the method of reflections
1. Introduction
2. The proof of Webster's theorem on algebraic mappings
3. Ellipsoids and biholomorphic maps between them
4. Smoothness of biholomorphic maps

Chapter 9. Moser's normal form
1. Moser's theorem for formal power series
2. Proof of Moser's theorem
3. The convergence of the Moser normal form

Chapter 10. Chern's theorem
1. The flat case
2. The general case
3. The ideas in the proof of Chern's theorem
4. Applications

Chapter 11. The complex Monge-Ampère equation
1. The Monge-Ampère equation
2. The Cheng-Yau theorem
3. Proof of boundary regularity
4. Chains

Chapter 12. The Bergman kernel
1. The asymptotic expansion
2. Christoffers' theorem
3. Parabolic invariant theory

References

0. Foreword by Fefferman. The goal of this article is to acquaint readers with analysis and geometry on smooth domains in $\mathbf{C}^n$. For domains with the simplest geometry (strictly pseudoconvex) a wealth of deep results came to light during the 1970s, and we restrict attention to this case. The state of our knowledge of more general (weakly pseudoconvex) smooth domains is much more primitive, although some outstanding results are known, notably on the Cauchy-Riemann equations (Kohn **[40]**) and the Poincaré metric (Cheng-Yau **[10]**). A natural problem is to extend the results presented here to more general domains.

One of our main themes is the close connection between the analysis and local geometry of domains. To understand the picture requires a lot of elementary background in geometry and differential equations. For completeness we have included a long exposition (Chapters 2–5) of the relevant background. Chapter 1 gives a brief introduction and Chapter 6 a detailed introduction to domains in $\mathbf{C}^n$, and finally Chapters 7–12 present the main results.

This paper grew out of a course I gave at Princeton during 1979–80, with notes taken by Beals and extended by Grossman. The contributions of both Beals and Grossman are pervasive, but responsibility for any errors lies with me. Since the course was given, relevant important results were discovered, especially by Kuranishi and by Lee and Melrose. We have incorporated here statements of their results. I spoke on some of the material in this paper at the recent symposium in honor of Poincaré; much of the modern work is rooted in his seminal ideas. The reader should see the excellent brief survey article by Wells [**66**], also presented at the Poincaré Symposium and overlapping strongly with our long exposition.

Special thanks are due Perry DiVerita, Lauri Hein, Maureen Kirkham, Annette Roselli, and Bonnie Tompsen who cheerfully typed the manuscript despite the pressure of a deadline long past due.

In studying several complex variables, I profited greatly from the deep insights of my colleagues, and it is a pleasure to thank them here. I am especially grateful to S. Bell, J. Faran, J. J. Kohn, J. Moser, L. Nirenberg, D. H. Phong, E. M. Stein, S. Webster, and S. T. Yau. Without a vast amount of work by Beals and Grossman, this paper could never have been completed. Without forceful prodding by F. Browder, it would not have been finished in the twentieth century.

CHAPTER 1. BRIEF INTRODUCTION

We shall study analogues in $\mathbf{C}^n$ of familiar ideas in one complex variable. The topics are as follows:

Cauchy-Riemann equations. A basic problem in several complex variables is to solve the inhomogeneous equations $\bar\partial u = \alpha$ with good bounds. Here u is a function of n complex variables $z_k = x_k + iy_k$, and $\bar\partial u$ stands for the n functions $\partial u/\partial \bar z_k = \frac{1}{2}(\partial u/\partial x_k + i\partial u/\partial y_k)$. The problem is important because, for instance, it allows us to patch local results into global theorems. To illustrate, let us try to find an analytic function $F(z)$ on a domain D which blows up only at a single boundary point p. As a first step, we find such a function F_0 defined only in a small neighborhood U of p. Next, take a smooth cutoff function ϕ supported in U and equal to one near p, and set $\alpha = \bar\partial(\phi F_0)$. Now α is globally defined on $\bar D$ and has no singularities anywhere, since $\phi = 1$ and F_0 is analytic near p. If we can solve the $\bar\partial$-equations with good bounds, then we can find a nice function u which is singular nowhere and satisfies $\bar\partial u = \alpha$. Therefore $F = \phi F_0 - u$ will be singular exactly at p, and will be analytic in D since $\bar\partial F = \bar\partial(\phi F_0) - \bar\partial u = \alpha - \alpha = 0$.

The $\bar\partial$ equations are overdetermined—there are n equations for one function—so they can be solved only when α satisfies consistency conditions. Also, the solution u is obviously not unique. If $\bar\partial u = \alpha$ and F is any analytic function on D, then also $\bar\partial(u + F) = \alpha$. So it is natural to try to solve $\bar\partial u = \alpha$ in D with the extra condition that u be orthogonal to the subspace $H(D)$ of analytic functions $\subseteq L^2(D)$. This is called *Kohn's solution* of $\bar\partial u = \alpha$; it minimizes the L^2-norm among all solutions.

There is also a family of Cauchy-Riemann equations for analytic funtions on the boundary ∂D of a domain in $\mathbf{C}^n$ $(n > 1)$. Imagine that F is a function defined in all of $\mathbf{C}^n$, but whose values are known to us only on ∂D. We can easily calculate the derivatives of F in directions tangent to ∂D, but we do not know the normal derivative $\partial F/\partial n$. Now suppose that F is analytic in a neighborhood of ∂D, so that we have the n Cauchy-Riemann equations $\partial F/\partial \bar{z}_k = 0$ on ∂D. We can solve one of these equations for the missing derivative $\partial F/\partial n$, and then substitute the result into the remaining $(n - 1)$ Cauchy-Riemann equations. Thus we obtain a system of $(n - 1)$ partial differential equations for the restriction of an analytic functions F to ∂D. One writes $\bar{\partial}_b F = 0$, and again we are interested in the inhomogeneous equation $\bar{\partial}_b F = \alpha$ on ∂D.

Cauchy integral formula. The idea of solving $\bar{\partial} u = \alpha$ with u orthogonal to $H(D)$ (the analytic functions) suggests that we study the orthogonal projection π: $L^2(D) \to H(D)$. One shows easily that π is given by an integral kernel, $\pi f(z) = \int_D K(z, w) f(w)\, dw$, where K is called the *Bergman kernel.* We shall explore the relation between the Bergman kernel and the geometry of the domain. The analogue of the Bergman kernel for the $\bar{\partial}_b$-problem is the *Szegö kernel* $K(z, w)$ which realizes the projection

$$\pi f(z) = \int_{\partial D} K(z, w) f(w)\, dw \tag{1}$$

from $L^2(\partial D)$ to the subspace $H^2(\partial D) = \{$Boundary values of analytic functions in $D\}$. For the unit disc in $\mathbf{C}^1$, the right-hand side of (1) is $(1/2\pi i)\oint_{|w|=1} f(w)/(z - w)\, dw$. So the Szegö kernel is the analogue of the Cauchy integral formula for domains in $\mathbf{C}^n$.

Dirichlet problem. Just as analytic functions are closely related to harmonic functions in one complex variable, so the problems $\bar{\partial} u = \alpha$, $\bar{\partial}_b u = \alpha$ are intimately related to certain second-order equations called $\square$, $\square_b$. To see how these arise, let us try to solve $\bar{\partial} u = \alpha$ in D with u orthogonal to analytic functions. A natural way to produce functions orthogonal to everything analytic is to start with w in the domain of the adjoint operator $\bar{\partial}^*$ and set $u = \bar{\partial}^* w$. If F is analytic, then $\langle u, F\rangle = \langle \bar{\partial}^* w, F\rangle = \langle w, \bar{\partial} F\rangle = 0$. The $\bar{\partial}$-equation therefore takes the form

$$\bar{\partial}\bar{\partial}^* w = \alpha, \tag{2}$$

$$w \in \operatorname{Domain}(\bar{\partial}^*). \tag{3}$$

The global condition $u \perp H(D)$ has now been replaced by (3), which is a boundary condition for the second-order differential equation (2). Since (2), (3) come from $\bar{\partial} u = \alpha$, we can hope for solutions only when α satisfies a consistency condition, which we write in the form $\bar{\partial}_1 \alpha = 0$. (Explicitly, $\partial u/\partial \bar{z}_k = \alpha_k$ can be solved only when $\partial \alpha_j/\partial \bar{z}_k - \partial \alpha_k/\partial \bar{z}_k = 0$.) For general α, possibly not satisfying the consistency condition, we replace (2), (3) by the boundary-value problem

$$\left(\bar{\partial}\bar{\partial}^* + \bar{\partial}_1^*\bar{\partial}_1\right) w = \alpha, \tag{4}$$

$$w \in \operatorname{Domain}(\bar{\partial}^*), \qquad \bar{\partial}_1 w \in \operatorname{Domain}\left(\bar{\partial}_1^*\right). \tag{5}$$

This is called the $\bar{\partial}$-Neumann problem. It can be solved for general α and it reduces to (2), (3) when $\bar{\partial}_1 \alpha = 0$. One finds that the second-order operator (4) is basically the Laplacian, but the boundary conditions (5) are more degenerate than Dirichlet or Neumann conditions and require deep analysis. The analogous construction for the $\bar{\partial}_b$-equation leads to a second-order equation $\square_b w = \alpha$ on ∂D. Here ∂D is a manifold without boundary, but $\square_b$ is not elliptic.

Riemann mapping theorem. Given two domains $D_1, D_2 \subseteq \mathbf{C}^n$, we want to know whether there is an analytic mapping Φ which carries D_1 one-to-one and onto D_2. (Φ is called *biholomorphic*.) In more than one complex variable, the answer is almost always "no". For example, a ball is not biholomorphic to an ellipsoid. This leads to *biholomorphic geometry*, the study of those concepts on domains and their boundaries which are preserved under biholomorphic maps.

The most interesting objects in the biholomorphic geometry of a smooth boundary ∂D are local invariants attached to points of ∂D, and a family of distinguished curves in ∂D called *chains*. The chains on the unit sphere S are the circles which arise by intersecting S with a complex line. On more general boundaries, the chains are solution curves of a system of second-order ordinary differential equations, much like geodesics on a hypersurface in $\mathbf{R}^n$.

The local invariants are not trivial to write down, but we can see at once that they exist. It is enough simply to count dimensions: To Nth order about p, a boundary ∂P looks like $\{\operatorname{Re} z_1 = f_N(\operatorname{Im} z_1, \operatorname{Re} z_2, \operatorname{Im} z_3, \ldots, \operatorname{Re} z_n, \operatorname{Im} z_n)\}$ for an Nth degree polynomial f_N. Thus ∂D is described to Nth order by a single Nth degree polynomial in $(2n - 1)$ variables. On the other hand, a biholomorphic map is described to Nth order by n polynomials of degree N in only n variables. For large N, one checks that the space of all possible boundaries has much higher dimension than the space of all possible biholomorphic maps.

It follows at once that many biholomorphic invariants may be attacked to the Taylor expansion of ∂D about p. In particular, a domain D can be biholomorphic to the unit ball only if its boundary satisfies a system of nonlinear partial differential equations. We shall write down these equations explicitly.

Schwartz reflection principle. A domain with real-analytic boundary may be written locally as $D = \{r(z, \bar{z}) < 0\}$ where $r(z, \bar{w})$ is a convergent power series in the independent variables $z, \bar{w}$. (For instance, if D is the ellipsoid $\Sigma_j 2\lambda_j \operatorname{Re}(z_j^2) + |z_j|^2 < 1$, then we may take $r(z, \bar{w}) = \Sigma_j(\lambda_j z_j^2 + \lambda_j \bar{w}_j^2 + z_j \bar{w}_j)$.) The power series $r(z, \bar{w})$ is determined by D up to multiplication by a nonvanishing factor, so the variety $V_w = \{z \in \mathbf{C}^n \mid r(z, \bar{w}) = 0\}$ is associated to D and w, independently of the choice of defining function r. In one complex variable, V_w is a point, and $w \to V_w$ is a conjugate-analytic reflection across ∂D; thus we recover the usual Schwartz reflection principle. What we obtain in higher dimensions is much stronger. For, each V_w is a codimension-one variety, and the family $\mathbb{S} = \{V_w \mid w \in \mathbf{C}^n\}$ is a biholomorphic invariant of the domain D. To see the power of this idea, suppose we try to classify the biholomorphic self-maps Φ of an ellipsoid. From the explicit $r(z, \bar{w})$ above, we see at once that $\mathbb{S}$ is a family of quadric hypersurfaces, and Φ has to carry each quadric in $\mathbb{S}$ to another quadric in $\mathbb{S}$. This is a severe restriction on Φ; for most ellipsoids Φ has to be linear.

Poincaré metric. On a domain $D \subseteq \mathbf{C}^n$ we look for a metric of constant negative curvature which degenerates at ∂D. Such a metric (with constant negative Ricci curvature) is given by

$$ds^2 = \sum_{j,k} \frac{\partial^2}{\partial z_j \partial \bar{z}_k} \log \frac{1}{u} dz_j \overline{dz_k},$$

where u is a solution of the complex Monge-Ampère equation

$$\det\left(\frac{\partial^2}{\partial z_j \partial \bar{z}_k} \log \frac{1}{u} \right) = u^{-(n+1)} \quad \text{in } D, \tag{6}$$

$$u = 0 \quad \text{at } \partial D.$$

For the unit disc in $\mathbf{C}^1$, $u(z) = 1 - |z|^2$, and ds^2 is the familiar Poincaré metric $ds^2 = |dz|^2/(1 - |z|^2)^2$. For suitable smooth domains $D \subseteq \mathbf{C}^n$, equation (6) has a unique solution u, which is known to be smooth in the interior. Near the boundary, u has an asymptotic expansion $u \sim \psi \Sigma_{k=0}^{\infty} \phi_k (\psi^{n+1} \log \psi)^k$, where ϕ_k are smooth functions on $\bar{D}$ and $\psi(z) =$ distance from z to ∂D. In particular, $u \in C^{n+2-\varepsilon}(\bar{D})$. The functions ϕ_k carry a lot of information on the local biholomorphic geometry of ∂D, and appear in the asymptotic expansion of the Bergman kernel.

Throughout this article, we restrict attention to strictly pseudoconvex domains in $\mathbf{C}^n$. This is by far the simplest class to study, and it includes many interesting examples. We recall the elementary definitions: A domain D is strictly pseudoconvex if its Levi form is strictly positive definite at every boundary point. The Levi form of $D = \{z \in \mathbf{C}^n \mid r(z) < 0\}$ with $r \in \mathbf{C}^\infty$, $r' \neq 0$ on ∂D is defined as the restriction of the quadratic form

$$(\xi_k) \to \sum_{j,k} \frac{\partial^2 r}{\partial z_j \partial \bar{z}_k}(p) \xi_j \overline{\xi_k}$$

to the subspace $\{(\xi_k) \in \mathbf{C}^n \mid \Sigma_k (\partial r / \partial z_k)(p) \cdot \xi_k = 0\}$. It is defined up to constant multiples, independently of the choice of r. Unless the Levi form is at least semidefinite at every boundary point, every analytic function on D continues analytically into a fixed $D^+ \supsetneq D$.

Our emphasis will be on the "big picture" and the interrelationships of the different ideas. A main theme is the analogy between domains $D \subseteq \mathbf{C}^n$ and curved Riemannian manifolds M:

	Riemannian M	Domains D
Simplest Case	$\mathbf{R}^n$ with Euclidean metric	unit ball
Analytic Problems	Laplace equation	$\bar{\partial}, \bar{\partial}_b, \Box, \Box_b$
Geometric Invariants	curvature	Chern-Moser invariants

In both settings, the strategy is to get a good understanding of the simplest case, then attack the general case either by approximating a curved M by a flat $\mathbf{R}^n$, or by approximating ∂D by a sphere. The analogy goes further than shown here, but now it is time to present the background information. In Chapter 6 we return to give a more detailed summary of the several complex variables that is our main goal.

CHAPTER 2. MECHANICS

1. Newton's equations and canonical transformations. This chapter contains an introduction to those parts of mechanics which will be needed later in the discussion of pseudodifferential operators (ψDOs) and Fourier integral operators (FIOs). We begin by considering the dynamics of interacting particles with N degrees of freedom, with masses m_i, and whose positions are described by vectors with components q_i. If the particles interact to form a conservative system, then the force exerted on the ith particle can be expressed as $-\partial V(q)/\partial q_i$, for some potential function $V(q)$. The dynamics are determined by Newton's equations

$$m_i \frac{d^2 q_i}{dt^2} = -\frac{\partial V(q)}{\partial q_i}, \qquad i = 1, \ldots, N. \tag{1}$$

We can expose some of the symmetries hidden in this equation, by introducing the conjugate variables

$$p_i = m_i \dot{q}_i, \qquad i = 1, \ldots, N,$$

and the Hamiltonian

$$H(q, p) = T(p) + V(q),$$

where $T(p)$ is the kinetic energy

$$T(p) = \frac{1}{2} \sum_{i=1}^{N} \frac{p_i^2}{m_i}.$$

In the new variables $(q_1, \ldots, q_N, p_1, \ldots, p_N)$ Newton's equations (1) become the Hamilton equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = \frac{-\partial H}{\partial q_i}, \qquad i = 1, \ldots, N. \tag{2}$$

EXAMPLE. The Kepler problem is concerned with a particle of mass m confined to a plane and moving in a central field with potential energy $V(r) = -k/r$. In polar coordinates (r, θ) the kinetic energy of the particle is

$$T = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2).$$

Recall that the angular momentum of the particle is $mr^2\dot{\theta}$. This suggests the coordinate transformation (the Legendre transformation)

$$(r, \theta, \dot{r}, \dot{\theta}) \to (r, \theta, p_r, p_\theta), \qquad p_r = m\dot{r}, \qquad p_\theta = mr^2\dot{\theta}; \tag{3}$$

the kinetic energy is now

$$T = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2}\right),$$

and the Hamiltonian becomes

(4) $$H = T + V = \left(p_r^2 + \frac{p_\theta^2}{r^2} \right) - \frac{1}{r}.$$

(Assume that $m = \frac{1}{2}$ and the gravitational constant $k = 1$.) We shall see below that Hamilton's equations (2) still hold in terms of the new p's, q's and H. Since the variable θ does not appear in the Hamiltonian (such a variable is called cyclic), Hamilton's equations (2) imply $0 = -\partial H/\partial\theta = \dot{p}_\theta$, i.e. the angular momentum $p_\theta = l$ is conserved. We can now integrate (3) to find

(5) $$\theta = 2l \int \frac{dt}{r^2}.$$

Applying (2) again gives

$$\frac{dH}{dt} = \sum \left(\frac{\partial H}{\partial q_i} \dot{q}_i + \frac{\partial H}{\partial p_i} \dot{p}_i \right) = 0,$$

and so the total energy $H = E$ is also conserved; hence from (3) and (4), $\dot{r}^2/4 + \omega(r) = E$, or

(6) $$\int dt = \int \frac{dr}{2\sqrt{E - \omega(r)}},$$

where $\omega(r) = l^2/r^2 - 1/r$ (the effective potential energy). Together the equations (5) and (6) provide a complete description of the path of the particle.

Note that the first step in solving the problem is to change from rectangular to polar coordinates

$$\Phi\colon (x, y) \to (r, \theta).$$

In rectangular coordinates the Hamiltonian involves all four variables $(x, y, \dot{x}, \dot{y})$. Since none of the variables is conserved, we are not able explicitly to integrate the system in these coordinates. This suggests that we look for maps $\Phi^\#$ of *phase space* $\mathbf{R}^{2n} = \{(q_1, \ldots, q_n, p_1, \ldots, p_n)\}$ which preserves the form of Hamilton's equation. We will see in example (2) below that the coordinate change Φ induces a map $\Phi^\#$ taking the Hamiltonian in the $(x, y, \dot{x}, \dot{y})$ variables into the Hamiltonian (4). This is why we are justified in using the coordinates $(r, \theta, p_r, p_\theta)$ to solve the problem.

We begin by defining the Poisson bracket $\{F, G\}$ of two functions $F(q, p)$, $G(q, p)$ on phase space

(7) $$\{F, G\} \equiv \sum_{k=1}^{n} \left(\frac{\partial F}{\partial p_k} \frac{\partial G}{\partial q_k} - \frac{\partial F}{\partial q_k} \frac{\partial G}{\partial p_k} \right).$$

Consider an integral curve $t \to z(t) = (q(t), p(t))$ of Hamilton's equations (2) and the restriction of a function $F(q, p)$ on phase space to this curve.

PROPOSITION 1.

(*) $$\frac{dF}{dt}(z(t)) = -\{H, F\}(z(t)).$$

PROOF.

$$\frac{dF}{dt}(q(t), p(t)) = \sum\left(\frac{\partial F}{\partial q_k}\dot{q}_k + \frac{\partial F}{\partial p_k}\dot{p}_k\right) = \left(\frac{\partial F}{\partial q_k}\frac{\partial H}{\partial p_k} + \frac{\partial F}{\partial p_k}\left(\frac{-\partial H}{\partial q_k}\right)\right). \quad \square$$

We define a *canonical transformation* Φ: $\mathbf{R}^{2n} \to \mathbf{R}^{2n}$ to be a map of phase space which preserves the Poisson brackets, i.e.

$$\{F, G\} \circ \Phi = \{F \circ \Phi, G \circ \Phi\}.$$

If we let F be one of the variables q_i, p_i, we see immediately from Proposition 1 that equation $(*)$ characterizes Hamiltonian paths, so that canonical transformations preserve the form of Hamilton's equations.

EXAMPLE 1.

$$\Phi\colon (q, p) \to (Q, P) \equiv (-p, q). \tag{8}$$

Since a canonical transformation Φ: $(q, p) \to (Q(q, p), P(q,p))$ preserves Poisson brackets, we find after comparing the Poisson brackets $\{P, Q\}$ and $\{p, q\}$ that

$$\{P_j, P_k\} = \{Q_j, Q_k\} = 0, \qquad \{P_j, Q_k\} = \delta_{jk}. \tag{9}$$

Conversely, we see by checking the definition that if (9) holds then the transformation is canonical. Equation (9) is a system of quadratic equations for the elements of the Jacobian matrix Φ'; in fact, these equations are just a restatement of the matrix equation

$$(\Phi')^t J(\Phi') = J, \tag{10}$$

where $J = \left(\begin{smallmatrix} 0 & I \\ -I & 0 \end{smallmatrix}\right)$. A matrix A is called symplectic if $A^tJA = J$. We have

PROPOSITION 2. *Φ is canonical $\Leftrightarrow$ Φ' is symplectic at every point.*

A simple calculation shows that a symplectic transformation preserves volume and orientation.

We will also need a third means of characterizing canonical transformations. This involves the *symplectic form*

$$\omega = \sum_{k=1}^{n} dp_k \wedge dq_k. \tag{11}$$

A straightforward calculation shows

PROPOSITION 3. *Φ is canonical $\Leftrightarrow$ $\Phi^*\omega = \omega$.*

EXAMPLE 2. A map Φ: $q \to Q(q)$ of just the position variables extends to give a canonical map

$$\Phi^{\#}\colon (q, p) \to (Q, P) \equiv \left(\Phi(q), \left[(\Phi'(q))^t\right]^{-1} p\right). \tag{12}$$

Indeed, it suffices to check that

$$\sum p_k dq_k = \sum P_k dQ_k, \tag{13}$$

since the exterior derivative of this equation gives $\Sigma dp_k \wedge dq_k = \Sigma dP_k \wedge dQ_k$ as required. We have

$$\sum P_k dQ_k = \langle P, dQ \rangle = \langle P, \Phi'(q)dq \rangle = \left\langle (\Phi'(q))^t P, dq \right\rangle.$$

Therefore if we put $p = (\Phi'(q))^t P$ or $P = [(\Phi'(q))^t]^{-1} p$, (13) holds and (12) becomes a canonical transformation. Note that $\Phi^{\#}$ sends the Hamiltonian

$$H = \frac{1}{2} \sum_j \frac{p_j^2}{2m_j} + V(q)$$

into the Hamiltonian

$$H = \frac{1}{2} \sum_{i,j} g^{ij}(Q) P_i P_j + \tilde{V}(Q),$$

where $\Sigma_j m_j \dot{q}_j^2 = \Sigma_{i,j} g_{ij}(Q)\dot{Q}_i \dot{Q}_j$, $g^{ij} = (g_{ij})^{-1}$ and $\tilde{V}(Q) = V(q)$. We can now finish our discussion of the Kepler problem. The change of variables Φ: $(x, y) \to (r, \theta)$ into polar coordinates induces a canonical transformation $\Phi^{\#}$ carrying the old Hamiltonian into the Hamiltonian (4). We are therefore justified in using the coordinates $(r, p_r, \theta, p_\theta)$ when calculating the orbit of the particle.

EXAMPLE 3. We end this section by giving one final example of a canonical transformation. This is a transformation used by Sundmann in a regularization of the three-body problem. We begin by considering two bodies in $\mathbf{R}^1$ separated by a distance q and put $p = \dot{q}$. The Hamiltonian for gravitational attraction is

$$H = p^2 - 1/q,$$

and we know that H remains constant as p, q evolve by Hamilton's equations. Thus, $H = E$ where E is the energy. For small q, this yields $\dot{q}^2 \sim 1/q$ and so $\dot{q} \sim q^{-1/2}$ or $q^{3/2} \sim t$. This gives

$$q \sim t^{2/3}, \qquad p \sim t^{-1/3}.$$

The map

$$\Phi\colon (q, p) \to (Q, P) = (-p^2 q, 1/p) \tag{14}$$

satisfies $dP \wedge dQ = (-1/p^2)dp \wedge -p^2 dq = dp \wedge dq$ and is therefore canonical. In these new coordinates, $P \sim t^{1/3}$, $Q \sim 1$ as we approach a collision at time $t = 0$.

The treatment of the three-body problem is simplified by using a change of clock. Suppose first that $H = 0$ on the path we are interested in. If we use a new Hamiltonian $H \to FH$, Hamilton's equations become

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \cdot F, \qquad \frac{dp_i}{dt} = \frac{-\partial H}{\partial q_i} \cdot F,$$

and we are led to the change of variables $t \to \tau$ so that $d\tau = dt/F$. In the general case, we have $H - E = 0$ along a given Hamiltonian path and so we can make a change of clock by using a new Hamiltonian

$$H \to F(H - E). \tag{15}$$

After these preliminaries we now turn to the three-body problem itself. We assume that two of the bodies are near collision, while the third body is well separated from the other two. Let q_1, q_2, q_3 measure the displacement between the two nearby bodies and put $p_i = \dot{q}_i$, $i = 1, 2, 3$. The Hamiltonian is

$$H = p_1^2 + p_2^2 + p_3^2 - \frac{1}{\left(q_1^2 + q_2^2 + q_3^2\right)^{1/2}} + \text{Junk}. \tag{16}$$

Along a given Hamiltonian path H is conserved, $H = E$. We introduce the canonical transformation

$$\Phi\colon (q, p) \to (Q, P) = \left(L(p)q, p/\|p\|^2\right), \tag{17}$$

where $L(p)$ is the linear transformation $L(p) = [(\partial P/\partial p)^t]^{-1} = \|p\|^2 R_p$, and R_p denotes reflection through the plane normal to p. This satisfies $\|Q\| = \|p\|^2\|q\|$ and sends the Hamiltonian (16) into

$$H = \|P\|^{-2} - \|P\|^{-2}\|Q\|^{-1} - E + \text{Junk}.$$

After a change of clock (15) with $F = \|P\|^2$, this becomes

$$\tilde{H} = 1 - \|Q\|^{-1} + \|P\|^2(\text{Junk} - E). \tag{18}$$

Along our path we have $\tilde{H} = 0$, while near collisions $\|p\| \gg 1$. Thus $\|P\| \ll 1$ and $\|Q\|^{-1} \approx 1$ near the collision, and therefore $\tilde{H}$ has a nice nonsingular solution. We can conclude that the solution of Newton's equations for the three-body problem continues in a natural way past simple collisions. This calculation will reappear in Chapter 11.

2. Generating functions. In the last section we saw that there are at least three ways to decide whether a transformation is canonical or not. On the other hand it is not so clear how to construct canonical transformations. One means of doing this is to use generating functions. Given a canonical transformation

$$\Phi\colon (q, p) \to (Q, P),$$

consider the graph of Φ

$$\Gamma = \{(q, p, Q, P)\colon \Phi(q, p) = (Q, P)\}.$$

By Proposition 3 we know that on Γ

$$d\left(\sum q_i dp_i + \sum P_i dQ_i\right) = -\left(\sum dp_i \wedge dq_i - \sum dP_i \wedge dQ_i\right) = 0,$$

and we can conclude that locally there exists a function S on Γ such that

$$\sum (q_i dp_i + P_i dQ_i) = dS.$$

Assume now that Φ is a small perturbation of the identity. This means that on Γ instead of the coordinates (q, p) we can use (p, Q). We have $dS(p, Q) = \Sigma((\partial S/\partial p_i)dp_i + (\partial S/\partial Q_i)dQ_i)$ and therefore

$$q_i = \frac{\partial S(p, Q)}{\partial p_i}, \qquad P_i = \frac{\partial S(p, Q)}{\partial Q_i}.$$

The function S is called a *generating function* for the canonical transformation.

EXAMPLE. The identity map Φ: $(q, p) \to (Q, P) = (q, p)$ arises from the generating function $S(p, Q) = \Sigma p_i Q_i$.

In fact each of the steps taking us from Φ to S can be reversed, giving us the

PROPOSITION 4. *If $S(p, Q)$ is a small perturbation of $\Sigma p_i Q_i$, then defining*

$$q_i = \frac{\partial S(p, Q)}{\partial p_i}, \qquad P_i = \frac{\partial S(p, Q)}{\partial Q_i} \tag{19}$$

gives a canonical transformation

$$\Phi\colon (q, p) \to (Q, P).$$

PROOF. The graph

$$\Gamma = \{(q, p, Q, P)\colon \text{equations (19) hold}\}$$

is a manifold. Because $S(p, Q)$ is a small perturbation of $\Sigma p_i Q_i$, we can by the implicit function theorem use either (q_i, p_i) or (p_i, Q_i) as coordinates. In particular, Γ is the graph of a map Φ: $(p, q) \to (P, Q)$. By (19) the identity

$$\sum_j P_j dQ_j + \sum q_j dp_j = dS(p, Q)$$

holds on Γ. Therefore

$$\sum_j dP_j \wedge dQ_j - \sum dp_j \wedge dq_j = 0$$

on Γ and Φ is a canonical transformation by Proposition 3. $\square$

3. The Hamilton-Jacobi equation. Consider the transformation of phase space

$$\Phi_t\colon (x, \xi) \to (y, \eta) \tag{20}$$

defined by flowing along the integral curves to Hamilton's equations

$$\frac{dy_j}{dt} = \frac{\partial H}{\partial \eta_j}(y, \eta), \qquad \frac{d\eta_j}{dt} = \frac{-\partial H}{\partial y_j}(y, \eta), \qquad y(0) = x, \eta(0) = \xi, \tag{21}$$

for a time t. This is called the *Hamiltonian flow.*

PROPOSITION 5. *The Hamiltonian flow* (20) *is canonical.*

PROOF. Note that we make this assertion only for small t, since we are only guaranteed a solution to (21) for small t. We want to show

$$\Phi^*\omega = \omega. \tag{22}$$

Using the semigroup property of Φ_t: $\Phi^*_{t_0+s}\omega = \Phi^*_{t_0}(\Phi^*_s\omega)$ and differentiating with respect to s gives

$$\frac{d}{dt}(\Phi_t^*\omega)\,\Big|_{t=t_0} = \Phi^*_{t_0}\frac{d}{ds}(\Phi^*_s\omega)\,\Big|_{s=0}.$$

We see that we need only check that (22) holds to first order in t at $t = 0$. Since

$$y_j = x_j + t\frac{\partial H}{\partial \xi_j}(x, \xi) + O(t^2), \qquad \eta_j = \xi_j + t\frac{\partial H}{\partial x_j}(x, \xi) + O(t^2),$$

we have

$$\sum d\eta_j \wedge dy_j - \sum d\xi_j \wedge dx_j = \sum_j \left(d\xi_j - t\sum_k \frac{\partial^2 H}{\partial x_j \partial x_k} dx_k - t\sum_k \frac{\partial^2 H}{\partial x_j \partial \xi_k} d\xi_k \right)$$

$$\wedge \left(dx_j + t\sum_k \frac{\partial^2 H}{\partial \xi_j \partial \xi_k} d\xi_k + t\sum_k \frac{\partial^2 H}{\partial \xi_j \partial x_k} dx_k \right) + O(t^2) - \sum d\xi_j \wedge dx_j$$

$$= O(t^2).$$

Thus (22) holds to first order at $t = 0$ and the proposition is proved.□

For small time t the canonical transformation Φ_t of (20) is near the identity and so has a generating function; denote this by $S_t(y, \xi)$. We are going to find a first order nonlinear partial differential equation which the generating function satisfies. We begin by looking at the lower order terms in t of $S_t(y, \xi)$. Equation (19) implies that $dS_t(y, \xi) = \Sigma x_j d\xi_j + \Sigma \eta_j dy_j$ and since

$$x_j = y_j - t\frac{\partial H}{\partial \xi_j}(y, \xi) + O(t^2), \qquad \eta_j = \xi_j - t\frac{\partial H}{\partial x_j}(x, \xi) + O(t^2),$$

we have

$$dS_t(y, \xi) = \sum (y_j d\xi_j + \xi_j dy_j) - t\sum \frac{\partial H}{\partial \xi_j}(y, \xi) d\xi_j$$

$$-t\sum \frac{\partial H}{\partial y_j}(y, \xi) dy_j + O(t^2),$$

so that

$$S_t(y, \xi) = \sum \xi_j y_j - tH(y, \xi) + O(t^2). \tag{22}$$

Before we can begin to derive the nonlinear partial differential equation we need one more preliminary. Consider two successive canonical transformations

$$(x, \xi) \overset{\Phi}{\to} (y, \eta) \overset{\Psi}{\to} (z, \zeta) \tag{23}$$

with generating functions

$$\begin{array}{ll} S(y, \xi) & \text{associated to } \Phi, \\ T(z, \eta) & \text{associated to } \Psi \end{array}$$

and graph

$$\Gamma = \{(x, \xi, y, \eta, z, \zeta) \colon \Phi(x, \xi) = (y, \eta) \text{ and } \Psi(y, \eta) = (z, \zeta)\}.$$

Let $\mathbb{S}(z, \xi)$ be the generating function associated with $\Psi \circ \Phi$.

PROPOSITION 5. *$\mathbb{S}(z, \xi) = S(y, \xi) + T(z, \eta) - \Sigma\eta_j y_j$ whenever $(x, \xi, y, \eta, z, \zeta) \in \Gamma$.*

PROOF. Equation (19) implies that on Γ

$$dS(y, \xi) = \sum x_j d\xi_j + \sum \eta_j dy_j, \qquad dT(z, \eta) = \sum y_j d\eta_j + \sum \zeta_j dz_j$$

and so

$$dS(y,\xi) + dT(z,\eta) = \sum x_j d\xi_j + d\left(\sum \eta_j y_j\right) + \sum \zeta_j dz_j$$

or

$$\mathbb{S}(z,\xi) = S(y,\xi) + T(z,\eta) - \sum \eta_j y_j. \quad \square$$

Now assume that the canonical transformations (23) arise as Hamiltonian flows $\Phi = \Phi_{t'}$, $\Psi = \Phi_t$ with $t \ll 1$. Then from (22) we have

$$T(z,\eta) = \sum \eta_j z_j - tH(z,\eta) + O(t^2)$$

and thus by Proposition 5

$$\begin{aligned}\mathbb{S}(z,\xi) &= S(y,\xi) + \sum \eta_j(z_j - y_j) - tH(z,\eta) + O(t^2)\\ &= S(z,\xi) - \sum \frac{\partial S}{\partial y_j}(z,\xi)(z_j - y_j) + O(t^2)\\ &\quad + \sum \eta_j(z_j - y_j) - tH(z,\eta) + O(t^2)\\ &= S(z,\xi) - tH(z,\eta) + O(t^2),\end{aligned}$$

since $\eta_j = \partial S/\partial y_j$. Therefore if we put $S_t(z,\xi) = \mathbb{S}(z,\xi)$, we have

$$\frac{\partial}{\partial t}(S_t(z,\xi)) = \lim_{t\to 0} \frac{\mathbb{S}(z,\xi) - S(z,\xi)}{t} = \lim_{t\to 0} - H(z,\eta) = -H\left(z, \frac{S_t(z,\xi)}{\partial z_j}\right).$$

We have proved

THEOREM 1. *If $S_t(z,\xi)$ is the generating function associated with a Hamiltonian flow Φ_t, then $S_t(z,\xi)$ satisfies the first order nonlinear partial differential equation (with parameter ξ)*

$$\frac{\partial}{\partial t} S_t(z,\xi) + H\left(z, \frac{\partial S_t}{\partial z_j}(z,\xi)\right) = 0. \tag{24}$$

Equation (24) is called the Hamilton-Jacobi equation; it will occur later in our study of the wave equation. To solve it, the theorem tells us that we must simply calculate the integral curves of H and let the solution be the generating function of the resulting flow.

Finally, we mention two general references for this chapter; the book by Arnold [**1**] and the book by Goldstein [**28**].

CHAPTER 3. ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS

In this chapter we will consider the general second-order linear equation

$$\sum a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \text{lower order terms} = f, \tag{1}$$

where (a_{ij}) is a positive definite, smoothly varying matrix. We can try to solve (1) by freezing the coefficients at a point x_0 so that it is reduced to the constant coefficient equation

$$\sum a_{ij}(x_0)\frac{\partial^2 u}{\partial x_i \partial x_j} + \text{lower order terms} = f.$$

This approach leads us to the study of singular integral operators (SIOs) and pseudodifferential operators (ψDOs). These are defined in §1 and some of their basic properties are explored in the next three sections. In §5 we prove a regularity theorem about elliptic operators on manifolds.

We will also consider the nonlinear equation

$$A\big(x, u(x), \partial u/\partial x_i, \partial^2 u/\partial x_i \partial x_j\big) = 0, \tag{2}$$

where $A(x, \omega, \omega_i, \omega_{ij})$ is a smooth function on $\mathbf{R}^n \times \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^{n^2}$. This defines an elliptic equation if

$$A_{jk} \equiv \big(\partial A/\partial \omega_{jk}\big)$$

is a positive definite matrix. Recall that the Lipschitz space $\Lambda(\alpha)$ consists of those functions f satisfying

$$\sup_x |f(x+h) - f(x)| \leqslant C|h|^\alpha, \tag{3.i}$$

$$\sup_x |f(x)| \leqslant C. \tag{3.ii}$$

We assume that $0 < \alpha < 1$. The norm of f is the infimum of the C satisfying (3). In §6 we prove the following regularity theorem from Schauder theory: If u is a solution of (2) and $\partial^2 u/\partial x_i \partial x_j \in \Lambda(\alpha)$, then $u \in C^\infty$. Furthermore the norm $\|u\|_{C^k}$ is bounded by a quantity determined only by A and $\|\partial^2 u/\partial x_i \partial x_j\|_{\Lambda(\alpha)}$. The proof requires that we examine the linear equation (1) when the coefficients (a_{ij}) are no longer assumed to be smooth but only in $\Lambda(\alpha)$.

1. SIOs and ψDOs. Consider first the equation

$$\Delta u = f \tag{4}$$

in $\mathbf{R}^n$. A particular solution of (4) is given by convolving f with the fundamental solution

$$u(x) = c_n \int_{\mathbf{R}^n} \frac{f(x-y)}{|y|^{n-2}}\,dy = c_n \int_{\mathbf{R}^n} \frac{f(y)}{|x-y|^{n-2}}\,dy.$$

Here, as later, c_n denotes a constant depending only on the dimension of the space. We have

THEOREM 1. *If $f \in \Lambda(\alpha)$ and u satisfies* (4), *then $\partial^2 u/\partial x_i \partial x_j \in \Lambda(\alpha)$.*

To prove this, we must take the second derivatives in x of

$$u(x) = \int_{\mathbf{R}^n} \frac{f(y)}{|x-y|^{n-2}}\,dy.$$

Formally this requires that f be integrated against a kernel which is homogeneous of degree $-n$ and therefore not integrable. This is an example of an SIO.

Consider kernels

$$K(x) = \Omega(x)/|x|^n \tag{5}$$

on $\mathbf{R}^n$, where

$$\begin{aligned} &\text{(i)} \quad \Omega(x) \text{ is homogeneous of degree } 0,\\ &\text{(ii)} \quad \Omega(x) \text{ is } C^\infty \text{ away from } 0,\\ &\text{(iii)} \quad \int_{a<|y|<b} K(y)\,dy = 0 \text{ for all } a, b. \end{aligned} \tag{6}$$

By definition an SIO is the operator

$$f(x) \to K * f(x) \equiv \lim_{\substack{\delta\to 0\\ m\to\infty}} \int_{\delta<|y|<m} K(y) f(x-y)\,dy. \tag{7}$$

Note that this limit obviously exists when $f \in C_0^\infty$.

EXAMPLE 1. For $f \in L^p(\mathbf{R}^n)$, $1 \leqslant p < \infty$, the Riesz transforms R_j are defined by

$$\begin{gathered} R_j(f)(x) = \lim_{\varepsilon\to 0} c_n \int_{|y|>\varepsilon} \frac{y_j}{|y|^{n+1}} f(x-y)\,dy,\\ c_n = \Gamma\left(\frac{n+1}{2}\right)\Big/\pi^{(n+1)/2}, \qquad j = 1,\ldots,n. \end{gathered} \tag{8}$$

In other words the R_j are defined by the kernel $K_j(x) = \Omega_j(x)/|x|^n$, with $\Omega_j(x) = c_n x_j/|x|$.

EXAMPLE 2. We claim

$$\partial^2 f/\partial x_j \partial x_k = -R_j R_k \Delta f. \tag{9}$$

Indeed, since the Fourier transform of $\partial f/\partial x_j$ is $-2\pi i x_j f(x)$, we have

$$\begin{aligned} \left(\frac{\partial^2 f}{\partial x_j \partial x_k}\right)^{\wedge}(x) &= -4\pi^2 x_j x_k \hat{f}(x) = -\left(\frac{i x_j}{|x|}\right)\left(\frac{i x_k}{|x|}\right)(-4\pi^2 |x|^2)\hat{f}(x)\\ &= -(R_j R_k \Delta f)^{\wedge}, \end{aligned}$$

verifying (9). Here we have used the fact that

$$(R_j f)^{\wedge}(x) = i\frac{x_j}{|x|}\hat{f}(x). \tag{10}$$

Equation (10) can be checked directly or can be deduced by analyzing the effect of rotations on the R_j. Equation (9) shows that if $\Delta u = f$, then

$$f \to \partial^2 u/\partial x_i \partial x_j \tag{11}$$

is a composition of two SIOs. In fact it is a linear combination of the identity operator and an SIO; but we won't use that.

In §2 we will prove that SIOs are bounded on $\Lambda(\alpha)$. From Example 2, we see that this proves Theorem 1. In fact SIOs are bounded on L^p, $1 < p < \infty$, but we will not prove that here. An alternate approach to solving elliptic equations is via the Fourier transform. Consider the equation

$$\Sigma a_{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j} = f. \tag{12}$$

Freezing the coefficients at x_0 gives $\Sigma a_{ij}(x_0)\partial^2 u/\partial x_i \partial x_j = f$, and then taking Fourier transforms yields $-(\Sigma a_{ij}(x_0)\xi_i\xi_j)\hat{u}(\xi) = \hat{f}(\xi)$, so that the solution is $u(x) = c_n \int e^{ix\cdot\xi}(-a^{-1}(x_0, \xi))\hat{f}(\xi)\, d\xi$, if we suitably interpret the integral to take into account the zeros of $a(x_0, \xi)$. We might hope that the variable coefficient equation (12) is well approximated by the equation with the coefficients frozen at x_0 so that we would have an approximate solution given by

$$u(x) \approx \int e^{ix\cdot\xi}(-a^{-1}(x, \xi))\hat{f}(\xi)d\xi. \tag{13}$$

In fact when (12) is elliptic, an excellent approximate solution is given by (13), which is an example of a ψDO.

The growth of the function $a^{-1}(x, \xi)$ is important and is captured by considering those functions $p(x, \xi) \in C^\infty(\mathbf{R}^n \times \mathbf{R}^n)$ with the property that, for any multi-indices α and β, there exists a constant $C_{\alpha,\beta}$ such that

$$\left|D_x^\alpha D_\xi^\beta p(x, \xi)\right| \leqslant C_{\alpha,\beta}(1 + |\xi|)^{m-\rho|\beta|+\delta|\alpha|}, \tag{14}$$

for all $(x, \xi) \in \mathbf{R}^n \times \mathbf{R}^n$. The function $p(x, \xi)$ is called an mth order symbol and the class of symbols just defined is denoted $S^m_{\rho,\delta}$. We assume $m, \rho, \delta \in \mathbf{R}$, and that $0 \leqslant \rho, \delta \leqslant 1$. Associated to any symbol $p(x, \xi)$ is a pseudodifferential operator

$$f(x) \to p(x, D)f(x) \equiv \int e^{ix\cdot\xi}p(x, \xi)\hat{f}(\xi)\, d\xi. \tag{15}$$

EXAMPLE 3. A polynomial in ξ

$$p(x, \xi) = \sum_{|\alpha| \leqslant m} a_\alpha(x)\xi^\alpha$$

is clearly in $S^m_{1,0}$. The Fourier inversion formula shows that $p(x, D)$ defined by (15) is nothing but a differential operator with symbol $p(x, \xi)$. In other words, differential operators are examples of ψDOs.

REMARK. To apply the argument leading to (13), we would need the operators to satisfy $p(x, D) \circ q(x, D) \approx pq(x, D)$ in some well-defined sense so that $p(x, D) \circ p^{-1}(x, D) \approx I$. In §4 we will derive an asymptotic expansion for the composition of ψDOs so that these formulas become precise. It turns out that an elliptic symbol $p(x, \xi) \in S^m_{1,0}$ has a parametrix whose main term is in $S^{-m}_{1,0}$. A parametrix of an operator $p(x, D)$ is an operator $q(x, D)$ satisfying

$$p(x, D) \circ q(x, D) \sim I, \qquad q(x, D) \circ p(x, D) \sim I. \tag{16}$$

We summarize by listing some of the important properties of these operators:

SIOs: bounded on L^p, $1 < p < \infty$,
bounded on $\Lambda(\alpha)$, $0 < \alpha < 1$;

ψDOs: 0th order operators are bounded on L^2,
there are formulas for compositions and adjoints.

The book by Stein **[57]** on singular integral operators and the books by Taylor **[61]** and Trèves **[63]** on pseudodifferential operators are good references for this chapter.

Finally, we will denote the symbol class $S^m_{1,0}$ by S^m.

2. Boundedness of SIOs on $\Lambda(\alpha)$. Let T be an SIO

$$Tf(x) = \int_{\mathbf{R}^n} K(x-y)f(y)\,dy = \int_{\mathbf{R}^n} K(y)f(x-y)\,dy,$$

and let $f \in \Lambda(\alpha)$.

THEOREM 2. *T is bounded on $\Lambda(\alpha)$.*

PROOF. Without loss of generality we can assume

$$\text{(17)}\qquad \begin{aligned} &\text{(i)} && f \text{ is supported in } |x| \leq 2,\\ &\text{(ii)} && |f(x)| \leq 1,\\ &\text{(iii)} && |f(x) - f(x')| \leq |x - x'|^\alpha. \end{aligned}$$

With these assumptions it suffices to prove the estimates (3.i) and (3.ii) for $\delta \equiv |x - x'| \leq 1/10$. We know by (5) that

$$\text{(18)}\qquad |K(x)| \leq \frac{c}{|x|^n}, \qquad \left|\frac{\partial K}{\partial x_i}(x)\right| \leq \frac{c}{|x|^{n+1}}$$

and (6.iii) implies

$$\begin{aligned} Tf(x) &= \int_{\mathbf{R}^n} K(x-y)f(y)\,dy = \int_{|x-y|\leq 100} K(x-y)f(y)\,dy \\ &= \int_{|x-y|\leq 100} K(x-y)[f(y) - f(x)]\,dy; \end{aligned}$$

therefore

$$\begin{aligned} |Tf(x)| &\leq \int_{|x-y|\leq 100} |K(x-y)|\cdot|f(y)-f(x)|\,dy \\ &\leq \int_{|x-y|\leq 100} \frac{c}{|x-y|^n}\cdot|x-y|^\alpha\,dy \\ &\leq \int_{|x-y|\leq 100} \frac{c}{|x-y|^{n-\alpha}}\,dy < \infty, \end{aligned}$$

showing Tf is bounded. We must now prove that $|Tf(x) - Tf(x')| \leq c|x - x'|^\alpha$.

We begin by splitting the integral

$$\begin{aligned} Tf(x) - Tf(x') &= \int_{\mathbf{R}^n} K(y)[f(x-y) - f(x'-y)]\,dy \\ &= \int_{|y|\leqslant 10\delta} K(y)[f(x-y) - f(x'-y)]\,dy \\ &\quad + \int_{|y|>10\delta} K(y)[f(x-y) - f(x'-y)]\,dy \\ &= I_1 + I_2, \end{aligned}$$

and the integral I_1 into

$$\begin{aligned} I_1 &= \int_{|y|\leqslant 10\delta} K(y)f(x-y)\,dy - \int_{|y|\leqslant 10\delta} K(y)f(x'-y)\,dy \\ &= I_{11} + I_{12}. \end{aligned}$$

Using (6.iii) again,

$$I_{11} = \int_{|y|\leqslant 10\delta} K(y)f(x-y)\,dy = \int_{|y|\leqslant 10\delta} K(y)[f(x-y) - f(x)]\,dy$$

so that as before

$$|I_{11}| \leqslant \int_{|y|\leqslant 10\delta} \frac{c}{|y|^n}\cdot|y|^\alpha\,dy = c'\delta^\alpha.$$

Similarly, $|I_{12}| \leqslant c'\delta^\alpha$ and therefore $|I_1| \leqslant 2c'\delta^\alpha$.

Define $K_\delta(y) \equiv K(y)\chi_{10\delta<y<100}$, where $\chi_{10\delta<y<100}$ is the characteristic function. Now

$$I_2 = \int_{\mathbf{R}^n} [K_\delta(x-y) - K_\delta(x'-y)]\cdot[f(y) - f(x)]\,dy.$$

For the integrand to be nonzero we must take $|x-y| > 9\delta$ and $|x'-y| > 9\delta$, in which case, by (18) and the mean value theorem we have

$$|K_\delta(x-y) - K_\delta(x'-y)| \leqslant \frac{c\delta}{|x-y|^{n+1}},$$

so that

$$|I_2| \leqslant \int_{|x-y|\geqslant 9\delta} \frac{c\delta\,|x-y|^\alpha}{|x-y|^{n+1}}\,dy \leqslant c'\delta^\alpha.$$

Therefore $I_1 + I_2 \leqslant c\delta^\alpha$ and the theorem is proved. □

3. Boundedness of ψDOs on L^2. Recall that a symbol $p(x,\xi) \in S^0_{1/2,1/2}$ satisfies estimates of the form

$$\left|\partial_x^\alpha \partial_\xi^\beta p(x,\xi)\right| \leqslant C_{\alpha,\beta}(1 + |\xi|)^{|\alpha|/2 - |\beta|/2}. \tag{19}$$

THEOREM 3. *$p(x,\xi) \in S^0_{1/2,1/2} \Rightarrow p(x,D)$ is bounded on L^2.*

Let Q be a cube in (x, ξ)-space whose sides are parallel to the axes and satisfying

$$\operatorname{diam}_x Q = 1, \qquad \operatorname{diam}_\xi Q = M.$$

LEMMA 1. *It is sufficient to prove Theorem 3 for $p(x, \xi)$ supported in Q, with bound independent of M.*

PROOF OF LEMMA 1. The proof consists of three steps. First we localize to cubes in phase space using a partition of unity. Then we patch together the operators associated with each of these cubes using a lemma due to Cotlar and Stein. Finally we must verify that the estimates occurring in the hypothesis of the Cotlar-Stein lemma are satisfied.

Let $\{Q_\nu\}$ be a partition of phase space into blocks of various sizes, centered at points (x_ν, ξ_ν), satisfying

$$\text{(20)} \qquad \text{(i) } \operatorname{diam}_x Q_\nu = 1, \qquad \text{(ii) } \operatorname{diam}_\xi Q_\nu \approx |\xi_\nu|\,.$$

For example we can partition x-space into cubes of diameter 1 whose sides are parallel to the axes and we can partition ξ-space into similar cubes whose diameters are given by the lengths of the dyadic intervals $[0,1], [1,2], [2,4], \ldots, [2^j, 2^{j+1}], \ldots$. Let $\{\phi_\nu(x, \xi)\}$ be a partition of unity of phase space such that

$$\text{(21)} \qquad \begin{aligned} &\text{(i) } \varphi_\nu(x, \xi) = 1 \text{ in a neighborhood of } (x_\nu, \xi_\nu), \\ &\text{(ii) } \operatorname{supp} \phi_\nu \subset Q_\nu^*, \end{aligned}$$

where the double Q_ν^* is obtained from Q_ν by doubling the diameter and keeping the center fixed. We can find such a partition of unity by translating and dilating a fixed cutoff function and then normalizing so that the sum is always one. We may take our partition to satisfy $\Sigma_\nu \chi_{Q_\nu^*} \leqslant C$.

Using the partition of unity we can break up the symbol into pieces

$$\text{(22)} \qquad p_\nu(x, \xi) \equiv \phi_\nu(x, \xi) p(x, \xi).$$

Note that $p_\nu(x, \xi)$ is supported in a cube Q_ν^* and that if the center ξ_ν of Q_ν is of the order $|\xi_\nu| \approx M$, then $p_\nu(x, \xi)$ satisfies the estimates

$$\text{(23)} \qquad \left|\partial_x^\alpha \partial_\beta^\xi p_\nu(x, \xi)\right| \leqslant C_{\alpha,\beta} M^{|\alpha|/2 - |\beta|/2}.$$

By hypothesis we can assume

$$\text{(24)} \qquad p_\nu(x, D) \text{ are uniformly bounded on } L^2.$$

We must show that $p(x, D) = \Sigma p_\nu(x, D)$ is bounded on L^2. This is a consequence of

LEMMA 2 (COTLAR-STEIN). *Let $\{A_j\}$ be a sequence of bounded operators on a Hilbert space satisfying*

$$\text{(25)} \qquad \begin{aligned} &\text{(i)} && \|A_j\| \leqslant B, \\ &\text{(ii)} && \|A_j^* A_k\| \leqslant C(j-k), \\ &\text{(iii)} && \|A_j A_k^*\| \leqslant C(j-k); \end{aligned}$$

then

$$\left\|\sum A_j\right\| \leqslant \sum (c(j))^{1/2}.$$

PROOF OF LEMMA 2. Note that

$$\|T^*T \cdots T^*T\| = \|T\|^m,$$

if the product on the left contains m terms. Put $T = \sum_{j=1}^N A_j$ so that

$$\|T\|^m \leqslant \sum_{j_1, \ldots, j_m = 1}^{N} \left\|A_{j_1}^* A_{j_2} \cdots A_{j_{m-1}}^* A_{j_m}\right\|.$$

By hypothesis each term on the right satisfies

$$\left\|A_{j_1}^* \cdots A_{j_m}\right\| \leqslant C(j_1 - j_2)C(j_3 - j_4) \cdots C(j_{m-1} - j_m),$$

$$\left\|A_{j_1}^* \cdots A_{j_m}\right\| \leqslant BC(j_2 - j_3)C(j_4 - j_5) \cdots C(j_{m-2} - j_{m-1})B$$

so that

$$\left\|A_{j_1}^* \cdots A_{j_m}\right\| \leqslant B\left[C(j_1 - j_2)C(j_2 - j_3) \cdots C(j_{m-1} - j_m)\right]^{1/2}.$$

Therefore

$$\|T\|^m \leqslant NB\left[\sum (c(j))^{1/2}\right]^{m-1},$$

$$\|T\| \leqslant (NB)^{1/m}\left[\sum (c(j))^{1/2}\right]^{1-(1/m)}.$$

Letting $m \to \infty$ gives $\|T\| \leqslant \Sigma(c(j))^{1/2}$. Since the right side is independent of N, we can let $N \to \infty$ and conclude

$$\left\|\sum A_j\right\| \leqslant \sum (c(j))^{1/2}. \quad \square$$

COROLLARY. *Instead of* (25), *assume*

$$\text{(25')} \qquad \left\|A_j^* A_k\right\| \leqslant C(j, k), \qquad \left\|A_j A_k^*\right\| \leqslant C(j, k),$$

with $\sup_j \Sigma_k \sqrt{C(j, k)} < \infty$. *Then*

$$\left\|\sum_j A_j\right\| \leqslant \max_j \sum_k \sqrt{C(j, k)}\,.$$

PROOF. The same argument with trivial modifications works. $\square$

All that remains is to verify the estimates (25′). Now

$$\left\langle p_j(y, D)g, f\right\rangle = \int e^{iy\cdot\eta} p_j(y, \eta)\hat{g}(\eta)\bar{f}(y)\, d\eta dy$$

$$= \int e^{i(y-x)\cdot\eta} p_j(y, \eta)g(x)\bar{f}(y)\, dx d\eta dy.$$

Therefore $p_j(x, D)$ has an adjoint $p_j^*(x, D)$

$$\text{(26)} \qquad p_j^*(x, D)f(x) = \int e^{i(x-y)\cdot\eta}\, \overline{p(y, \eta)}\, f(y)\, d\eta dy,$$

and

$$p_j^* p_k f(x) = \int e^{i(x-y)\cdot\eta}\, \overline{p_j(y,\eta)}\, e^{i\xi\cdot y} p_k(y,\xi) \hat{f}(\xi) d\xi d\eta dy$$

$$= \int e^{ix\cdot\eta}\Big[e^{i(\xi-\eta)\cdot y}\, \overline{p_j(y,\eta)}\, p_k(y,\xi) \hat{f}(\xi)\, d\xi dy\Big]\, d\eta,$$

so that by Plancherel's theorem we need only estimate

$$\int e^{i(\xi-\eta)\cdot y}\, \overline{p_j(y,\eta)}\, p_k(y,\xi) \hat{f}(\xi)\, d\xi dy. \tag{27}$$

There are three cases.

Case (i). Suppose that the y-support of Q_j^* is distinct from the y-support of Q_k^*. Then the support of $p_j(y,\eta)$ and $p_k(y,\xi)$ are distinct and (27) vanishes.

Case (ii). Consider all those Q_k^* that intersect a fixed Q_j^* and consider the corresponding $p_k(y,\xi)$ and $\overline{p_j(y,\eta)}$. By (21.ii) there are at most $3^{2n}-1$ such $p_k(y,\xi)$. Moreover

$$\|p_j^*(y,D)p_k(y,D)\|_2 \leqslant \|p_j^*(y,D)\|_2 \cdot \|p_k(y,D)\|_2,$$

and so the contribution to the sum $\Sigma(c(j,k))^{1/2}$ is bounded.

Case (iii). Assume that $\overline{p_j(y,\eta)}$ and $p_k(y,\xi)$ are supported in blocks which share some y-support and whose centers satisfy $|\eta_j| \approx M$ and $|\xi_k| \approx N$. Now

$$\frac{(-\Delta_y)^s e^{iy\cdot(\xi-\eta)}}{(|\xi-\eta|)^{2s}} = e^{iy\cdot(\xi-\eta)},$$

while from (23) we have

$$(-\Delta_y)^s\big[\overline{p_j(y,\eta)}\, p_k(y,\xi)\big] \leqslant (M+N)^s;$$

therefore after integrating (27) by parts we find

$$\begin{aligned}
\|p_j^* p_k f\|_2 &= \left\| \int e^{i(\xi-\eta)\cdot y}\, \overline{p_j(y,\eta)}\, p_k(y,\xi)\hat{f}(\xi)\, d\xi dy \right\|_2 \\
&= \left\| \int (|\xi-\eta|)^{-2s} e^{i(\xi-\eta)\cdot y} \hat{f}(\xi)(-\Delta_y)^s\big[\overline{p_j(y,\eta)}\, p_k(y,\xi)\big]\, dy d\xi \right\|_2 \\
&\leqslant C_s(|\eta_j - \xi_k|)^{-2s} \cdot \|f\|_2 \cdot (M+N)^s \\
&\leqslant \frac{C_s}{M^s+N^s} \cdot \|f\|_2 .
\end{aligned}$$

This shows that we can essentially neglect the contribution to the sum $\Sigma(c(j,k))^{1/2}$ arising from these blocks. We have estimated $p_j^*(y,D)p_k(y,D)$; a similar argument yields an estimate for $p_j(y,D)p_k^*(y,D)$.

We have shown that the sum $\Sigma(c(j,k))^{1/2}$ is bounded. By the corollary to Lemma 2, the operator $p(x,D) = \Sigma_j p_j(x,D)$ is bounded on L^2. This finishes the proof of Lemma 1. □

PROOF OF THEOREM 3. By Lemma 1, we can assume $p(x, \xi)$ is supported on a block Q in (x, ξ)-space of diameter $1 \times M$. We will rerun the proof of Lemma 1, this time using a finer localization. Let $\{Q_\nu\}$ be a partition of Q into blocks with

$$\operatorname{diam}_x Q_\nu = M^{-1/2}, \qquad \operatorname{diam}_\xi Q_\nu = M^{1/2},$$

and let $1 = \Sigma\phi_\nu(x, \xi)$ be a partition of unity with

$$\phi_\nu(x, \xi) = 1 \quad \text{near the center of } Q_\nu,$$
$$\operatorname{supp} \phi_\nu \subset Q_\nu^*.$$

As before, put

$$p(x, D) = \Sigma p_\nu(x, D) \quad \text{where } p_\nu(x, \xi) \equiv \phi_\nu(x, \xi) p(x, \xi).$$

We must now verify the hypotheses of Cotlar's lemma. We begin by showing that each of the $p_\nu(x, D)$ is bounded on L^2. Suppose that $p_0(x, \xi) \in C^\infty$ and is supported in $|x| < 10$, $|\xi| < 10$ so that $|p_0(x, D)u(x)| \leq C\|u\|_2 \chi_{|x|<10}$, where χ is the characteristic function. By integrating we get

$$\|p_0(x, D)u\|_2 \leq C\|u\|_2. \tag{28}$$

Let (x_ν, ξ_ν) be the center of Q_ν. Using the change of variables

$$(x, \xi) \to (y, \eta) \equiv \left(M^{1/2}(x - x_\nu), M^{-1/2}(\xi - \xi_\nu)\right), \tag{29}$$

we can change the support of $p_\nu(x, \xi)$ to $|y| < 10$, $|\eta| < 10$ and apply estimate (28). This together with the fact that $p_\nu \in S^0_{1/2,1/2}$ uniformly in ν imply

$$\|p_\nu(x, D)u\|_2 \leq C\|u\|_2,$$

where C is independent of ν.

We must now estimate $\|p_j^*(x, D)p_k(x, D)\|_2$. Using formula (26) for the adjoint we have

$$\begin{aligned} p_j^* p_k u(x) &= \int e^{i(x-z)\cdot\zeta}\,\overline{p_j(z, \zeta)}\,[p_k(z, D)u(z)]\,dz d\zeta \\ &= \int e^{i[(x-z)\cdot\zeta + (z-y)\cdot\xi]}\,\overline{p_j(z, \zeta)}\,p_k(z, \xi)u(y)\,dy d\xi dz d\zeta. \end{aligned} \tag{30}$$

We proceed here just as we did at the end of Lemma 1. Let $(x_j, \xi_j), (x_k, \xi_k)$ be the centers of the cubes Q_j^*, Q_k^* supporting $\overline{p_j}$ and p_k. There are three cases.

Case (i). If $|x_j - x_k| \geq CM^{-1/2}$, then $\overline{p_j}p_k = 0$ and $\|p_j^* p_k\|_2 = 0$.

Case (ii). Assume $|x_j - x_k| < CM^{-1/2}$ but $|\xi_j - \xi_k| \geq CM^{1/2}$. We rewrite (30) as

$$\begin{aligned} p_j^* p_k u(x) &= \int e^{i[(x-z)\cdot\zeta + z\cdot\xi]}\,\overline{p_j(z, \zeta)}\,p_k(z, \xi)\hat{u}(\xi)\,d\xi dz d\zeta \\ &= \int e^{ix\cdot\xi} p_{jk}(x, \xi)\hat{u}(\xi)\,d\xi = p_{jk}(x, D)u(x), \end{aligned}$$

where

$$p_{jk}(x, \xi) = \int e^{i[(x-z)\cdot(\zeta-\xi)]}\,\overline{p_j(z, \zeta)}\,p_k(z, \xi)\,dz d\zeta. \tag{31}$$

To estimate $p_{jk}(x, \xi)$, we use the identities

$$(32.\mathrm{i}) \qquad -|\zeta - \xi|^{-2}\Delta_z e^{i[(x-z)\cdot(\zeta-\xi)]} = e^{i[(x-z)\cdot(\zeta-\xi)]},$$

$$(32.\mathrm{ii}) \qquad -|x - z|^{-2}\Delta_\zeta e^{i[(x-z)\cdot(\zeta-\xi)]} = e^{i[(x-z)\cdot(\zeta-\xi)]}.$$

To use these, note that $|\zeta - \xi| \neq 0$ in the support of the integrand of (31); and for $|x - x_j| \geq CM^{-1/2}$, we have also that $|x - z| \neq 0$ in the support of the integrand of (31).

If we substitute the expressions (32.i) and (32.ii) for the exponential in (31) and integrate by parts repeatedly, we obtain

$$(33.\mathrm{i}) \quad p_{jk}(x, \xi) = \int e^{i[(x-z)\cdot(\zeta-\xi)]}\left(-|\zeta - \xi|^{-2}\Delta_z\right)^s\left[\overline{p_j(z, \zeta)}\, p_k(z, \xi)\right] dz d\zeta,$$

$$(33.\mathrm{ii}) \quad p_{jk}(x, \xi) = \int e^{i[(x-z)\cdot(\zeta-\xi)]}\left(-|x - z|^{-2}\Delta_\zeta\right)^s\left[\overline{p_j(z, \zeta)}\, p_k(z, \xi)\right] dz d\zeta.$$

Equation (33.i) holds for all (x, ξ) and (33.ii) holds for $|x - x_j| \geq CM^{-1/2}$.

Since the integrands are supported on the cube Q_j^* of volume ~ 1, and since $|\xi - \zeta| \sim |\xi_j - \xi_k|$ in (33.i) and $|x - z| \sim |x - x_j|$ in (33.ii), our estimates on the derivatives of $p_j(z, \zeta)$, $p_k(z, \xi)$ now yield

$$|p_{jk}(x, \xi)| \leq C_s\left(\frac{M^{1/2}}{|\xi_j - \xi_k|}\right)^{2s} \quad \text{for all } (x, \xi),$$

$$|p_{jk}(x, \xi)| \leq C_s\left(\frac{M^{-1/2}}{|x - x_j|}\right)^{2s} \quad \text{for all } |x - x_j| \geq CM^{-1/2}.$$

Therefore

$$|p_{jk}(x, \xi)| \leq C_s\left(M^{-1/2}|\xi_j - \xi_k| + M^{1/2}|x - x_j|\right)^{-2s} \quad \text{for all } (x, \xi)$$

while a glance at the definition shows that $p_{jk}(x, \xi)$ is supported in the projection Q_k^ξ of Q_k^* onto ξ-space. So we can estimate

$$\begin{aligned}|p_j^*(x, D)p_k(x, D)u(x)| &\leq \left|\int e^{ix\cdot\xi}p_{jk}(x, \xi)\hat{u}(\xi)\, d\xi\right| \\ &\leq C_s\left(M^{-1/2}|\xi_j - \xi_k| + M^{1/2}|x - x_j|\right)^{-2s}\int_{Q_k^\xi}|\hat{u}(\xi)|\, d\xi,\end{aligned}$$

from which we get immediately

$$\|p_j^*(x, D)p_k(x, D)u\|_2 \leq C_s\left(M^{-1/2}|\xi_j - \xi_k|\right)^{-s}\|u\|_2;$$

in other words

$$\|p_j^*(x, D)p_k(x, D)\|_2 \leq C_s\left(M^{-1/2}|\xi_j - \xi_k|\right)^{-s}.$$

Similarly, we find

$$\|p_j(x, D)p_k^*(x, D)\|_2 \leq C_s\left(M^{-1/2}|\xi_j - \xi_k|\right)^{-s}.$$

Case (iii). The terms arising when $|x_j - x_k| < CM^{-1/2}$ and $|\xi_j - \xi_k| < CM^{-1/2}$ contribute only a finite amount to the sum appearing in the Cotlar-Stein lemma and need not be estimated.

We have considered all the cases and can now invoke the Cotler-Stein lemma to conclude that $p(x, D) = \Sigma p_\nu(x, D)$ is bounded on L^2. This proves Theorem 3. $\square$

We have just shown that pseudodifferential operators arising from 0th order symbols are bounded on L^2; on the other hand, in general, mth order symbols do not yield operators that are bounded on L^2. But consider the space

$$H^k(\mathbf{R}^n) = \{u \in L^2(\mathbf{R}^n) \colon D^\alpha u \in L^2(\mathbf{R}^n), \text{ for } |\alpha| \leqslant k\},$$

for k a positive integer. Using Theorem 3 it is not hard to show that $p(x, \xi) \in S^m$ implies

$$p(x, D) \colon H^k(\mathbf{R}^n) \to H^{k-m}(\mathbf{R}^n)$$

is bounded. See [**61**, p. 51], for a proof.

Note that by the Plancherel theorem

$$D^\alpha u \in L^2(\mathbf{R}^n), \qquad |\alpha| \leqslant k \Leftrightarrow \xi^\alpha \hat{u}(\xi) \in L^2(\mathbf{R}^n), \qquad |\alpha| \leqslant k,$$

or equivalently,

$$\Leftrightarrow (1 + |\xi|)^k \hat{u}(\xi) \in L^2(\mathbf{R}^n).$$

Therefore we can generalize the spaces $H^k(\mathbf{R}^n)$ by defining for $s \in \mathbf{R}$,

$$H^s(\mathbf{R}^n) = \left\{ \begin{array}{ll} \text{tempered distributions } u\colon & \text{(i) } \hat{u} \in L^2_{\text{loc}}(\mathbf{R}^n) \\ & \text{(ii) } (1 + |\xi|)^s \hat{u} \in L^2(\mathbf{R}^n) \end{array} \right\}$$

and

$$\|u\|_{H^s} = \left[\int \left(1 + |\xi|^2\right)^s |\hat{u}(\xi)|^2 \, d\xi \right]^{1/2}.$$

These are the *Sobolev spaces*.

4. Stationary phase and the calculus of ψDOs. In this section we will have to deal with integrals of the form

$$F_M(x, \eta) = \int e^{iM(x-y)\cdot(\eta-\xi)} \psi(x, y, \xi, \eta) \, dy d\xi. \tag{34}$$

This integral converges if ψ has compact support. The estimates in Lemma 4 below allow us to pass to the limit and interpret the integral for a general symbol ψ.

We now prove the important composition formula.

THEOREM 4. *If $a(x, \xi) \in S^m$ and $b(x, \xi) \in S^n$, then $a(x, D) \circ b(x, D)$ is a pseudodifferential operator, whose symbol $a \circ b$ has an asymptotic expansion*

$$a \circ b \sim \sum_{0 \leqslant |\alpha|} \frac{1}{\alpha!} \left(\frac{1}{i} \partial_\xi \right)^\alpha a (\partial_x)^\alpha b.$$

More precisely, we have for each N the estimate

$$a \circ b - \sum_{0 \leqslant |\alpha| < N} \frac{1}{\alpha!} \left(\frac{1}{i} \partial_\xi \right)^\alpha a (\partial_x)^\alpha b \in S^{m+n-N}. \tag{35}$$

PROOF. Using a localization argument like that in Lemma 1, we can assume that $a(x, \xi)$ and $b(x, \xi)$ are supported in a ring $M \leqslant |\xi| \leqslant 2M$.

With

$$a(x, D)u(x) = \iint e^{i(x-y)\cdot\xi} a(x, \xi) u(y)\, dy d\xi,$$

$$b(y, D)u(y) = \iint e^{i\eta\cdot y} b(y, \eta) \hat{u}(\eta)\, d\eta,$$

the composition becomes

$$a(x, D) \circ b(x, D)u(x) = \int e^{i(x-y)\cdot\xi + y\cdot\eta} a(x, \xi) b(y, \eta) \hat{u}(\eta)\, d\eta dy d\xi$$

$$= \int e^{ix\cdot\eta} (a \circ b)(x, \eta) \hat{u}(\eta)\, d\eta,$$

where we have defined

$$(a \circ b)(x, \eta) \equiv \int e^{i(x-y)\cdot(\xi-\eta)} a(x, \xi) b(y, \eta)\, dy d\xi. \tag{36}$$

The integral (36) converges when $a(x, \xi)$ and $b(y, \eta)$ have compact support. As we remarked at the beginning of the section, we will be able to drop the assumption of compact support after we prove Lemma 4. According to the asymptotic expansion, the main contributions to the integral (36) occur when $(y, \xi) = (x, \eta)$.

This is an illustration of the principle of stationary phase, which holds more generally for integrals of the form

$$F(t) = \int e^{it\phi(z)} \psi(z)\, dz,$$

where $\phi(z) \in C^\infty(\mathbf{R}^n)$ and $\psi(z) \in C_0^\infty(\mathbf{R}^n)$. We are interested in the asymptotic behavior of $F(t)$ as $t \to \infty$. Using the identity

$$\frac{(\partial/\partial z_j) e^{it\phi(z)}}{it(\partial/\partial z_j)\phi(z)} = e^{it\phi(z)},$$

and integrating by parts k times gives

$$F(t) = \left(\frac{1}{it} \right)^k \int e^{it\phi(z)} \frac{\psi(z)}{\left[(\partial/\partial z_j)\phi(z) \right]^k} dz.$$

If $(\partial/\partial z_j)\phi(z) \neq 0$ on the support of ψ, then $F(t) \to 0$ very fast as $|t| \to \infty$. Breaking up $\psi = \Sigma_l \psi_l$ with a partition of unity so that supp ψ_l is small shows that we may neglect the region where $\partial_z \phi(z) \neq 0$. Therefore the only important points are those z such that $\partial_z \phi(z) = 0$. These are the *stationary points*.

We will use this principle in the form

LEMMA 4. *Let*

$$F_M(x, \eta) = \iint e^{iM(x-y)\cdot(\eta-\xi)}\psi(x, y, \xi, \eta)\, dyd\xi,$$

where ψ satisfies

(i) $\psi \in C_0^\infty$ *with support in* $\{|x - y| + |\eta - \xi| < 1\}$,

(ii) ψ *vanishes to order k' at the stationary point* $(y, \xi) = (x, \eta)$.

Then $\partial_{x,\eta}^\alpha F_M = O(M^{-k})$ for $|\alpha| \leq k$ as $M \to \infty$. Here k may be made arbitrarily large by taking k' large.

PROOF. Observe

$$\frac{\left[-\left(\partial_y^2 + \partial_\xi^2\right)\right]^s e^{iM(x-y)\cdot(\eta-\xi)}}{\left[M^2\left(|\eta - \xi|^2 + |x - y|^2\right)\right]^s} = e^{iM(x-y)\cdot(\eta-\xi)};$$

therefore after integrating by parts we get

$$F_M(x, \eta) = \int \frac{e^{iM(x-y)\cdot(\eta-\xi)}}{M^{2s}}\left[-\left(\partial_y^2 + \partial_\xi^2\right)\right]^s\left\{\frac{\psi(x, y, \xi, \eta)}{\left[|\eta - \xi|^2 + |x - y|^2\right]^s}\right\} dyd\xi.$$

If ψ vanishes to high enough order at $(x, \xi) = (y, \eta)$, then

$$\frac{\psi(x, y, \xi, \eta)}{\left[|\eta - \xi|^2 + |x - y|^2\right]^s} \in C^{2s},$$

so that

$$\left|\left(\partial_y^2 + \partial_\xi^2\right)^s\left\{\frac{\psi(x, y, \xi, \eta)}{\left[|\eta - \xi|^2 + |x - y|^2\right]^s}\right\}\right| \leq C_s.$$

Consequently, recalling that ψ is supported in $|y| \leq 1$, $|\xi| \leq 1$, we find

(37) $$|F_M(x, \eta)| \leq C_s/M^{2s}.$$

We can obtain similar estimates for $\partial_{x,\eta}^\alpha F_M$ by noting $\partial_{x,\eta}^\alpha F_M(x, \eta)$ is a sum of terms

$$\int e^{iM(x-y)\cdot(\xi-\eta)}M^\beta(x - y)^\gamma(\xi - \eta)^\delta\psi_\mu(x, y, \xi, \eta)\, dyd\xi,$$

with $\beta \leq |\alpha|$. These terms are

$$M^\beta \cdot \int e^{iM(x-y)\cdot(\xi-\eta)}\tilde{\psi}(x, y, \xi, \eta)\, dy\, d\xi,$$

with $\tilde{\psi} = (x - y)^\gamma(\xi - \eta)^\delta\psi_\mu$, so (37) applies. Therefore,

$$\left|\partial_{x,\eta}^\alpha F_M\right| \leq \frac{C_s'}{M^{2s}} \cdot M^{|\alpha|}.$$

Taking s large enough completes the proof. $\square$

We now return to the proof of the composition formula. Let $\theta \in C_0^\infty$ be a cutoff function

$$\theta(x, \xi, y, \eta) = \begin{cases} 1 & \text{for } |x + y| + M^{-1}|\xi - \eta| \leq 1, \\ 0 & \text{for } |x + y| + M^{-1}|\xi - \eta| \geq 2 \end{cases}$$

and consider the integral

$$I_\theta = \iint e^{i(x-y)\cdot(\xi-\eta)}\theta(x, \xi, y, \eta)a(x, \xi)\, b(y, \eta)\, dyd\xi.$$

This is the integral (36) with the cutoff function θ introduced so that Lemma 4 applies. It will turn out that putting in θ has negligible effect on the integral (36). By Taylor's theorem the product $a \cdot b$ satisfies

$$a(x, \xi)b(y, \eta) \sim \sum_{|\alpha|, |\beta| \geq 0} \left[\frac{1}{\alpha!}\partial_\eta^\alpha a(x, \eta)(\xi - \eta)^\alpha\right]\left[\frac{1}{\beta!}\partial_x^\beta b(x, \eta)(y - x)^\beta\right].$$

If we truncate this sum after sufficiently many terms, the remainder satisfies the hypotheses of Lemma 4 and hence may be neglected. Thus I_θ is essentially

$$(38) \quad \sum_{0 \leq |\alpha|, |\beta| \leq N} \frac{1}{\alpha!\beta!} \cdot \left[\int e^{i(x-y)\cdot(\xi-\eta)}\theta(x, \xi, y, \eta)(\xi - \eta)^\alpha(y - x)^\beta\, dyd\xi\right]\left(\partial_\eta^\alpha a \partial_x^\beta b\, |_{(x,\eta)}\right).$$

To evaluate this integral we can remove $\theta(x, \xi, y, \eta)$ since a simple stationary phase argument shows that the effect is only $O(M^{-s})$. This leaves us with a finite sum of integrals of the form

$$J = \iint e^{i(x-y)\cdot(\xi-\eta)}(\xi - \eta)^\alpha(y - x)^\beta\, dyd\xi,$$

which we interpret as

$$\lim_{\varepsilon_k, \delta_k \to 0^+} \iint e^{i(x-y)\cdot(\xi-\eta)}(\xi - \eta)^\alpha(y - x)^\beta e^{-\Sigma_k \varepsilon_k(x_k - y_k)^2 + \delta_k(\xi_k - \eta_k)^2}\, dyd\xi.$$

It is sufficient to evaluate the translated integral $\iint e^{iy\cdot\xi}y^\beta\xi^\alpha\, dyd\xi$. By analyzing the effect of dilations we see that the integral vanishes unless $\alpha = \beta$. Now the Dirac δ function satisfies

$$\left(\partial_y^k\right)\delta(y) = \int e^{iy\cdot\xi}(i\xi)^k\, dy,$$

therefore

$$\iint e^{iy\cdot\xi}y^k\xi^k\, dyd\xi = \frac{1}{(i)^k}\int y^k\left(\delta^{(k)}(y)\right) dy = \frac{1}{(i)^k}k!,$$

and hence $J = k!/(i)^k$. This establishes the correct coefficients in the asymptotic expansion (38) and finishes the proof. □

COROLLARY. *Let U be an open set and $u \in C^\infty(U)$. Then $p(x, D)u(x) \in C^\infty(U)$.*

PROOF. Let $\phi, \psi \in C_0^\infty$ with $\psi = 1$ on $\operatorname{supp} \phi$. It is enough to show that $p(x, D)u$ is smooth in $\operatorname{supp} \phi$ whenever u is smooth in a neighborhood of $\operatorname{supp}(\psi)$. We write $\phi(x)p(x, D)u = \phi(x)p(x, D)(\psi u) + \phi(x)p(x, D)\cdot(1 - \psi(x))u$. The first term on the right is obviously smooth since ψu is smooth and compactly supported. To handle the second term we note that $\phi(x)p(x, D)\cdot(1 - \psi(x))$ is a composition of three pseudodifferential operators whose symbols have no common support. The composition law (Theorem 4) for ψDOs shows that $\phi(x)p(x, D)(1 - \psi(x))$ is a smoothing operator ($\in S^{-\infty}$), so the second term must also be smooth. Therefore $\phi(x)p(x, D)u$ is smooth, and the proof is complete. □

Although pseudodifferential operators, unlike differential operators, do not necessarily decrease the support of functions on which they act, they do according to the corollary decrease the singular support of the functions on which they act, i.e., the complement of the open set on which the function is smooth. Such an operator is called *pseudolocal.*

5. Elliptic regularity. We begin by sketching a proof that elliptic operators have parametrices (Equation 16). A symbol $a(x, \xi) \in S^m$ is called elliptic if

$$c(1 + |\xi|)^m \leqslant |a(x, \xi)| \leqslant c'(1 + |\xi|)^m.$$

Step 1. Put $b_1(x, \xi) = \theta(\xi)/a(x, \xi)$, where $\theta(\xi)$ is a cutoff function that vanishes in a bounded set containing the zeros of $a(x, \xi)$ and is 1 elsewhere. It is easy to check that $b_1(x, \xi)$ satisfies the estimates required to make $b_1 \in S^{-m}$. By Theorem 4,

$$a(x, D) \circ b_1(x, D) = \mathrm{I} + e_1(x, D),$$

where I is the identity and $e_1(x, \xi) \in S^{-1}$.

Step 2. Put $b_2(x, \xi) = -\theta(\xi)e_1(x, \xi)/a(x, \xi) \in S^{-m-1}$ so that

$$a(x, D) \circ b_2(x, D) = -e_1(x, D) + e_2(x, D),$$

where $e_2(x, \xi) \in S^{-2}$.

Step k. In general put $b_k(x, \xi) = -\theta(\xi)e_{k-1}(x, \xi)/a(x, \xi) \in S^{-m-k+1}$ so that

$$a(x, D) \circ b_k(x, D) = -e_{k-1}(x, D) + e_k(x, D),$$

where $e_k(x, \xi) \in S^{-k}$

Formally setting $b(x, \xi) \equiv \Sigma_1^\infty b_k(x, \xi)$, we have

$$a(x, D) \circ b(x, D) = \mathrm{I} + e(x, D),$$

where $e(x, \xi) \in S^{-\infty} \equiv \bigcap_m S^m$. We need

LEMMA 5. *Given a formal sum* $\Sigma_{k=1}^\infty b_k$, *with* $b_k(x, \xi) \in S^{m-k+1}$, *there exists* $b \in S^m$ *such that*

$$b - \left(\sum_1^N b_k\right) \in S^{m-N}.$$

SKETCH OF PROOF. Set $b = \Sigma\chi_k(\xi)b_k$, where $b_k \in C^\infty$ and

$$\chi_k(\xi) = \begin{cases} 1, & |\xi| > R_k, \\ 0, & |\xi| < R_k/2. \end{cases}$$

If $R_k \nearrow \infty$ sufficiently rapidly, then $b(x, \xi) \in S^m$ as required. □

We have proved

THEOREM 5. *An elliptic symbol $a \in S^m$ has a parametrix $b \in S^{-m}$ satisfying*

$$a(x, D) \circ b(x, D) = \mathrm{I} + e(x, D), \qquad b(x, D) \circ a(x, D) = \mathrm{I} + \tilde{e}(x, D),$$

where $e(x, D)$, $\tilde{e}(x, D) \in S^{-\infty} \equiv \bigcap_m S^m$ (smoothing errors). (See Remark below.)

REMARK. Strictly speaking, we only checked

$$a(x, D) \circ b(x, D) = \mathrm{I} + e(x, D)$$

where $e \in S^{-\infty}$. Similarly, we can construct a symbol $\tilde{b}$ so that

$$\tilde{b}(x, D) \circ a(x, D) = \mathrm{I} + \tilde{e}(x, D)$$

with $\tilde{e} \in S^{-\infty}$. Modulo errors in $S^{-\infty}$ we have

$$\tilde{b} = \tilde{b} \circ (a \circ b) = (\tilde{b} \circ a) \circ b = b.$$

So b is both a left and right parametrix for a.

We will now discuss briefly a regularity theorem for elliptic systems of equations on compact manifolds. First we examine the invariance of pseudodifferential operators under change of variables. Let τ: $U_1 \to U_2$ be a diffeomorphism between two open sets in $\mathbf{R}^n$. Fix a symbol $a(x, \xi) \in S^m$ with compact support (with respect to x) in U_1. We want to define a symbol $b(x, \xi)$ in U_2 that arises from $a(x, \xi)$ under the action of τ. For f supported in U_2, define

$$f_\tau(x) \equiv f(\tau^{-1}x);$$

$f_\tau(x)$ is supported in U_1.

THEOREM 6. *There exists a symbol $b(x, \xi) \in S^m$ such that*
(i) $(a(x, D)f_\tau)_{\tau^{-1}}(x) = b(x, D)f(x)$,
(ii) $b(x, \xi) = a(\tau(x), [(\partial\tau/\partial x)^t]^{-1}\xi) + S^{m-1}$.

Here $[(\partial\tau/\partial x)^t]^{-1}$ is the inverse of the transpose of the Jacobian of τ. We will see in the next chapter that this is a special case of a theorem describing the composition of two Fourier integral operators. Although we could give a direct proof here, we will instead note that this theorem is a corollary of Theorem 4 in Chapter 4.

We can now define a pseudodifferential operator on a manifold. First use a partition of unity so that the symbol and the functions on which it operates are supported in a local coordinate neighborhood (U_α, ϕ_α). Then Theorem 6 applies and we see by letting τ be the change of variables $\phi_\beta \circ \varphi_\alpha^{-1}$ that the principal symbol is invariantly defined on the cotangent bundle.

Let $u_1, \ldots, u_p$ and $v_1, \ldots, v_p$ be functions defined on a manifold and consider the matrix equation

$$A \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},$$

where A is a matrix of pseudodifferential operators (a_{ij}) defined on the manifold and satisfying

(i) $a_{ij} \in S^m$,

(ii) the inverse matrix (a^{ij}) has entries in S^{-m}.

Condition (ii) amounts to saying that $\det(a_{ij})$ is an elliptic symbol of the appropriate order.

THEOREM 7.

(i) $\ker A$ *and* $\operatorname{coker} A$ *are finite dimensional and composed of smooth functions.*

(ii) $\operatorname{ran} A$ *is closed.*

(iii) $Au = v$ *has a solution* u *if* $v \perp \operatorname{coker} A$.

(iv) $Au = v$ *and* $v \in C^\infty \Rightarrow u \in C^\infty$.

SKETCH OF PROOF. Recall that an operator T is called Fredholm if there exists an operator S such that $ST - \mathrm{I}$ and $TS - \mathrm{I}$ are compact. We have the standard

Fact. If T is Fredholm, then

(i) $\ker T$ and $\operatorname{coker} T$ are finite dimensional.

(ii) $\operatorname{ran} T$ is closed.

By the assumptions (i) and (ii) above about A and by Theorem 5, we can find a matrix B of pseudodifferential operators such that $AB - \mathrm{I}$ and $BA - \mathrm{I}$ are smoothing. We claim $AB - \mathrm{I}$ and $BA - \mathrm{I}$ are compact. Indeed both these operators can be represented by a matrix of integral operators with smooth kernels. For the symbols of the operators are in the intersection $\bigcap_m S^{-m}$. In particular, the kernel is in L^2. Therefore the operator is Hilbert-Schmidt and $AB - \mathrm{I}$ and $BA - \mathrm{I}$ are compact.

This together with the fact above proves the first half of (i) and (ii). Since A has a closed range and since the closure of the range of A is the orthogonal complement of $\ker A^* = \operatorname{coker} A$, we get (iii). Now $(BA - \mathrm{I})u \in C^\infty$ and by hypothesis $Au \in C^\infty$. By the corollary to Theorem 4, $BAu \in C^\infty$, therefore $u \in C^\infty$. This gives (iv) and the fact that $u \in \ker A \Rightarrow u \in C^\infty$. To prove the rest of (i), just note that a similar argument shows $u \in \ker T^* \Rightarrow u \in C^\infty$. $\square$

6. Schauder theory. Consider the nonlinear equation

$$A\left(x, u, \partial u/\partial x_i, \partial^2 u/\partial x_i \partial x_j\right) = 0,$$

where

$$\text{(39)} \qquad \begin{aligned} &\text{(i)} \quad A(x, \omega, \omega_i, \omega_{ij}) \text{ is smooth,} \\ &\text{(ii)} \quad \left(\partial A/\partial \omega_{ij}\right) > 0 \text{ for all } x, \omega, \omega_i, \omega_{ij}. \end{aligned}$$

If $u \in C^2$ and $\partial^2 u/\partial x_i \partial x_j \in \Lambda(\alpha)$, we write $u \in C^{2+\alpha}$.

THEOREM 8. *If $u \in C^{2+\alpha}$ is a solution of* (39), *then $u \in C^\infty$ with $\|u\|_{C^k}$ bounded only by a quantity determined by A and $\|u\|_C^{2+\alpha}$.*

Here

$$\|u\|_{C^k} \equiv \sum_{l=0}^{k} \max_{|\alpha|=l} \|D^\alpha u\|_{C^0}, \qquad \|u\|_{C^{k+\alpha}} \equiv \|u\|_{C^k} + \max_{|\alpha|=l} \|D^\alpha u\|_{\Lambda(\alpha)},$$

where $\|u\|_{C^0(\Omega)} = \sup_{\bar\Omega} |u(x)|$. We stress that this theorem asserts neither the existence of a solution nor the existence of a $\|\cdot\|_{C^{2+\alpha}}$ bound on the solution, rather just the regularity of such a solution. In practice finding a $\|\cdot\|_{C^{2+\alpha}}$ bound on the solution is the most difficult part of the analysis.

We motivate the proof by differentiating (39):

$$0 = \frac{\partial}{\partial x_k}\left(A\left(x, u, \frac{\partial u}{\partial x_i}, \frac{\partial^2 u}{\partial x_i \partial x_j}\right)\right)$$
$$= \sum_{ij} A_{\omega_{ij}}(x, u, u_i, u_{ij}) \frac{\partial^3 u}{\partial x_i \partial x_j \partial x_k} + \text{lower order terms.}$$

The first term is the only one involving third order differentiations of u. The coefficients $a_{ij}(x) \equiv A_{\omega_{ij}}(x, u, u_i, u_{ij})$ of this term are not C^∞ but rather only in $\Lambda(\alpha)$. The proof consists of four steps. The first two steps prove results on linear PDE's with rough coefficients, that permit us to conclude that $\partial^3 u/\partial x_i \partial x_j \partial x_k \in \Lambda(\alpha)$. The third step is a bootstrapping argument that allows us to repeat this procedure and reach the desired conclusion. The final step removes one of the initial a priori assumptions.

Step 1. Assume that u is supported in $|x| \leq 3/2$ and is a C^∞ solution of

$$\left[\Delta + \sum b_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j}\right] u = f, \qquad |x| \leq 2,$$

where

(i) $f, b_{ij} \in C^\infty$,

(ii) $\|f\|_{\Lambda(\alpha)} \leq B$, $\|b_{ij}\|_{\Lambda(\alpha)} \leq \delta$, $\delta > 0$.

Then $\|\partial^2 u/\partial x_i \partial x_j\|_{\Lambda(\alpha)} \leq B'$, where B' depends on B and δ.

PROOF. By (9) we can rewrite (40) using the Riesz transforms as

$$\left[\mathrm{I} - \sum b_{ij}(x) R_i R_j\right] \Delta u = f.$$

By Theorem 2, $R_i R_j$ are bounded on $\Lambda(\alpha)$. Now $\Sigma b_{ij}(x) R_i R_j$ has a small norm as an operator on $\Lambda(\alpha)$; therefore $\mathrm{I} - \Sigma b_{ij}(x) R_i R_j$ is a small perturbation of the identity, and the required estimate follows from Theorem 1.

Note that by assuming

(iii) $\|u\|_{\Lambda(\alpha)} \leq B$, $\|\partial u/\partial x_i\|_{\Lambda(\alpha)} \leq B$,

we can drop the assumption that u be supported in $|x| \leq 3/2$. Indeed let θ be an appropriate cutoff function and consider

$$L(u\theta) = \theta f + \text{remainder.}$$

Assumption (iii) guarantees that remainder $\in \Lambda(\alpha)$ and satisfies the required bounds. $\square$

Step 2. Assume that u is a C^∞ solution of

$$\sum a_{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j} = f, \qquad (a_{ij}(x)) > \delta \mathrm{I},\, \delta > 0,$$

with

(i) $\|u\|_{\Lambda(\alpha)} \leqslant B$, $\|\partial u/\partial x_i\|_{\Lambda(\alpha)} \leqslant B$.

Also assume

(ii) f, $a_{ij} \in C^\infty$ in $|x| \leqslant 2$,

(iii) $\|a_{ij}\|_{\Lambda(\alpha)} \leqslant B$.

Then $\|\partial^2 u/\partial x_i \partial x_j\|_{\Lambda(\alpha)} \leqslant B'$, where B' depends only on B and δ.

Proof. Let $\mathfrak{B}$ be any ball of radius $\eta > 0$ such that $\mathfrak{B} \subset \{|x| \leqslant 2\}$. To simplify the notation, assume that the center of $\mathfrak{B}$ is 0. We will make an estimate on the inner half $\mathfrak{B}_0$ of $\mathfrak{B}$. Let $y = 2\eta^{-1}x$, so that $\mathfrak{B}_0 = \{|y| \leqslant 2\}$, and put $\tilde{u}(y) = u(\eta y/2)$, so that $\Sigma a_{ij}(\eta y/2)\partial^2\tilde{u}/\partial x_i \partial x_j = \eta^2 f(\eta y/2) = \tilde{f}(y)$. Now on $|y| \leqslant 2$ we have

$$\|\tilde{f}\|_{\Lambda(\alpha)} \leqslant \|f\|_{\Lambda(\alpha)}, \qquad \|a_{ij}(\eta y/2) - a_{ij}(0)\|_{\Lambda(\alpha)} \leqslant B\eta^\alpha.$$

After a linear change of coordinates (so that $(a_{ij}(0)) = \mathrm{I}$) we can take η small enough so that the result above applies and conclude

$$\left\|\frac{\partial^2 u}{\partial y_i \partial y_j}\right\|_{\Lambda(\alpha)} \leqslant B' \quad \text{on } |y| \leqslant 1,$$

or

$$\left\|\frac{\partial^2 u}{\partial x_i \partial x_j}\right\|_{\Lambda(\alpha)} \leqslant B'\eta^{-2\alpha} \quad \text{for } x \in \mathfrak{B}_0.$$

To finish the proof, cover $\{|x| \leqslant 1\}$ by these $\mathfrak{B}_0$'s. The conclusion then follows since the η are bounded below. $\square$

Step 3. Take $u \in C^\infty$ in $|x| \leqslant 2$ that satisfies the estimate

$$\|u\|_{C^{2+\alpha}} \leqslant B$$

and consider the equation

$$A\left(x, u, \frac{\partial u}{\partial x_i}, \frac{\partial^2 u}{\partial x_i \partial x_j}\right) = 0. \tag{40}$$

Differentiating (40) gives

$$0 = \sum_{ij}\left(A_{\omega_{ij}}\left(x, u, \frac{\partial u}{\partial x_i}, \frac{\partial^2 u}{\partial x_i \partial x_j}\right)\right)\frac{\partial^3 u}{\partial x_i \partial x_j \partial x_k} + \text{lower order terms},$$

where the lower order terms involve derivatives of u of order $\leqslant 2$. Put $G_{ij}(x) = A_{\omega_{ij}}(x, u, \partial u/\partial x_i, \partial^2 u/\partial x_i \partial x_j)$, and rewrite the equation above as

$$\sum_{ij} G_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}\left(\frac{\partial u}{\partial x_k}\right) = f',$$

where $f' \in \Lambda(\alpha)$, and, moreover, the bounds on f and G_{ij} depend only on the bounds for the original equation (3). By Step 1, $\partial u/\partial x_k \in C^{2+\alpha}$, or $u \in C^{3+\alpha}$ on a slightly smaller ball with the required bounds on its norm. We complete the bootstrap by using the induction on the degree k of differentiation: assume $\|u\|_{C^{k+\alpha}} \leqslant B$, where B depends on the original $C^{2+\alpha}$ bound. Let D be a differential monomial of order $k-1$ and apply D to (3)

$$0 = \sum_{ij} A_{\omega_{ij}}\left(x, u, \frac{\partial u}{\partial x_i}, \frac{\partial^2 u}{\partial x_i \partial x_j}\right)\frac{\partial^2}{\partial x_i \partial x_j}(Du) + \text{lower order terms},$$

where the lower terms involve differentiations of u of order $\leqslant k$. Just as above, we have $\sum_{ij} A_{ij}(x)(\partial^2/\partial x_i \partial x_j)(Du) = f$, with $\|f\|_{\Lambda(\alpha)} \leqslant \text{const } B$, or $\|Du\|_{C^{2+\alpha}} \leqslant B^1$, giving

$$\|u\|_{C^{k+1+\alpha}} \leqslant B^1,$$

as required. □

Step 4. It is not difficult to remove the a priori assumption that $u \in C^\infty$ from Step 2. To remove the assumption from Step 3 assume that $u \in C^{2+\alpha}$ is any solution of

$$A\left(x, u, \frac{\partial u}{\partial x_i}, \frac{\partial^2 u}{\partial x_i \partial x_j}\right) = 0.$$

Replace the derivatives in this equation by difference quotients

$$\begin{aligned} 0 &= \frac{A(u_{ij}(x+h)) - A(u_{ij}(x))}{h} \\ &= \frac{1}{h}\int_0^1 \frac{\partial}{\partial t}\left\{A(tu_{ij}(x+h) + (1-t)u_{ij}(x))\right\} dt \\ &= \sum_{ij}\left\{\int_0^1 A_{\omega_{ij}}(tu_{ij}(x+h) + (1-t)u_{ij}(x))\, dt\right\}\left(\frac{u_{ij}(x+h) - u_{ij}(x)}{h}\right) \\ &\quad + \text{other terms.} \end{aligned}$$

This is the only place where we require $(A_{\omega_{ij}}) > 0$; in fact we require only that $(A_{\omega_{ij}}) > 0$ near the solution u. Now apply the theory of elliptic PDE's with rough coefficients to these difference coefficients and conclude that $u \in C^{3+\alpha}$. Iterating this argument removes the a priori assumption for $A_{\omega_{ij}} = 0$; a similar argument works in general. This finishes Step 4 and completes the proof of the Schauder Theorem. □

7. Boundary value problems. So far we have studied elliptic equations locally on $\mathbf{R}^n$ and on manifolds without boundary. We now show how to use pseudodifferential operators to study elliptic equations with boundary conditions. To illustrate the ideas, we shall discuss the problem

$$Lu = 0 \quad \text{in } D \subseteq \mathbf{R}^n, \tag{41}$$

$$Xu = f \quad \text{on } \partial D, \tag{42}$$

where L is an elliptic second order partial differential operator, and X is a complex vector field.

Our plan is as follows. Set $u_+ = u$ in D; 0 outside D. For (41) to hold, we must have $Lu_+ = g$ on $\mathbf{R}^n$, where g is a distribution on $\mathbf{R}^n$ whose support lies in ∂D. In other words, our solution is given by

$$u_+ = L^{-1}g, \quad g \text{ a distribution supported on } \partial D. \tag{43}$$

Note that (41) now holds automatically whenever u_+ is defined by (43). We shall pick g so that (42) holds also. It turns out that finding such a g amounts to solving a pseudodifferential equation on the manifold ∂D. Once g is known, formula (43) gives the solution u to our boundary problem.

To carry out this plan, we first have to study what kind of distributions g arise by applying L to u_+. It is convenient to work in a coordinate system in which ∂D is straightened out. Thus, assume in local coordinates:

$$D = \{(x', x_n) \in \mathbf{R}^{n-1} \times \mathbf{R}^1 \mid x_n > 0\}, \tag{44}$$

$$L = -\sum_{jk} a_{jk}(x) \frac{\partial^2}{\partial x_j \partial x_k} + \cdots, \qquad (a_{jk}) > 0,$$

$$X = \sum_k b_k(x') \frac{\partial}{\partial x_k}.$$

In our coordinates, we have the distribution equations

$$\frac{\partial}{\partial x_j} u_+ = \left(\frac{\partial}{\partial x_j} u\right)_+ + u(x',0)\delta(x_n)\cdot\delta_{jk}, \tag{45}$$

$$\frac{\partial^2}{\partial x_j \partial x_k} u_+ = \left(\frac{\partial^2 u}{\partial x_j \partial x_k}\right)_+ \quad \text{if } j, k \neq n, \tag{46}$$

$$\frac{\partial^2}{\partial x_j \partial x_n} u_+ = \left(\frac{\partial^2 u}{\partial x_j \partial x_n}\right)_+ + \frac{\partial u}{\partial x_j}(x',0)\delta(x_n) \quad \text{if } j \neq n, \tag{47}$$

$$\frac{\partial^2}{\partial x_n^2} u_+ = \left(\frac{\partial^2 u}{\partial x_n^2}\right)_+ + \frac{\partial u}{\partial x_n}(x',0)\delta(x_n) + u(x',0)\delta'(x_n). \tag{48}$$

Here $\delta(t)$ is the Dirac delta, while δ_{jk} is the Kronecker delta. Equation (45) follows by differentiating $u_+ = u\chi_{\{x_n>0\}}$, while successive applications of (45) yield (46)–(48). Now equations (44)–(48) show that

$$Lu_+ = (Lu)_+ + g_0(x')\delta(x_n) + g_1(x')\delta'(x_n),$$

with $g_1 = -a_{nn}(x',0) \cdot u(x',0)$. In our application, we want $(Lu)_+ = 0$, so that

$$u_+ = L^{-1}\{g_0(x')\delta(x_n) + g_1(x')\delta'(x_n)\} \quad \text{on } R^n. \tag{49}$$

The two unknown functions g_0, g_1 will be determined by the boundary condition (49) and the consistency requirement that u_+ defined by (49) must vanish for $x_n < 0$.

To carry this out, we now have to compute $L^{-1}\{g_0(x')\delta(x_n) + g_1(x')\delta'(x_n)\}$ using the formula for L^{-1} as a pseudodifferential operator. The answer is most easily expressed if in addition to (44), we normalize our coordinate system to satisfy

$$a_{jn}(x',0) = 0 \quad \text{for } j < n. \tag{50}$$

A coordinate system satisfying only (44) may be transformed locally to satisfy both (44) and (50) by the simple change of variable $(x', x_n) \to (x' + x_n F(x'), x_n)$ for suitable F. With normalizations taken care of, we can now state the result of the calculations with (49).

PROPOSITION 1. *Let D, L, X be as in* (44), (50), *and define*

$$u_+ = L^{-1}\{g_0(x')\delta(x_n) + g_1(x')\delta'(x_n)\}.$$

If

$$\lim_{x_n \to 0-} u_+(x', x_n) = 0 \quad \textit{and} \quad \lim_{x_n \to 0+} Xu_+(x', x_n) = f(x_1),$$

then g_0, g_1 satisfy the pseudodifferential equations

$$g_0 = \tilde{p}(x', D')g_1 \quad \textit{with } \tilde{p}(x', \xi') = -a_{nn}^{-1/2}\Big(\sum_{j,k<n} a_{jk}\xi_j\xi_k\Big)^{1/2} \mod S^0 \tag{51}$$

$$p^+(x', D')g_1 = f \tag{52}$$

$$\textit{with } p^+(x', \xi') = \frac{b_n}{a_{nn}^{3/2}}\Big(\sum_{j,k<n} a_{jk}\xi_j\xi_k\Big)^{1/2} + i\sum_{k<n}\frac{b_k}{a_{nn}}\xi_k \mod S^0.$$

Here b_n and a_{jk} are to be evaluated at $(x',0)$.

The proposition tells us that our boundary-value problem (41), (42) is reduced to the pseudodifferential equation (51) on ∂D. In fact, once g_1 is obtained by solving (52), we can then use (51) to find g_0 and then substitute g_0, g_1 into (49) to obtain our solution u of (41), (42). Let us illustrate this procedure in a simple example. Suppose X has real coefficients and is everywhere transverse to the boundary ($b_n(x',0) \neq 0$). Then the symbol p^+ in (52) is elliptic, since already

$$\operatorname{Re} p^+ = b_n a_{nn}^{-3/2}\Big(\sum_{j,k<n} a_{jk}\xi_j\xi_k\Big)^{1/2} \mod S^0$$

is elliptic. Theorem 7 now applies to equation (52), so that we can read off as a consequence

Elliptic regularity of the Neumann problem. Equations (41), (42) have a solution for all f orthogonal to a finite-dimensional cokernel. The solution is unique modulo a finite-dimensional kernel. If f is smooth, then u is smooth up to the boundary. The kernel and cokernel both consist of smooth functions.

If X has complex coefficients, then equation (52) becomes quite hard. In general it is not understood. Significant regularity theorems for (52) have been obtained by Egorov [**17**, **19**, **20**], Hörmander [**36**] and Kohn [**40**]. Egorov and Hörmander imposed hypotheses insuring that the S^0 error in (52) may be regarded as a trivial perturbation. Kohn's work [**40**] deals with the $\bar{\partial}$-Neumann problem, a system for which the first-order part of p^+ is positive semidefinite, and the zero-order correction term plays an essential role. We shall return later to discuss the $\bar{\partial}$-Neumann problem in its simplest case (strictly pseudoconvex domains).

PROOF OF THE PROPOSITION. We have to compute u_+ from formula (49). Now the main term in the symbol of the operator L^{-1} is $1/a(x, \xi)$ where $a(x, \xi) = \Sigma_{j,k} a_{jk}(x)\xi_j\xi_k$, so the main term in (49) is the integral

$$u_+(x', x_n) \approx \int_{R^n} \frac{e^{ix_n\xi_n} e^{ix'\cdot\xi'}}{a(x', x_n, \xi', \xi_n)} [\hat{g}_0(\xi') - i\xi_n \hat{g}_1(\xi')]\, d\xi' d\xi_n. \tag{53}$$

For fixed real x', x_n, ξ', the function $\xi_n \to 1/(a(x', x_n, \xi', \xi_n))$ is of course meromorphic with two simple poles $\xi_n = \tau_+(x', x_n, \xi')$ and $\xi_n = \tau_-(x', x_n, \xi')$ lying in the upper and lower half-planes respectively. Therefore, we can evaluate the ξ_n-integral in (53) by residues. The result for $x_n < 0$ is

$$u_+(x', x_n) \approx \int_{R^{n-1}} \frac{e^{ix_n\xi_-} e^{ix'\cdot\xi'}}{(\partial a/\partial \xi_n)(x', x_n, \xi', \xi_-)} [\hat{g}_0(\xi') - i\tau_- \hat{g}_1(\xi')]\, d\xi',$$

so that

$$\lim_{x_n \to 0-} u_+(x', x_n) \approx \int_{R^{n-1}} \frac{e^{ix'\cdot\xi'}}{(\partial a/\partial \xi_n)(x', 0, \xi', \tau_-)} \hat{g}_0(\xi')\, d\xi' + \int_{R^{n-1}} \frac{e^{ix'\cdot\xi'}(-i\tau_-)}{(\partial a/\partial \xi_n)(x', 0, \xi', \tau_-)} \hat{g}_1(\xi')\, d\xi'. \tag{54}$$

The integrals on the right in (54) are pseudodifferential operators applied to g_0, g_1. In particular, one checks easily that τ_- and $(\partial a/\partial \xi_n)(x', x_n, \xi', \tau_-)$ are first-order elliptic symbols for x_n near 0. (Recall that $a \sim a_{nn}\xi_n^2 + \Sigma_{j,k<n} a_{jk}\xi_j\xi_k$ by (50).)

Now the formulas (53), (54) are only approximate because we used only the principal symbol $1/(a(x, \xi))$ for L^{-1} in place of its full symbol. However, a glance at the recursive procedure for computing L^{-1} shows that the lower-order correction terms in the symbol for L^{-1} also continue analytically in ξ_n to meromorphic functions with poles only at $\tau_\pm$. Therefore a residue calculation like the one we carried out yields the corrected form of formula (54)

$$\lim_{x_n \to 0-} u_+ = p_0(x', D')g_0 + p_1(x', D')g_1 \tag{55}$$

where

$$p_0(x', \xi') = \frac{1}{(\partial a/\partial \xi_n)(x', 0, \xi', \tau_-)} \quad \bmod S^{-2},$$

$$p_1(x', \xi') = \frac{i\tau_-}{(\partial a/\partial \xi_n)(x', 0, \xi', \tau_-)} \quad \bmod S^{-1}.$$

Again note that p_0 is elliptic of order -1 while p_1 is elliptic of order 0. If $\lim_{x_n \to 0-} u_+(x', x_n) = 0$, then (55) shows that $p_0(x', D')g_0 = -p_1(x', D')g_1$, so that (since p_0 is elliptic) $g_0 = \tilde{p}(x', D')g_1$ with $\tilde{p} = -(p_1/p_0) \bmod S^0$. A computation with (44), (50) easily gives the formula (51) for $\tilde{p}$.

Next we want to calculate $\lim_{x_n \to 0+} Xu_+(x', x_n)$.

We can save part of the work by recalling that $g_1 = -a_{nn}(x', 0)u(x', 0)$. (See the discussion just before (49).) It follows at once that

$$\lim_{x_n \to 0+} \sum_{k<n} b_k(x') \frac{\partial u_+}{\partial x_k}(x', x_n) = \sum_{k<n} b_k(x') \frac{\partial}{\partial x_k}\left\{ \frac{-g_1(x')}{a_{nn}(x', 0)} \right\}.$$

Comparing this with the definition (44) of X, we see that it remains to find $\lim_{x_n \to 0+} (\partial/\partial x_n) u_+(x', x_n)$. To compute this, we differentiate (53) under the integral sign and again use residues to evaluate the ξ_n-integral. This time since $x_n > 0$, it is $\tau_+(x', x_n, \xi')$ that enters. Again the lower-order correction terms in L^{-1} are not important. The result is

$$\lim_{x_n \to 0+} \frac{\partial}{\partial x_n} u_+(x', x_n) = p_0^\#(x', D')g_0 + p_1^\#(x', D')g_1,$$

where $p_k^\#(x', \xi') = -(-i\tau_+)^{k+1}/(\partial a/\partial \xi_n)(x', 0, \xi', \tau_+) \bmod$ lower terms, $k = 0, 1$. Substituting (51) to eliminate g_0, we obtain a formula for

$$\lim_{x_n \to 0+} (\partial/\partial x_n)\, u_+(x', x_n)$$

as a pseudodifferential operator applied to g_1. Therefore

$$\lim_{x_n \to 0+} Xu_+(x', x_n) = b_n(x') \lim_{x_n \to 0+} \frac{\partial}{\partial x_n} u_+(x', x_n) + \lim_{x_n \to 0+} \sum_{k<n} b_k(x') \frac{\partial u_+}{\partial x_k}(x', x_n)$$

is expressed as a pseudodifferential operator applied to g_1. Carrying out the details using (50) to simplify formulas, we arrive at formula (52). □

REMARK. If L^{-1} is expressed as a (variable-coefficient) singular integral instead of a pseudodifferential operator, then it is very easy to read off the formula for $u_+(x', x_n)$, $x_n \neq 0$. We have no need of the residue calculations used above. The price we pay for this is that it takes some work to see what happens as $x_n \to 0\pm$, whereas the passage to the limit in the proof given above is immediate.

Finally, note that it makes no difference to replace (41, 42) by the seemingly more general problem

$$\begin{aligned} Lu &= f_1 \quad \text{in } D, \\ Xu &= f_2 \quad \text{on } \partial D. \end{aligned}$$

The reason is that we can extend f_1 to a convenient function $\tilde{f}_1$ on $\mathbf{R}^n$, solve $Lv = \tilde{f}_1$ in $\mathbf{R}^n$, and then observe that $(u - v)$ satisfies

$$\begin{aligned} L(u-v) &= 0 \quad \text{in } D, \\ X(u-v) &= f_2 - Xv \equiv f \quad \text{on } \partial D. \end{aligned}$$

CHAPTER 4. THE WAVE EQUATION

Let $(b_{ij}(x))$ be a positive definite, smoothly varying matrix and let

$$\Box = \frac{\partial^2}{\partial t^2} - \sum_{i,j=1}^{n} b_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j}$$

denote the wave operator. This chapter is concerned with the linear hyperbolic equation

$$(1) \qquad \Box u = 0, \qquad u\big|_{t=0} = f, \qquad \frac{\partial u}{\partial t}\bigg|_{t=0} = g.$$

In the last chapter we saw how ψDOs provided parametrices for elliptic operators; §1 of this chapter uses techniques from geometrical optics to construct a parametrix for the wave operator. This is an example of a Fourier Integral Operator (FIO). §2 sketches how this parametrix can be used to prove a theorem of Hörmander about the asymptotic distribution of the eigenvalues of selfadjoint elliptic operators with positive principal symbols. The next section defines (local) Fourier integral operators and amplifies the discussion in the third chapter on the principle of stationary phase. Egorov's theorem is proved in §4. This theorem gives a formula for the conjugation of a ψDO by an FIO. The calculus of FIOs is discussed in the next section. §6 completes the proof of Hörmander's theorem on eigenvalues, and the final section contains a few remarks about the global theory of FIOs and FIOs with complex phase. A good general reference for these topics is the book by Trèves [**63**].

1. A parametrix for $\Box$. Our calculations will be a bit simpler if we use

$$g^{ij}(x) = \begin{cases} 1, & i = j = 0, \\ -b_{ij}(x), & 1 \leqslant i, j \leqslant n, \\ 0, & \text{otherwise}, \end{cases}$$

so that

$$\Box = \sum_{i,j=0}^{n} g^{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j};$$

for convenience we will assume that the matrix $(g^{ij}(x))$ is symmetric. Consider first the constant coefficient wave equation $\partial^2/\partial t^2 - \partial^2/\partial x_1^2 - \cdots - \partial^2/\partial x_n^2$ and its plane wave solutions

$$u_\xi(x,t) = e^{i[x\cdot\xi \pm t|\xi|]}. \tag{2}$$

If the initial datum $f(x) = u(x,0)$ is suitable, then the Fourier decomposition $f(x) = \int_{\mathbf{R}^n} e^{ix\cdot\xi}\hat{f}(\xi)\,d\xi$ expresses f as a "sum" of plane waves $e^{ix\cdot\xi}$. Since $u_\xi(x,0) = e^{ix\cdot\xi}$, we can solve our original boundary value problem by superposition.

Turning now to the boundary value problem (1), we rewrite (2) as

$$e^{i\lambda[x\cdot\xi^0 \pm t|\xi^0|]},$$

where $\xi = \lambda(\xi/|\xi|) = \lambda\xi^0$. This suggests that we try to solve $\Box u = 0$ using a wave

$$u(x,t) = a(x,t,\lambda,\xi^0)e^{i\lambda S(x,t,\xi^0)}, \tag{3.i}$$

with

$$a(x,t,\lambda,\xi^0) = \sum_{k=0}^{\infty} a_k(x,t,\xi^0)\lambda^{-k}. \tag{3.ii}$$

We require

$$a(x,0,\lambda,\xi^0) = 1, \qquad S(x,0,\xi^0) = x\cdot\xi^0, \tag{3.iii}$$

so that (3) reduces to a plane wave at $t = 0$. It turns out that it is reasonable to assume that $a(x,t,\lambda,\xi^0) \in S^0$ so that it satisfies estimates of the form

$$\left|\partial_x^\alpha \partial_{\lambda\xi^0}^\beta \partial_t^\gamma a\right| \leqslant C_{\alpha\beta\gamma}\left(1 + |\lambda\xi^0|\right)^{-|\beta|}, \tag{4}$$

and to assume that

(5) (i) S is a smooth, real-valued function,
(ii) for $\xi^0 \neq 0$, $\nabla_x S$ is never 0 on the support of $a(x,t,\lambda,\xi^0)$.

Since

$$\frac{\partial}{\partial x_j}(ae^{i\lambda S}) = \frac{\partial a}{\partial x_j}e^{i\lambda S} + i\lambda\frac{\partial S}{\partial x_j}ae^{i\lambda S},$$

$$\frac{\partial^2}{\partial x_i \partial x_j}(ae^{i\lambda S}) = e^{i\lambda S}\left[\frac{\partial^2 a}{\partial x_i\partial x_j} + i\lambda\left(\frac{\partial a}{\partial x_j}\frac{\partial S}{\partial x_i} + \frac{\partial a}{\partial x_i}\frac{\partial S}{\partial x_j} + \frac{\partial^2 S}{\partial x_i \partial x_j}a\right) - \lambda^2\frac{\partial S}{\partial x_i}\frac{\partial S}{\partial x_j}a\right],$$

we have

$$\begin{aligned}\Box(ae^{i\lambda S}) &= \sum g^{ij}(x)\frac{\partial^2}{\partial x_i\partial x_j}(ae^{i\lambda S}) \\ &= e^{i\lambda S}\sum g^{ij}(x) \\ &\quad\times\left[\frac{\partial^2 a}{\partial x_i \partial x_j} + 2i\lambda\frac{\partial a}{\partial x_j}\frac{\partial S}{\partial x_i} + i\lambda\frac{\partial^2 S}{\partial x_i\partial x_j}a - \lambda^2\frac{\partial S}{\partial x_i}\frac{\partial S}{\partial x_j}a\right].\end{aligned} \tag{6}$$

We will choose $S(x, t, \xi^0)$ and $a_k(x, t, \xi^0)$ so that the right-hand side of (6) is 0. Beginning with the highest-order term, we require $\Sigma_{i,j=0}^n g^{ij}(x)(\partial S/\partial x_i)(\partial S/\partial x_j) = 0$, or in the earlier notation

$$\left(\frac{\partial S}{\partial t}\right)^2 - \sum_{i,j=1}^{n} b_{ij}(x)\frac{\partial S}{\partial x_i}\frac{\partial S}{\partial x_j} = 0,$$

which becomes

$$(7) \qquad \frac{\partial S}{\partial t} = \pm\left(\sum_{i,j=0}^{n} b_{ij}(x)\frac{\partial S}{\partial x_i}\frac{\partial S}{\partial x_j}\right)^{1/2} = H\left(x, \frac{\partial S}{\partial x_i}\right),$$

if we define $H(x, \zeta) = \pm(\Sigma b_{ij}(x)\zeta_i\zeta_j)^{1/2}$. To solve this equation, subject to the initial condition (3.iii), we choose, say, the positive sign and invoke Theorem 1 of Chapter 2. According to Theorem 1, the flow determined by the Hamiltonian $H(x, \zeta)$ yields for small time a family of canonical transformations Φ_t: $(y, \eta) \to (x, \zeta)$ whose generating functions $S_t(x, \eta)$ satisfies (7). Equation (7) is called the eikonal equation.

Having defined $S(x, t)$, we now ask that

$$\sum g^{ij}(x)\left(\frac{\partial^2 a}{\partial x_i \partial x_j} + 2i\lambda\frac{\partial a}{\partial x_j}\frac{\partial S}{\partial x_i} + i\lambda\frac{\partial^2 S}{\partial x_i\partial x_j}a\right) = 0.$$

Collecting terms of first order in λ and using the expansion (3.ii) of $a(x, t, \lambda, \xi^0)$ yields the next requirement

$$\sum\left(2g^{ij}(x)\frac{\partial S}{\partial x_i}\frac{\partial a_0}{\partial x_j} + g^{ij}(x)\frac{\partial^2 S}{\partial x_i\partial x_j}a_0\right) = 0.$$

This is a first-order linear partial differential equation for a_0 called the transport equation. Since at $t = 0$ (and hence for small time), $\Sigma g^{ij}(x)(\partial S/\partial x_i)(\partial/\partial x_j)$ is a nondegenerate vector field, we can solve the transport equation by integrating along the integral curves $\dot{x}_i = \Sigma_j g^{ij}(x)\partial S/\partial x_j$. We use the initial condition $a_0(x, 0, \xi^0) = 1$ so that (3.iii) is satisfied. Note, that except for a time change, the integral curves of $\dot{x}_i = \Sigma_j g^{ij}(x)\partial S/\partial x_j$ are the same as the integral curves determined by the Hamiltonian $H(x, \zeta)$ and agree with the geodesics in the metric in which $\Sigma b_{ij}\partial^2/\partial x_i \partial x_j + \cdots$ is the Laplacian.

In general, if we have found $a_0, \ldots, a_k$, collecting terms in (6) produces the requirement

$$(8) \quad \sum g^{ij}(x)\frac{\partial^2 a_k}{\partial x_i\partial x_j} + 2\sum g^{ij}(x)\frac{\partial S}{\partial x_i}\frac{\partial a_{k+1}}{\partial x_j} + \sum g^{ij}(x)\frac{\partial^2 S}{\partial x_i\partial x_j}a_{k+1} = 0.$$

This is a first-order linear ordinary differential equation for a_{k+1}. That is, (8) has the form $Xa_{k+1} + \phi a_{k+1} = b$ for known functions ϕ, b and a vector field X. It is called the transport equation and is solved just as above. For the initial conditions we can take $a_k(x, 0, \xi^0) = 0$, for $k = 2, 3, \ldots$.

THEOREM 1. *A solution of the wave equation* (1) *is given for small time by*

$$T^t u \equiv \int_{\mathbf{R}^n} \sum_{\pm} a_{\pm}(x, t, \xi) e^{iS_{\pm}(x,t,\xi)} \hat{f}(\xi)\, d\xi + \varepsilon f \tag{9}$$

$$+ \int_{\mathbf{R}^n} \sum_{\pm} a'_{\pm}(x, t, \xi) e^{iS_{\pm}(x,t,\xi)} \hat{g}(\xi)\, d\xi + \varepsilon' g,$$

where $a_{\pm}$, $a'_{\pm}$ *are symbols satisfying the estimates*

$$\left|\partial_x^\alpha \partial_\xi^\beta \partial_t^\gamma a_{\pm}\right| \leq C_{\alpha\beta\gamma}(1 + |\xi|)^{-|\beta|}, \qquad \left|\partial_x^\alpha \partial_\xi^\beta \partial_t^\gamma a'_{\pm}\right| \leq C_{\alpha\beta\gamma}(1 + |\xi|)^{-1-|\beta|},$$

S satisfies the requirements (5), *and* ε, ε' *are smoothing operators.*

PROOF. Recall that a smoothing operator is a ψDO $\varepsilon(x, D)$ with symbol $\varepsilon(x, \xi) \in S^{-\infty} \equiv \bigcap_m S^m$. First apply the construction above to the boundary value problem

$$\square u = 0, \qquad u\big|_{t=0} = f, \qquad \left.\frac{\partial u}{\partial t}\right|_{t=0} = 0,$$

by breaking f into plane waves using the Fourier decomposition and letting a_k be the functions determined by the transport equations. To achieve $\partial u/\partial t = 0$, we use linear combinations of our two solutions of $\square u = 0$, corresponding to $S_{\pm}$ which arise from the two choices of sign in the eikonal equation. By Lemma 5 of Chapter 3, there exists a symbol $a(x, \xi)$ such that formally $a - \sum_{k=0}^{\infty} a_k = \varepsilon$, where $\varepsilon \in S^{-\infty}$. See the lemma for the precise statement. The proof is completed by repeating this procedure for the boundary value problem

$$\square u = 0, \qquad u\big|_{t=0} = 0, \qquad \left.\frac{\partial u}{\partial t}\right|_{t=0} = g,$$

and adding the results. $\square$

As a corollary of this theorem we can prove a result about the propagation of singularities for the wave equation $\square u = 0$. Let Φ_τ be the Hamiltonian flow associated with the generating function $S_\tau(x, \xi)$ in (9).

THEOREM 2. *Assume that the initial data* $u\big|_{t=0}$, $\partial u/\partial t\big|_{t=0}$ *of* $\square u = 0$ *are supported at* $x = 0$. *Then the singularities of* $T^t u$ *lie on the light cone*

$$\{x \colon \exists \eta \text{ with } \Phi_\tau(0, \eta) = (x, \xi) \text{ for some } \xi\}.$$

We recall that the light cone is given by geodesics starting at x.

PROOF. Since up to a smooth error

$$T^t u = \int a(x, t, \eta) e^{iS(x,t,\eta)} \hat{u}(0, \eta)\, d\eta$$

$$+ \int a'(x, t, \eta) e^{iS(x,t,\eta)} \frac{\partial \hat{u}}{\partial t}(0, \eta)\, d\eta,$$

we know by the principle of stationary phase that the singular points of $T^t u$ occur where $\partial S/\partial \eta_j = 0$. In general, if $\Phi_t\colon (y, \eta) \to (x, \xi)$ is the Hamiltonian flow associated with the generating function $S_t(x, \eta)$, then $\partial S/\partial \eta_j = y_j$. So in our case, the singularities can occur only at those points x such that $(0, \eta) \to (x, \xi)$, for some η, ξ. But this is precisely the light cone. $\square$

2. Hörmander's Theorem. The operator $T^t u$ in (9) is an example of a Fourier integral operator. We have seen in Theorem 2 how this representation can be used to prove theorems about the propagation of singularities. In this section we will show how it can be used to prove a theorem about eigenvalue asymptotics. We will give a detailed proof of this theorem in §6 after we have developed some of the machinery necessary in order to study FIOs.

We begin with an example. Let M be the n-dimensional torus $S^1 \times \cdots \times S^1$ and $A = -\Delta = -\Sigma_{i=1}^n(\partial^2/\partial\theta_i^2)$ the Laplacian. The eigenvectors are $e^{im\cdot\theta}$, where $m = (m_1, \ldots, m_n) \in \mathbf{Z}^n$, and the eigenvalues are $|m|^2$. Let

$$N(\lambda) = \#\{\text{eigenvalues of } A < \lambda^2\}. \tag{10}$$

Now the number of integer lattice points $N'(\lambda)$ inside a ball B of radius λ is

$$N'(\lambda) = (\text{volume } B) + O(\lambda^{n-1}).$$

Indeed by approximating the ball from above and below by unions of unit cubes we see that $N'(\lambda) - (\text{volume } B)$ is bounded by the area of the boundary of B. Since volume $B = c_n\lambda^n$ we have

$$N(\lambda) = c_n\lambda^n + O(\lambda^{n-1}).$$

We will see later how to improve the error term.

This is illustrative of the general case. Let M be a compact n-dimensional manifold and $A = a(x, D)$ a mth-order elliptic, selfadjoint pseudodifferential operator whose principal symbol $a_m = \lim_{\lambda\to\infty}\lambda^{-m}a(x, \lambda\xi)$ is positive. Put

$$N(\lambda) = \#\{\text{eigenvalues of } A < \lambda\}. \tag{11}$$

THEOREM 3 (HÖRMANDER) [33].

$$N(\lambda) = c_{a,n}\lambda^{n/m} + O(\lambda^{(n-1)/m}), \tag{12}$$

where $c_{a,n} = c_n$ measure$\{(x, \xi) \in T^*M\colon a_m(x, \xi) < 1\}$.

Since $N(\lambda)$ in (10) counts the number of eigenvalues $< \lambda^2$, we see after replacing λ by λ^2 that the estimate above for the torus is a special case of (12). It should be noted that Theorem 3 is part of an extensive literature on eigenvalues going back to H. Weyl, and that the connection with the wave equation is also quite old. See [33] for detailed references. To simplify the discussion, we will assume in this section that $a(x, D)$ is first order. We begin by using the eigenfunctions of $a(x, D)$ to solve the "half-wave" equation

$$\left(\frac{1}{i}\frac{\partial}{\partial t} + A\right)u = 0, \qquad u\big|_{t=0} = f \tag{13}$$

on $M \times [0, T)$. Let $\psi_1, \psi_2, \ldots$ be the eigenfunctions of A and $\lambda_1, \lambda_2, \ldots$ the eigenvalues. If the initial data has an expansion $f = \Sigma c_k\psi_k$, then the solution of

(13) is $u(x, t) = \Sigma e^{i\lambda_k t} c_k \psi_k$. Formally the operator

$$U^t: f \to u(x, t) = \sum e^{i\lambda_k t} c_k \psi_k$$

is diagonal with entries $e^{i\lambda_k t}$ and its trace is $\Sigma e^{i\lambda_k t}$.

On the other hand the construction in §1 may be adopted to (13) to yield a parametrix

$$T^t: f \to u(x, t) = \int b(x, t, \eta) e^{iS(x,t,\eta)} \hat{f}(\eta)\, d\eta.$$

All the information about the distribution of the eigenvalues of A is contained in the measure

$$\mu = \sum \delta_{\lambda_k}, \tag{14}$$

where δ_λ is the delta function. Interpreted as a distribution the Fourier transform of μ is $\hat{\mu} = \Sigma_k e^{i\lambda_k t}$. We will study the asymptotic distribution of the eigenvalues of A by comparing U^t and T^t and calculating their traces (in an appropriate sense).

To be precise take $\phi \in C_0^\infty(\{|t| < t_0\})$ and consider the operator $f \to \int_{-\infty}^{\infty} \phi(t) T^t f\, dt$. It turns out that this operator is of trace class; so we may define a distribution F acting on test functions $\phi \in C_0^\infty$ by

$$\langle F, \phi \rangle \equiv \text{Trace} \int_{-\infty}^{\infty} \phi(t) T^t\, dt.$$

We calculate that $\langle F, \phi \rangle = \langle \hat{\mu}, \phi \rangle$ for ϕ supported in $(-t_0, t_0)$ and hence that $(\theta F)\hat{} = \mu * \hat{\theta}$, for a suitable cutoff function $\theta \in C_0^\infty$.

We now take advantage of the Fourier integral operator representation of the solution. Recall that when the integral operator $f(x) \to \int K(x, y) f(y)\, dy$ has a nice kernel $K(x, y)$, we have $\text{Trace} = \int K(x, x)\, dx$. In our case, for $\phi \in C_0^\infty(-t_0, t_0)$,

$$\left(\int \phi(t) T^t \right) (f(x)) = \iiint \phi(t) b(x, t, \eta) e^{i[S(x,t,\eta) - \eta \cdot y]} f(y)\, dt d\eta dy,$$

so that

$$\langle F, \phi \rangle = \iiint \phi(t) b(x, t, \eta) e^{i[S(x,t,\eta) - \eta \cdot x]}\, dt d\eta dx. \tag{15}$$

By the principle of stationary phase, the important contributions are the points (x, η) at which

$$\partial S/\partial \eta_j = x_j, \qquad \partial S/\partial x_j = \eta_j.$$

Note that these are the points which at time t have been carried back to themselves by the Hamiltonian flow associated with the generating function $S(x, t, \eta)$ in (15). After working a bit on (15), we find

$$\int_0^\lambda d[\hat{\theta} * \mu] = c_{a,n} \lambda^n + O(\lambda^{n-1}) \quad \text{as } \lambda \to \infty,$$

so that by applying a Tauberian theorem we get the desired estimate

$$N(\lambda) = \int_0^\lambda d\mu = c_{a,n} \lambda^n + O(\lambda^{n-1}).$$

3. Fourier integral operators and stationary phase. In this section we will have to work with operators of the form

$$Af(y) = \int e^{iS(y,\xi)} a(y,\xi) \hat{f}(\xi)\, d\xi$$
$$= \iint e^{i[S(y,\xi) - x\cdot\xi]} a(y,\xi) f(x)\, dx d\xi,$$

where

$$(16) \qquad \begin{array}{ll} \text{(i)} & a(y,\xi) \in S^m, \\ \text{(ii)} & f(x) \in C_0^\infty(\mathbf{R}^n), \\ \text{(iii)} & T(x,\xi; y) \equiv S(y,\xi) - x\cdot\xi \text{ is smooth, real valued,} \\ & \text{with } d_{x,\xi}T \neq 0, \text{ for any } (x,\xi). \end{array}$$

Condition (iii) implies that $T(x, \xi; y)$ does not have an extremum with respect to x and ξ, for any value of the parameter y; we will relax this condition shortly. Our first task is to give a precise meaning to the integral (16), which will not be absolutely convergent in general.

Let $\theta \in C_0^\infty(\mathbf{R}^n)$ be a cutoff function that is 1 near the origin and 0 elsewhere and put

$$a_\varepsilon(y,\xi) = a(y,\xi)\theta(\xi/\varepsilon),$$

so that the product $a_\varepsilon f$ has compact support in both x and ξ. If $m + n < 0$, then $Af(y)$ is absolutely convergent. If this is not the case, then let L be a first-order differential operator on $\mathbf{R}^n$ with smooth coefficients satisfying $LT/i = 1$. We can always find such an L locally: a partition of unity then produces a globally defined L. We have $Le^{iT} = e^{iT}$ and therefore integrating by parts formally k times gives

$$A_\varepsilon f(y) \equiv \iint e^{iT(x,\xi;y)} ({}^tL)^k [a_\varepsilon(y,\xi) f(x)]\, dx d\xi,$$

where ${}^tL$ is the transpose of L and maps ${}^tL$: $S^m \to S^{m-1}$. By taking k large enough, we get $m + n - k < 0$, so that the integral above becomes absolutely convergent uniformly in ε. In this case, $A_\varepsilon f$ has a limit Af as $a_\varepsilon(y,\xi) \to a(y,\xi)$, for $\varepsilon \to \infty$. We take this as the definition of the integral (16). Note that this procedure defines a map

$$C_0^\infty(\mathbf{R}^n) \to C^\infty(\mathbf{R}^n),$$
$$f \to Af,$$

for each fixed $a(y, \xi)$ and $T(x, \xi; y)$, which is a compactly supported distribution of order k. This is because at most k derivatives of f appear in the integral $A_\varepsilon f$. By taking duals we can extend this map to a map from compactly supported distributions to general distributions $\mathcal{E}'(\mathbf{R}^n) \to \mathcal{D}'(\mathbf{R}^n)$. In particular, we have defined Af for

$$(16.\text{ii})' \qquad f(x) \in C^\infty(\mathbf{R}^n).$$

This is called an *oscillatory integral*.

REMARK. This first order operator L involves dividing by the derivatives of T, which is where we used the hypothesis (16.iii).

We now relax the requirement (16.iii) and assume

(16.iii)′ $\quad T(x, \xi; y) \equiv S(y, \xi) - x \cdot \xi$ has a unique nondegenerate critical point $(x_0, \xi_0; y)$.

We repeat that the critical point is with respect to x and ξ; y is just a parameter. In §4 of Chapter 3 we saw how the principle of stationary phase applied to the integral

$$F(y, \eta) = \iint e^{iM(x-y)\cdot(\xi-\eta)} a(y, \xi) f(x)\, dx d\xi. \tag{17}$$

This is the integral appearing in Lemma 4 of Chapter 3 with $\psi(x, y, \xi, \eta) = a(y, \xi)f(x)$ and with x and y interchanged. The Morse lemma with parameters (which is our next proposition) will allow us to find the main term in an asymptotic expansion of the composition of two integral operators (16) by reducing it to this form.

PROPOSITION 1. *Suppose that $\phi_\lambda(z)$ is smooth and satisfies*

$$\phi_{\lambda_0}(0) = \phi'_{\lambda_0}(0) = 0,$$

$$\phi''_{\lambda_0}(0) = \begin{pmatrix} 1 & & & & & 0 \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ 0 & & & & & -1 \end{pmatrix}$$

where $z \in \mathbf{R}^n$, $\lambda \in \mathbf{R}^m$. Then near $z = 0$, $\lambda = \lambda_0$, there exists a smooth change of variables $z \to \tilde{z}$, where $\tilde{z}$ also depends smoothly on λ, such that

$$\phi_\lambda(\tilde{z}) = \tilde{z}_1^2 + \cdots + \tilde{z}_k^2 - \tilde{z}_{k+1}^2 - \cdots - \tilde{z}_n^2 + C(\lambda).$$

PROOF. Put $(z_1, \ldots, z_n) = (z_1, z')$ and let $f_\lambda(z')$ be the solution of the equation

$$\frac{\partial}{\partial z_1} \phi_\lambda(z_1, z') = 0,$$

with $z_1 = f_\lambda(z')$. By the implicit function theorem $f_\lambda(z')$ is smooth in λ and z'. After a change of variables $(z_1, z') \to (z_1 - f_\lambda(z'), z')$, we can assume $f_\lambda = 0$, i.e.,

$$\frac{\partial}{\partial z_1} \phi_\lambda(z_1, z') = 0 \quad \text{when } z_1 = 0.$$

Now by Taylor's theorem we can write

$$\phi_\lambda(z_1, z') = \phi_\lambda(0, z') + a_\lambda(z_1, z') z_1^2,$$

where $a_\lambda(z_1, z')$ is smooth and $a_\lambda(0,0) > 0$. Letting $\tilde{z}_1 = (a_\lambda(z_1, z'))^{1/2} z_1$, $\tilde{z}' = z'$ gives

$$\phi_\lambda(\tilde{z}) = \tilde{z}_1^2 + \tilde{\phi}_\lambda(\tilde{z}'),$$

allowing us to finish the proof inductively. □

In fact this proposition allows us to deal with integrals slightly more general than (16); but, before we define these integrals, we must introduce an appropriate class of symbols. Let $S^m(1 \times M \times 1 \times N)$ denote those smooth functions $a(x, \xi, y, \eta)$ which are supported on blocks of Q of dimension

$$\operatorname{diam}_x = 1, \qquad \operatorname{diam}_y = 1, \qquad \operatorname{diam}_\xi = M, \qquad \operatorname{diam}_\eta = N,$$

and which satisfy estimates of the form

$$\left|\partial_x^\alpha \partial_\xi^\beta \partial_y^\gamma \partial_\eta^\delta a(x, \xi, y, \eta)\right| \leq C_{\alpha\beta\gamma\delta} M^{m-|\beta|} N^{-|\delta|}, \tag{18}$$

for $(x, \xi, y, \eta) \in Q$. Note that this is a simple generalization of the estimates (23) in §3 of Chapter 3. The symbol class $S^m(1 \times M)$ is defined in the obvious way. Now consider integrals of the form

$$F(y, \eta) = \iint e^{iT(x,\xi,y,\eta)} a(x, \xi, y, \eta)\, dx d\xi, \tag{19}$$

where

$$\begin{array}{lll} & \text{(i)} & a(x, \xi, y, \eta) \in S^m(1 \times M \times 1 \times N), \\ & \text{(ii)} & T \text{ is real valued,} \\ (20) & \text{(iii)} & T \in S^1(1 \times M \times 1 \times N), \\ & \text{(iv)} & T(x, \xi, y, \eta) \text{ is a small perturbation of} \\ & & y \cdot \xi + x \cdot \eta - x \cdot \xi \text{ in the topology of } S^1(1 \times M \times 1 \times N). \end{array}$$

Condition (iv) means that there exist a small constant ε and a large constant E such that

$$\left|\partial_x^\alpha \partial_\xi^\beta \partial_y^\gamma \partial_\eta^\delta (T - (y \cdot \xi + x \cdot \eta - x \cdot \xi))\right| \leq \varepsilon M^{1-|\beta|} N^{-|\delta|},$$

for $|\alpha|, |\beta|, |\gamma|, |\delta| \leq E$.

PROPOSITION 2. *Assume that $T(x, \xi, y, \eta)$ and $a(x, \xi, y, \eta)$ satisfy* (20) *and that T has a unique nondegenerate critical point $(x_0(y, \eta), \xi_0(y, \eta), y, \eta)$, with respect to x and ξ; the critical point may depend on the parameters y and η. Then the main term in an asymptotic expansion of the integral* (19) *is*

$$(\textit{Jacobian factor}) \quad a(x_0(y, \eta), \xi_0(y, \eta), y, \eta) e^{iT(x_0, \xi_0, y, \eta)}. \tag{21}$$

PROOF. Putting $\lambda = (y, \eta/M)$, $z = (x, \xi/M)$ and applying Proposition 1 to (19) gives

$$F(y, \eta) = \iint e^{i(x - x_0(y,\eta)) \cdot (\xi - \xi_0(y,\eta))} \tilde{a}(x, \xi, y, \eta)\, dx d\xi.$$

The proof is completed by invoking Lemma 4 in §4 of Chapter 3. □

REMARK. In particular this proposition gives us the main term in an asymptotic expansion of (16) around a critical point satisfying (16.iii)$'$.

A *Fourier integral operator* is an operator

$$Af(y) = \int e^{iS(y,\xi)} a(y,\xi) \hat{f}(\xi)\, d\xi \tag{22}$$
$$= \iint e^{i[S(y,\xi) - x\cdot\xi]} a(y,\xi) f(x)\, dx d\xi$$

where

$$\begin{array}{ll} \text{(i)} & a(y,\xi) \in S^m(1 \times M), \\ \text{(ii)} & S(y,\xi) \text{ is real valued and homogeneous of degree 1} \\ & \text{in } \xi \text{ outside } |\xi| < 1, \\ \text{(iii)} & S(y,\xi) \in S^1(1 \times M), \\ \text{(iv)} & S(y,\xi) - y\cdot\xi \text{ is small in the topology of } S^1. \end{array} \tag{23}$$

As before (iv) means that there exists a small constant ε and a large constant E such that

$$\left|\partial_y^\alpha \partial_\xi^\beta (S(y,\xi) - y\cdot\xi)\right| \leqslant \varepsilon M^{1-|\beta|}, \tag{24}$$

for all $|\alpha|, |\beta| \leqslant E$. The function $a(y,\xi)$ is called the *amplitude* and $S(y,\xi)$ is called the *phase function*.

EXAMPLE 1. If $a(y,\xi) \in S^m$, then the pseudodifferential operator

$$a(y,D) f(y) = \int e^{iy\cdot\xi} a(y,\xi) \hat{f}(\xi)\, d\xi$$
$$= \iint e^{i[y\cdot\xi - x\cdot\xi]} a(y,\xi) f(x)\, dx d\xi$$

is clearly a Fourier integral operator.

EXAMPLE 2. Let Φ: $(x,\xi) \to (y,\eta)$ be a canonical transformation with generating function $S(y,\xi)$ and let $a(y,\xi) \in S^m$ be a symbol. Define the Fourier integral operator U^Φ by

$$U^\Phi f(y) = \int e^{iS(y,\xi)} a(y,\xi) \hat{f}(\xi)\, d\xi \tag{25}$$
$$= \iint e^{i[S(y,\xi) - x\cdot\xi]} a(y,\xi) f(x)\, dx d\xi.$$

A short calculation shows that $S(y,\xi)$ satisfies the requirements (23.ii), (23.iii), (23.iv).

EXAMPLE 3. Let $a(y,\xi)$, Φ and $S(y,\xi)$ have the same meaning as in the last example.

The adjoint operator of the Fourier integral operator (25) is given by

$$(U^\Phi)^* f(x) = \iint e^{-i[S(y,\xi) - x\cdot\xi]} \bar{a}(y,\xi) f(y)\, dy d\xi. \tag{26}$$

We shall see that $(U^\Phi)^*$ is a Fourier integral operator arising from the canonical transformation Φ^{-1}.

EXAMPLE 4. For fixed t, the parametrix for $\Box u = 0$ described in Theorem 1 is the sum of two Fourier integral operators, the first of which is

$$\begin{aligned} A^t f(y) &= \int e^{iS(y,t,\xi)} a(y, t, \xi) \hat{f}(\xi)\, d\xi \\ &= \iint e^{i[S(y,t,\xi) - x\cdot\xi]} a(y, t, \xi) f(x)\, dx d\xi. \end{aligned}$$

According to Theorem 1, $a(y, t, \xi) \in S^0$; since $S_t(y, \xi) = S(y, t, \xi)$ arises as a canonical transformation, it satisfies the necessary requirements by Example 2. The second Fourier integral operator is handled in the same way.

The next proposition clarifies the nature of the requirements (23) of the phase function $S(y, \xi)$.

PROPOSITION 3. *Conditions* (23.ii), (23.iii), (23.iv) *imply that* $S(y, \xi)$ *is the generating function for a canonical transformation* Φ: $(x, \xi) \to (y, \eta)$ *satisfying*

$$\text{(27)} \qquad \begin{aligned} &\text{(i)} \quad \Phi \text{ is homogeneous of degree } 1 \text{ in } \xi \quad \text{outside } |\xi| < 1, \\ &\qquad\quad \text{i.e., for } \lambda > 0,\ \Phi: (x, \lambda\xi) \to (y, \lambda\eta), \\ &\text{(ii)} \quad \text{for } |\xi| = 1,\ \Phi \text{ is a small perturbation of the identity.} \end{aligned}$$

PROOF. This follows from Proposition 4 of Chapter 2. $\Box$

Recall that as a corollary to Theorem 4 of Chapter 3 we proved that pseudodifferential operators decrease the singular support of functions on which they act. We conclude this section by extending this discussion to Fourier integral operators. We begin with an important refinement of the notion of singular support; this will also play a role in §8.

If u is a distribution on $\mathbf{R}^n$ and $(x_0, \xi_0) \in \mathbf{R}^n \times \mathbf{R}^n = T^*(\mathbf{R}^n)$ with $\xi_0 \neq 0$, then we call u regular at (x_0, ξ_0) if $\sigma(x, D)u \in C^\infty$ for some symbol $\sigma(x, \xi) \in S^0$ for which $\sigma(x_0, \lambda\xi_0)$ stays bounded away from zero as $\lambda \to \infty$. The *wave-front set* $\mathrm{WF}(u)$ consists of all (x_0, ξ_0) at which u is not regular. This refines the notion of singular support because $x^0 \in \operatorname{sing\,supp}(u)$ iff $(x^0, \xi^0) \in \mathrm{WF}(u)$ for some ξ_0. If $x^0 \in \operatorname{sing\,supp}(u)$, then for a cutoff function ϕ supported near x_0 we have $\phi u \notin C^\infty$; but $\mathrm{WF}(u)$ tells us the directions in which $(\phi u)\hat{}(\xi)$ fails to decrease rapidly. A Fourier integral operator, $Tf(y) = \int a(y, \xi) e^{iS(y,\xi)} \hat{f}(\xi)\, d\xi$, with $S =$ generating function of Φ, may be written as $Tf(y) = \int K(y, x) f(x)\, dx$ for a "Fourier integral distribution" K on $\mathbf{R}^n \times \mathbf{R}^n$.

Later on it will be useful to know the wave-front set of K. By stationary phase we can calculate $(\phi K)\hat{}$, and the answer is $\mathrm{WF}(K) = \{(y, \eta, x, \xi) \mid \Phi(x, \xi) = (y, \eta)\}$. In particular $\mathrm{WF}(Tf) \subset \Phi\mathrm{WF}(f)$, which strengthens and generalizes the corollary to Theorem 4 of Chapter 3.

4. Egorov's theorem. Changing variables $x \to y$ in the equation $p(x, D)u = f$ yields a new equation $\tilde{p}(y, D)\tilde{u} = \tilde{f}$ which may be easier to solve. Egorov's idea was to consider more generally canonical transformations Φ: $(x, \xi) \to (y, \eta)$ of the symbol $p(x, \xi)$ so that in the new variables $\tilde{p} = p \circ \Phi$ and to show that the equations $p(x, D)u = f$ and $\tilde{p}(y, D)\tilde{u} = \tilde{f}$ are still equivalent. For example, if the principal symbol $p(x, \xi)$ is real and has zeros in ξ of at most first order (in this case $p(x, \xi)$ is said to be of real principal type), then locally there is a canonical transformation such that $p(x, \xi) = \eta_1$. Thus equations of real principal type are reduced to

$$\frac{1}{i}\frac{\partial}{\partial x_1}u(x) = f(x).$$

Note that the wave equation is of real principal type.

THEOREM 4 (EGOROV) [**18**]. *Suppose $a(y, \xi) \in S^0$ is elliptic and that Φ is a canonical transformation satisfying*

$$(28)\qquad \begin{array}{ll} \text{(i)} & \Phi\colon (x, \lambda\xi) \to (y, \lambda\eta) \textit{ for } \lambda > 0, \\ \text{(ii)} & \textit{for } |\xi| = 1, \Phi \textit{ is a small perturbation of the identity.} \end{array}$$

Let U^Φ and $(U^\Phi)^*$ be the FIO associated with $a(y, \xi)$ and Φ by equations (23) *and* (24) *respectively. Let $p(y, \eta) \in S^m$ be a symbol. Then*

$$(29)\qquad \begin{array}{ll} \text{(i)} & \tilde{p}(x, D) \equiv (U^\Phi)^* \circ p(y, D) \circ (U^\Phi) \textit{ satisfies} \\ & \tilde{p}(x, \xi) \in S^m, \textit{ and} \\ \text{(ii)} & \textit{by choosing } a(x, \xi) \textit{ appropriately, we have} \\ & \tilde{p}(x, \xi) = p \circ \Phi(x, \xi) \bmod S^{m-1}. \end{array}$$

PROOF. By adapting the localization argument used in Lemma 1 of Chapter 3, we can assume

$$a(x, \xi) \in S^m(2 \times 2M), \qquad p(y, \eta) \in S^m(1 \times M).$$

From equation (26) and the definition of a pseudodifferential operator, we have

$$(U^\Phi)^*f(x) = \iint \bar{a}(y, \xi)e^{-i[S(y,\xi)-x\cdot\xi]}f(y)\,dyd\xi$$

and

$$p(y, D)g(y) = \iint p(y, \eta)e^{i\eta(y-y')}g(y')\,dy'd\eta;$$

therefore, since

$$U^\Phi f(y) = \int a(y, \xi)e^{iS(y,\xi)}\hat{f}(\xi)\,d\xi,$$

we have

$$(U^\Phi)^*p(y, D)U^\Phi f(x) = \int \bar{a}(y, \xi)e^{i[x\cdot\xi-S(y,\xi)]} \cdot p(y, \eta)e^{i\eta(y-y')}a(y', \xi')e^{iS(y',\xi')}\hat{f}(\xi')\,dyd\xi dy'd\eta d\xi',$$

where each of these integrals is interpreted as an oscillatory integral. Define $\tilde{p}(x, \xi')$ by

$$(U^\Phi)^* p(y, D) U^\Phi f(x) = \int \tilde{p}(x, \xi') e^{ix\cdot\xi} \hat{f}(\xi')\, d\xi',$$

so that

$$\tilde{p}(x, \xi') = \iiiint \bar{a}(y, \xi') a(y', \xi) p(y, \eta) e^{iT} dy d\xi' dy' d\eta, \tag{30}$$

where

$$T = x \cdot (\xi' - \xi) - S(y, \xi') + S(y', \xi) + \eta \cdot (y - y').$$

We will complete the proof by applying to equation (30) a stationary phase argument using Proposition 2. Unfortunately, Proposition 2, as it stands, does not apply; indeed, we must replace the symbol class $S^m(1 \times M \times 1 \times N)$ by the symbol class $S^m(1 \times M \times 1 \times N \times 1 \times P)$, which is defined in the obvious way. It is clear that $\bar{a}(y, \xi')a(y', \xi)p(y, \eta) \in S^m$ and that T is real valued. Since by hypothesis Φ is a small perturbation of the identity, $S(y, \xi')$ and $S(y', \xi)$ are small perturbations of $y \cdot \xi'$ and $y' \cdot \xi$, respectively, for these are the generating functions of the identity. This implies that $T \in S^1$ and that T is a small perturbation of

$$T_{\text{lin}} = x \cdot \xi - y \cdot \xi' + y' \cdot \xi - x' \cdot \xi + \eta \cdot y - \eta \cdot y'$$

in the topology of S^1. This completes the verification of the hypotheses of Proposition 2. Now we must locate the stationary points of T.

The equation $\nabla_{(y, \xi', y', \eta)} T = 0$ is equivalent to

$$\frac{\partial S}{\partial y_j}(y, \xi') = \eta_j, \qquad \frac{\partial S}{\partial \xi_j'}(y, \xi') = x_j,$$

$$\frac{\partial S}{\partial y_j'}(y', \xi) = \eta_j, \qquad y_j = y_j'.$$

Using the fact that S is a small perturbation of the generating function of the identity, we get

$$y = y', \qquad \xi = \xi', \tag{31.i}$$

$$\partial S / \partial y_j = \eta_j, \qquad \partial S / \partial \xi_j = x_j. \tag{31.ii}$$

Comparing (31.ii) and Proposition 4 of Chapter 2, we find that

$$\Phi(x, \xi) = (y, \eta), \tag{31.iii}$$

where Φ is the canonical transformation associated with S. We can now evaluate the integral (30) using Proposition 2 and conclude that the principal symbol of $\tilde{p}(x, \xi')$ is

$$(\text{Jacobian factor}) \quad (\bar{a}(y, \xi) a(y, \xi)) p(\Phi(x, \xi)).$$

This proves (29.i). Now choose $a(y, \xi) \in S^0$ so that

$$(\text{Jacobian factor}) \quad (\bar{a}(y, \xi) a(y, \xi)) = 1 \mod S^{m-1},$$

giving

$$\tilde{p}(x, \xi) = p \circ \Phi(x, \xi) \mod S^{m-1}. \quad \square$$

Note that instead of the hypotheses (28) in Egorov's theorem we could have simply stated the theorem for the Fourier integral operators defined by (22) and (23). Proposition 3 would then have given us a canonical transformation Φ satisfying (28). This was not done since we wish to emphasize that the canonical transformation can be chosen to suit the problem at hand. This is illustrated by the next theorem. Recall that a symbol $p(x, \xi)$ is of *principal type* if its principal symbol p_m satisfies for fixed x

$$p_m(x, \xi) = 0 \quad \text{and} \quad \xi \neq 0 \Rightarrow \nabla_\xi p_m(x, \xi) \neq 0. \tag{32}$$

THEOREM 5. *If $p(x, \xi)$ is real and of principal type, then locally there exists a canonical transformation Φ: $(x, \xi) \to (y, \eta)$ such that*

$$p(x, \xi) = \eta_1.$$

PROOF. See [**63**, vol. 2, p. 468]. $\square$

5. The composition formula. Fourier integral operators enjoy a calculus just as pseudodifferential operators do. In this section we prove a composition formula for Fourier integral operators corresponding to different phase functions. Consider

$$Au(y) = \int a(y, \xi) e^{iS(y,\xi)} \hat{u}(\xi)\, d\xi,$$

$$Bv(z) = \int b(z, \eta) e^{iT(z,\eta)} \hat{v}(\eta)\, d\eta,$$

where the amplitudes $a(y, \xi) \in S^m(1 \times M)$, $b(z, \eta) \in S^{m'}(1 \times M)$ and the phases $S(y, \xi)$, $T(z, \eta)$ satisfy the requirements (23). By Proposition 3, S and T are the generating functions of canonical transformations Φ and Ψ. Let Γ be the graph

$$\Gamma = \left\{ (x, \xi, y, \eta, z, \zeta) \colon \begin{array}{l} \Phi\colon (x, \xi) \to (y, \eta) \text{ and} \\ \Psi\colon (y, \eta) \to (z, \zeta) \end{array} \right\}$$

THEOREM 6. *$B \circ A$ is an FIO of the form*

$$(B \circ Au)(z) = \int c(z, \xi) e^{iS\square T(z,\xi)} \hat{u}(\xi)\, d\xi,$$

where

$$\begin{array}{ll} \text{(i)} & S \square T(z, \xi) \textit{ is the generating function for} \\ & \textit{the canonical transformation } \Psi \circ \Phi. \\ \text{(ii)} & c(z, \xi) \in S^{m+m'}. \\ \text{(iii)} & \text{Mod } S^{m+m'-1}, c(z, \xi) = \theta(z, \xi) b(z, \eta) a(y, \xi) \textit{ on } \Gamma, \\ & \textit{with } \theta(z, \xi) \textit{ an elliptic symbol, depending on} \\ & S, T \textit{ but not on } a, b. \end{array} \tag{33}$$

PROOF. We will argue formally first. We have

$$Bv(z) = \iint b(z,\eta) e^{i[T(z,\eta)-\eta\cdot y]} v(y)\, dy d\eta$$

and, therefore, putting $v = Au$ gives

$$(B \circ Au)(z) = \iiint b(z,\eta) a(y,\xi) e^{i[T(z,\eta)+S(y,\xi)-\eta\cdot y]} \hat{u}(\xi)\, d\xi d\eta dy. \tag{34}$$

To complete the calculation we will compute

$$F(z,\xi) = \iint b(z,\eta) a(y,\xi) e^{i[T(z,\eta)+S(y,\xi)-\eta\cdot y]}\, dy d\eta$$

using stationary phase. Put $R = T(z,\eta) + S(y,\xi) - \eta\cdot y$. The important points are the critical points, where

$$\partial R/\partial y = 0, \qquad \partial R/\partial \eta = 0,$$

or

$$\eta = \partial S(y,\xi)/\partial y, \qquad y = \partial T(z,\eta)/\partial \eta.$$

By Proposition 4 of Chapter 2, we see

$$\Phi\colon (x,\xi) \to (y,\eta), \qquad \Psi\colon (y,\eta) \to (z,\zeta),$$

showing that the stationary points coincide with the graph Γ. Proposition 5 of Chapter 2 tells us that R is the generating function of $\Psi \circ \Phi$. We can now apply Proposition 2 and conclude that $F(z,\xi)$ is asymptotic to

$$F(z,\xi) \sim \theta(z,\xi) b(z,\eta) a(y,\xi)\,|_{\Gamma} e^{iS \square T(z,\xi)},$$

and this gives

$$(B \circ Au)(z) = \int c(z,\xi) e^{iS\square T(z,\xi)} \hat{u}(\xi)\, d\xi,$$

with the amplitude and the phase satisfying (33).

To justify the application of Proposition 2, we can mimic the proof used to justify the use of stationary phase in Egorov's theorem. We have already tacitly assumed that the symbols have been localized and are elements of $S^m(1 \times M)$. As before, this follows from a modification of the localization argument in Lemma 1 of Chapter 3.

We must now check hypotheses (20) of Proposition 2. Clearly $b(z,\eta)a(y,\xi)$ is a symbol of the proper type and R is real valued. By (23.iii) we know $S, T \in S^1(1 \times M)$; therefore $R = T(z,\eta) + S(y,\xi) - \eta\cdot y \in S^1(1 \times M \times 1 \times M)$. Finally, since $S(y,\xi) - y\cdot\xi$ and $T(z,\eta) - z\cdot\eta$ are small in the topology of S^1, we see that R is a small perturbation of

$$R_{\text{linear}} = z\cdot\eta + y\cdot\xi - \eta\cdot y$$

in the topology of S^1. This verifies all the hypotheses of Proposition 2. All that remains is to check that the symbol $\theta(z,\xi)$ arising from the Jacobian factor is elliptic. This is a consequence of the proof of Proposition 2. $\square$

EXAMPLE 1. Let $T = y \cdot \xi$ so that B is a pseudodifferential operator. Then $B \circ A$ is a Fourier integral operator whose amplitude is multiplied by the symbol $b(y, \xi)$ of the pseudodifferential operator B.

EXAMPLE 2. Let U^{Φ} be the Fourier integral operator defined by the elliptic symbol $a \in S^0$ and the canonical transformation Φ according to (25). Thinking of the pseudodifferential operator $b(x, D)$ as a Fourier integral operator and applying the composition formula twice to

$$(U^{\Phi^{-1}})b(x, D)(U^{\Phi})$$

gives another proof of Egorov's theorem.

6. Proof of Hörmander's theorem. Let M be a compact n-dimensional manifold and $A = a(x, D)$ an mth order elliptic, selfadjoint operator whose principal symbol is positive. Recall that

$$N(\lambda) = \#\{\text{eigenvalues of } A < \lambda\}.$$

In this section we will prove

THEOREM 3 (HÖRMANDER). $N(\lambda) = c_{a,n}\lambda^{n/m} + O(\lambda^{(n-1)/m})$.

PROOF. Note that we will not actually compute the constant $c_{a,n}$. Assume to start with that A is first order. The main idea is to study the half-wave equation

$$\left(\frac{1}{i}\frac{\partial}{\partial t} + A\right)u = 0, \qquad u\,|_{t=0} = f \tag{35}$$

on $M \times \mathbf{R}^n$.

Step 0. In the spirit of §1 we shall write down a parametrix for (35) locally in $\mathbf{R}^n$. This is the heart of the analysis.

We look for a solution of (35) given by

$$u(t, y) = \int e^{iS(t,y,\xi)}\hat{F}_t(\xi)\,d\xi = U_tF_t \quad \text{with } F_t \text{ to be determined.}$$

The composition law for Fourier integral operators lets us compute $A(y, D_y)U_t$ and yields

$$\begin{aligned}\left[\frac{1}{i}\frac{\partial}{\partial t} + A(y, D_y)\right]u \\ = \int \frac{1}{i}\left\{\frac{\partial S}{\partial t} + A\left(y, \frac{\partial S}{\partial y}\right) + (\text{lower-order symbol})\right\}e^{iS(t,y,\xi)}\hat{F}_t(\xi)\,d\xi \\ + \int e^{iS(t,y,\xi)}\frac{1}{i}\frac{\partial \hat{F}_t}{\partial t}(\xi)\,d\xi.\end{aligned}$$

Therefore if $S(t, y, \xi)$ is taken to satisfy the Hamilton-Jacobi equation, we obtain

$$U_t^{-1}\left[\frac{1}{i}\frac{\partial}{\partial t} + A(y, D_y)\right]u = \left[\frac{1}{i}\frac{\partial}{\partial t} + B_t(x, D_x)\right]F_t$$

with $B_t(x, D)$ a zero-order pseudodifferential operator depending smoothly on t.

So $u = U_t F_t$ satisfies (35), provided F_t satisfies

$$\left[\frac{1}{i}\frac{\partial}{\partial t} + B_t(x, D_x)\right]F_t = 0, \qquad F\,|_{t=0} = f. \tag{35a}$$

However, it is easy to write down a parametrix for (35a). We just set $F_t = b(t, x, D_x)f$ where $b(t, x, \xi)$ is a symbol with the asymptotic expansion $b(t, x, \xi) \sim \sum_{\mu=0}^{\infty} b_\mu(t, x, \xi)$, $b_\mu \in S^{-\mu}$.

Equations (35a) hold modulo smoothing operators if the symbols b_μ satisfy

$$\left[\frac{1}{i}\frac{\partial}{\partial t} + B_t(x, \xi)\right]b_\mu = g_\mu(t, x, \xi), \qquad b_\mu\,|_{t=0} = \begin{cases} 1, & \mu = 0, \\ 0, & \mu > 0, \end{cases} \tag{35b}$$

where $g_\mu \in S^{-\mu}$ is determined by $b_{\mu'}$ ($\mu' < \mu$). Since $B_t \in S^0$, the integrating factor for the ordinary differential equation (35b) is elliptic in S^0. Thus (35b) can be solved with $b_\mu \in S^{-\mu}$, and we have our parametrix for (35a). The resulting parametrix $u = U_t F_t = U_t b(t, x, D_x)f$ for (35) has the form

$$u(t, y) = \int b(t, y, \xi) e^{iS(t,y,\xi)} \hat{f}(\xi)\, d\xi \quad \text{with } b \in S^0,$$

in view of the composition law for Fourier integral operators.

Step 1. We shall solve (35) on the manifold $M \times [-T, T]$ by patching together the local solutions derived in Step 0. Let $\{U_j\}$ be a cover of M by coordinate patches and let $\{\phi_j\}$ be a partition of unity subordinate to this cover. The construction of Step 0 yields a parametrix T^t for (35) in the following way. We break up the initial data f using the ϕ_j and let T_j^t be the parametrix constructed in Step 0 for

$$\left(\frac{1}{i}\frac{\partial}{\partial t} + a(x, D)\right)u = 0, \qquad u\,|_{t=0} = \phi_j f. \tag{36}$$

Put $u_j(x, t) = T_j^t \phi_j f(x)$. Observe that $u_j(x, t)$ is defined only for $x \in U_j$. We glue together the $u_j(x, t)$ by using a second partition of unity $\{\psi_j\}$ satisfying

(i) $\psi_j = 1$ on a neighborhood of supp ϕ_j,

(ii) $\operatorname{supp}(\psi_j) \subset U_j$

and by defining $u(x, t)$ by

$$u = \left(\sum_j \psi_j T_j^t \phi_j\right)(f).$$

Since $\operatorname{supp}(\phi_i f) \subseteq \operatorname{supp}(\phi_i)$, we have $\operatorname{sing\,supp}(T_j^t \phi_j f) \subseteq \{\psi_j = 1\}$, for $|t| \ll 1$; this means that the cutoff functions ψ_j introduce only C^∞ errors, and, therefore $(\sum_j \psi_j T_j^t \phi_j)$ is a parametrix for the half-wave equation (35) on $M \times [-T, T]$.

We now show that it is sufficient to prove the theorem for a first-order operator.

PROPOSITION 4. *Let $a(x, \xi) \in S^m$ be an elliptic symbol with positive principal part and with $a(x, D)$ selfadjoint. Then there exists a symbol $b(x, \xi) \in S^1$ satisfying*

(i) *$b(x, D)$ is selfadjoint,*

(ii) *$[b(x, D)]^m - a(x, D)$ is a bounded operator,*

(iii) *$b - a^{1/m} \in S^0$.*

PROOF. Since $a(x, \xi)$ is positive and elliptic, $b_1(x, \xi) \equiv a^{1/m}(x, \xi)$ satisfies

$$b_1^m - a \in S^{m-1},$$

with $b_1(x, \xi) \in S^1$. Now choose $b_0(x, \xi) \in S^0$ so that

$$(b_1 + b_0)^m - a \in S^{m-2}.$$

In general choose $b_{-k}(x, \xi) \in S^{-k}$ recursively so that

$$(b_1 + b_0 + \cdots + b_{-k})^m - a \in S^{m-(k+2)}.$$

By Lemma 5 of Chapter 3, there exists $b(x, \xi) \in S^1$ satisfying

$$b \sim b_1 + b_0 + b_{-1} + \cdots, \qquad [b(x, D)]^m - a = (\text{bounded operator}).$$

One checks that $b(x, D)$ is selfadjoint if $a(x, D)$ is. $\square$

This means that the eigenvalues of $[b(x, D)]^m$ and $a(x, D)$ differ by a bounded amount; therefore, the asymptotic behavior of the eigenvalues of these operators is the same. Note that even if we had started out with (35) a partial differential equation, this procedure would still lead to a pseudodifferential equation. In summary, instead of studying the eigenvalues of $a(x, D)$ by solving (35) for higher-order operators, we can solve (36) for $a(x, D)$ a first-order operator.

Step 2. Let T^t: $f \to u(\cdot, t)$ be the parametrix for (36) described above. To simplify the notation, we have denoted the initial data $\phi_i f$ by f. The operator maps f to the solution u and has the form

$$T^t f(x) = \int b(x, t, \eta) e^{i[S(x,t,\eta) - y\cdot\eta]} f(y)\, dy d\eta + \varepsilon_t f, \tag{37}$$

where ε_t is a smoothing operator $b(x, t, \eta) \in S^0$, and S_t is the generating function for the Hamiltonian flow Φ_t: $(y, \eta) \to (x, \xi)$, determined by the Hamiltonian $H(x, \xi) = a(x, \xi)$, a = symbol of A.

Let $\phi(t) \in C_0^\infty(-T, T)$ and consider the operator

$$T_\phi = \int_{-T}^{T} \phi(t) T^t\, dt.$$

We claim

$$\begin{aligned} &\text{(i)} && T_\phi f(y) = \int_M K(y, x) f(x)\, dx, \text{ with } K \in C^\infty(M \times M),\\ (38)\quad &\text{(ii)} && \operatorname{Trace} T_\phi = \int_M K(y, y)\, dy,\\ &\text{(iii)} && |\operatorname{Trace} T_\phi| \leq c \|\phi\|_{C^k}, \text{ for } k \text{ large},\\ &\text{(iv)} && \text{the map } F\colon \phi \to \operatorname{Trace} T_\phi \text{ is a distribution.}\end{aligned}$$

We begin with the verifiction of (i). From (37) we have

$$T_\phi f(x) = \int \phi(t) b(x, t, \eta) e^{i[S(x,t,\eta) - y\cdot\eta]} f(y)\, dy d\eta dt$$
$$+ \int \phi(t)(\varepsilon_t f)\, dt.$$

The kernel of the second integral is obviously smooth; the kernel of the first integral is

$$K(x, y) = \iint \phi(t) b(x, t, \eta) e^{i[S(x,t,\eta) - y\cdot\eta]}\, d\eta\, dt.$$

In polar coordinates $\eta = \rho\eta^0 = |\eta|\eta^0$, this becomes

$$K(x, y) = \iiint \phi(t) b(x, t, \rho, \eta^0) e^{i\rho[S(x,t,\eta^0) - y\cdot\eta^0]} \rho^{n-1}\, d\rho\, d\eta^0 dt. \tag{39}$$

By changing the smoothing operator $\varepsilon_t f$, we can assume that b is supported in $\rho > 1$. Theorem 1 of Chapter 2 states that the generating function S_t satisfies the Hamilton-Jacobi equation

$$\frac{\partial S_t}{\partial t}(x, \eta^0) = H(x, \xi) = a_1(x, \xi),$$

where Φ_t: $(y, \eta) \to (x, \xi)$. By hypothesis the principal symbol $a_1(x, \xi)$ of $a(x, \xi)$ is positive; therefore,

$$\frac{\partial}{\partial t}\left[S_t - \eta^0 \cdot y\right] > 0. \tag{40}$$

This shows that the exponential in (39) has no stationary points; we can conclude by stationary phase that the kernel $K(x, y)$ is smooth. Since we are on a compact manifold, we have $K(x, y) \in C^\infty(M \times M)$. This proves (38.i).

The implications (i) $\Rightarrow$ (ii) and (iii) $\Rightarrow$ (iv) follow immediately from the definitions. Thus it remains to verify (iii). However, by going carefully through the proof of (i) we can see that $|K(y, x)| \leq C\|\phi\|_{C^k}$ for k large. (Essentially, k is determined by the number of integrations by parts needed to apply stationary phase to (39).) Now (iii) follows at once from (ii).

Let λ_k and ψ_k be the eigenvalues and eigenfunctions of $A = a(x, D)$. Put $\mu = \Sigma_k \delta_{\lambda_k}$. We have just shown that F: $\phi \to \operatorname{Trace} T_\phi$ is a well-defined distribution. We claim

$$\langle F, \phi \rangle = \langle \hat{\mu}, \phi \rangle \quad \text{for } \phi \in C_0^\infty(-T, T), \tag{41}$$

where $\hat{\mu}$ is the Fourier transform of the distribution μ. Indeed if the initial data has an expansion $f = \Sigma c_k \psi_k$, then $u(x, t) = \Sigma e^{it\lambda_k} c_k \psi_k$, and we see that

$$T_\phi f = \sum_k \left[\int \phi(t) e^{it\lambda_k}\, dt\right] c_k \phi_k;$$

i.e., T_ϕ is diagonal with entries $\int \phi(t) e^{it\lambda_k}\, dt$. Therefore

$$\operatorname{Trace} T_\phi = \sum_k \left[\int \phi(t) e^{it\lambda_k}\, dt\right] = \sum_k \hat{\phi}(\lambda_k) = \langle \mu, \hat{\phi} \rangle,$$

as desired.

Step 3. For $\phi \in C_0^\infty(-T, T)$ we claim

$$\langle F, \phi \rangle = \left\langle \sum_{k=0}^{\infty} \frac{c_k}{t^{n-k}} + h(t), \phi \right\rangle, \tag{42}$$

where $h(t) \in C^\infty$ and c_k are constants which are defined below. Indeed, by integrating (39) we find

$$\langle F, \phi \rangle = \operatorname{Trace} T_\phi$$
$$= \iiiint \phi(t) b(y, t, \rho, \eta^0) e^{i\rho[S(y,t,\eta^0) - \eta^0 \cdot y]} \rho^{n-1}\, d\rho\, d\eta^0 dt dy.$$

We can evaluate this by substituting the expansion

$$b(y, t, \rho, \eta^0) = \sum_k b_k(y, t, \eta^0) \rho^{-k}$$

and using the identity

$$\int_{0+}^{\infty} \rho^{n-1-k} e^{iz\rho}\, d\rho = \frac{\text{const}}{z^{n-k}},$$

with $z = S_t(y, \eta^0) - y \cdot \eta^0$. Here $1/z^\mu$ is the distribution defined by

$$\frac{1}{z^\mu} = \lim_{\varepsilon \to 0^+} \frac{1}{(z + i\varepsilon)^\mu}.$$

Now $\partial S/\partial t = a$ and by hypothesis the principal symbol a_1 is positive; also recall $S_t(y, \eta^0) = y \cdot \eta^0$, when $t = 0$; therefore

$$z \sim te(y, \eta^0),$$

where $e(y, \eta^0)$ is smooth and nonvanishing. This leaves

$$\langle F, \phi \rangle = \int \phi(t) \left\{ \sum_{k=0}^{\infty} \frac{\tilde{c}_k(t)}{t^{n-k}} \right\} dt, \tag{43}$$

with

$$\tilde{c}_k(t) = \int \frac{b_k(y, t, \eta^0)}{\left((S_t(y, \eta^0) - y \cdot \eta^0)/t \right)^{n-k}}\, dy\, d\eta^0$$

smooth in t. This proves (42).

Now we can calculate the Fourier transform of $1/t^\mu$ with the table

Function	Fourier transform
sign λ	$c(1/t)$
1	δ_0
$\chi_{[0,\infty)}$	$c\left(\frac{1}{t} + \delta_0\right) = \lim_{\varepsilon \to 0^+} \frac{c}{t + i\varepsilon}$
$\lambda^{n-1}\chi_{[0,\infty)}$	$\lim_{\varepsilon \to 0^+} \frac{c}{(t + i\varepsilon)^n}$

We have

$$\langle F, \phi \rangle = \int \phi(t) \left\{ \frac{c_0}{t^n} + (\text{less singular terms}) \right\} dt \tag{44}$$

for $\phi \in C_0^\infty(-T, T)$.

Step 4. We shall complete the proof using a Tauberian theorem. Equation (42) determines the distribution

$$F_\theta = \theta\hat{\mu} \tag{45}$$

for $\theta \in C_0^\infty(-T, T)$. We may pick θ so that

$$\begin{aligned} &\text{(i)} \quad \hat{\theta} \geqslant 0 \text{ and even,} \\ &\text{(ii)} \quad \hat{\theta}(\lambda) \geqslant \tfrac{1}{10} \text{ on } |\lambda| \leqslant 1, \\ &\text{(iii)} \quad \int \hat{\theta}(\lambda)\, d\lambda = 1. \end{aligned} \tag{46}$$

Now from (44) and (45) and the table above,

$$\hat{F}_\theta(\lambda) = \mu * \hat{\theta}(\lambda) = c_n \lambda^{n-1} \chi_{[0,\infty)}(\lambda) + O(\lambda^{n-2}), \qquad \lambda \to \infty, \tag{46a}$$

so

$$(\mu * \hat{\theta})[0, \lambda] = c\lambda^n + O(\lambda^{n-1}), \qquad \lambda \to \infty. \tag{47}$$

We claim

$$\mu(\lambda - 1, \lambda) = O(\lambda^{n-1}), \qquad \lambda \to \infty. \tag{48}$$

Indeed

$$\hat{F}_\theta(\lambda) = \int \hat{\theta}(\nu)\, d\mu(\lambda - \nu) \leqslant c\lambda^{n-1}$$

by (46a), so that (46.i), (46.ii) imply

$$\frac{1}{10}\mu(\lambda - 1, \lambda) \leqslant \int_{|\nu| \leqslant 1} \hat{\theta}(\nu)\, d\mu(\lambda - \nu) \leqslant c\lambda^{n-1},$$

which is (48).

We want to estimate

$$\begin{aligned} \mu[0, \lambda] &= \int_{-\infty}^{\infty} \chi_{[0,\lambda]}(\nu)\, d\mu(\nu) \\ &= \int_{-\infty}^{\infty} \chi_{[0,\lambda]} * \hat{\theta}(\nu)\, d\mu(\nu) + \int_{-\infty}^{\infty} \left[\chi_{[0,\lambda]} - \chi_{[0,\lambda]} * \hat{\theta}\right](\nu)\, d\mu(\nu) \\ &= I + II. \end{aligned}$$

The first integral is estimated using (47)

$$I = c\lambda^n + O(\lambda^{n-1}), \qquad \lambda \to \infty;$$

for the second we have using (48)

$$\begin{aligned} II &\leqslant \int_{-\infty}^{\infty} \frac{c_N\, d\mu(\nu)}{(1 + |\lambda - \nu|)^N} = \sum_{\text{integers } k} \int_{|\nu - k| \leqslant 1/2} \frac{c_N\, d\mu(\nu)}{(1 + |\lambda - \nu|)^N} \\ &\leqslant \sum_{\text{integers } k} \frac{c_N}{(1 + |\lambda - k|)^N} \cdot O(|k|^{n-1} + 1) \\ &= O(\lambda^{n-1}), \qquad \lambda \to \infty, \end{aligned}$$

if we take N large enough. This gives

$$N(\lambda) = \mu[0, \lambda] = c\lambda^n + O(\lambda^{n-1}), \qquad \lambda \to \infty,$$

and finishes the proof of Hörmander's theorem. □

If $A = (-\Delta)^{1/2}$ and $M =$ the n-sphere, then according to Hörmander's theorem

$$N(\lambda) = c_n\lambda^n + O(\lambda^{n-1}).$$

A direct count using spherical harmonics shows that the error term cannot be improved. On the other hand consider $A = -\Delta$ acting on the n-torus. We know from the discussion following equation (10) that

$$N(\lambda) = c_n\lambda^n + O(\lambda^{n-1}),$$

where $N(\lambda) = \{\text{eigenvalues of } A < \lambda^2\}$. Since eigenvalues are square roots of integers, and since there are approximately $\sim \lambda$ square roots of integers between λ and $\lambda + 100$, we see that $N(\lambda)$ is a step function with $\sim \lambda$ jumps in $[\lambda, \lambda + 100]$ and an increment of $\sim \lambda^{n-1}$ in that interval. Therefore at least one jump of $N(\lambda)$ exceeds λ^{n-2}; this means that we cannot have an error better than

$$N(\lambda) = c_n\lambda^n + O(\lambda^{n-2}).$$

In fact, we have the following theorem which is due to Van der Corput, Hlawka, Herz and Hardy-Littlewood; see **[30]**.

THEOREM 7. *$N(\lambda) = c_n\lambda^n + O(\lambda^{n-1-a})$, where $a = a(n) \to 1$, as $n \to \infty$.*

PROOF. Recall that with $f \in \mathcal{S}(\mathbf{R}^n)$, the Poisson summation formula states

$$\sum_{k \in \mathbf{Z}^m} f(k) = c_n \sum_{k \in \mathbf{Z}^m} \hat{f}(k). \tag{49}$$

We will need this formula later; for now fix an approximation to the identity $\phi \in C_0^\infty$ satisfying

(i) $\phi \geqslant 0$,

(ii) $\phi(x)$ supported in $|x| \leqslant 1$,

(iii) $\int \phi(x)\,dx = 1$,

and define

$$\phi_\zeta(x) = \zeta^{-n}\phi(x/\zeta), \tag{50.i}$$

$$N(\lambda, \zeta) = \sum_{k \in \mathbf{Z}^n} \chi_{B(\lambda)} * \phi_\zeta(k), \tag{50.ii}$$

where $\chi_{B(\lambda)}$ is the characteristic function of the ball $\{x \in \mathbf{R}^n\colon |x| \leqslant \lambda\}$. Note that if $|k| < \lambda - 10\zeta$ or $|k| > \lambda + 10\zeta$, then $\chi_{B(\lambda)} * \phi_\zeta = \chi_{B(\lambda)}$ and so

$$N(\lambda - 10\zeta, \zeta) \leqslant N(\lambda) \leqslant N(\lambda + 10\zeta, \zeta). \tag{51}$$

By (49) and (50)

$$N(\lambda, \zeta) = \sum_{k \in \mathbf{Z}^m} \hat{\chi}_{B(\lambda)}(k)\hat{\phi}_\zeta(k) = \lambda^n \sum_{k \in \mathbf{Z}^m} \hat{\chi}(k\lambda)\hat{\phi}_\zeta(k), \tag{52}$$

where $\chi = \chi_{B(1)}$ and $k = 0$ is the main term. We need the following estimates

$$|\hat{\chi}(k\lambda)| \leq \frac{c_n}{(|k|\lambda)^{(n+1)/2}}, \qquad |\hat{\phi}(\zeta)| \leq \frac{c_m}{(1+|\zeta|)^m}.$$

The second follows from the fact that $\phi \in \mathcal{S}$ and the first follows from writing

$$\int_{\mathbf{R}^n} e^{i\lambda\cdot x}\chi(x)\,dx = \int_{|x|\leq 1} e^{i\lambda\cdot x}\,dx = \int_{-1}^{1} e^{i|\lambda|x_1}\left(1 - x_1^2\right)^{\text{power}} dx_1.$$

To evaluate the integral we expand $(1 - x_1^2)^{\text{power}}$ in powers of $1 - x$ near $x = 1$ and in powers of $1 + x$ near $x = -1$. This gives asymptotic formulas for the contribution near $x = \pm 1$. After the usual argument with integration by parts and stationary phase, we see that apart from the contributions near $x = \pm 1$ the integral decreases rapidly as $\lambda \to \infty$. For details, see **[30]**. With these estimates (52) becomes

$$N(\lambda, \zeta) = \omega_n\lambda^n + O\left(\lambda^n\left\{\sum_{\substack{k\neq 0\\ k\in\mathbf{Z}^m}}\left[\frac{c_n}{(|k|\lambda)^{(n+1)/2}}\,\frac{c_m}{(1+|k|\zeta)^m}\right]\right\}\right),$$

where $\omega_n = \text{volume}\{|x|\leq 1\}$.

Next we estimate the term in braces $\{\ \}$. The main contribution occurs when $|k| < \zeta^{-1}$ and so the error is

$$\begin{aligned}
\lambda^n \sum_{\substack{0\neq k\in\mathbf{Z}^m\\ |k|\leq\zeta^{-1}}} \frac{c_n}{(|k|\lambda)^{(n+1)/2}} &\approx \lambda^{(n-1)/2} \sum_{\substack{0\neq k\in\mathbf{Z}^m\\ |k|\leq\zeta^{-1}}} \frac{1}{(1+|k|)^{(n+1)/2}} \\
&\approx \lambda^{(n-1)/2}\int_{|z|<\zeta^{-1}} \frac{dz}{(1+|z|)^{(n+1)/2}} \\
&\approx \lambda^{(n-1)/2}\zeta^{-((n-1)/2)}.
\end{aligned}$$

So

$$N(\lambda, \zeta) = \omega_n\lambda^n + O\left(\lambda^{(n-1)/2}\zeta^{-((n-1)/2)}\right).$$

Substituting $\lambda \pm 10\zeta$ in place of λ here and combining with (51), we find that

$$N(\lambda) = \omega_n\lambda^n + O\left(\lambda^{n-1}\zeta + \lambda^{(n-2)/2}\zeta^{-((n-1)/2)}\right).$$

Finally if we pick ζ to minimize the remainder, we get

$$N(\lambda) = \omega_n\lambda^n + O\left(\lambda^{n-2+2/(n+1)}\right). \quad \square$$

7. Introduction to global Fourier integral operators. For short time $|t| \ll 1$ we solved the wave equation by

$$T^t u(y) = \int a(y, \xi)e^{i[S_t(y,\xi) - \xi\cdot x]}u(x)\,dx\,d\xi,$$

where S_t is the generating function of a Hamiltonian flow Φ_t. However, for long times t, Φ_t will not have a generating function. The starting point of the global theory of Fourier integral operators (Hörmander [34]; see also Maslov [45]) is that even for long times t we can solve the wave equation by an operator

$$T^t u(y) = \int a(y, x, \zeta) e^{i\phi_t(y,x,\zeta)} u(x)\, dx d\zeta, \tag{53}$$

where ϕ_t is homogeneous in ζ of degree 1, and smooth in x, y, $\zeta/|\zeta|$.

The idea is that the t-axis is covered by overlapping small intervals $\{I_\alpha\}$; in each I_α we can represent T^t in the form (53) with suitable a, ϕ_t. However, if we pass from I_α to an overlapping interval I_β then the form of the representation (53) may change completely. Thus, the amplitude a, the phase function ϕ, and even the dimension of the space of dummy variables ζ may be different for I_α and I_β. For small I_α near $t = 0$ we can take $\phi_t(y, x, \zeta) = S_t(y, \zeta) - \zeta \cdot x$, but for most of the t-intervals I_β this will not be possible.

To carry out this idea, a main issue is clearly to decide when two integrals of the form (53) define approximately the same operator. Since $Tu(y) = \int K(y, x) u(x)\, dx$ with

$$K(y, x) = \int a(y, x, \zeta) e^{i\phi(y,x,\zeta)}\, d\zeta, \tag{54}$$

it is enough to classify distributions of the form (54). So we consider locally in $\mathbf{R}^n$ two distributions

$$K(z) = \int a(z, \tau) e^{i\phi(z,\tau)}\, d\tau, \tag{55}$$

$$K'(z) = \int a'(z, \tau') e^{i\phi'(z,\tau')}\, d\tau'. \tag{56}$$

Our assumptions on a and ϕ are as follows:

— $\phi(z, \tau)$ is defined on $\mathbf{R}^n \times \mathbf{R}^N$, real valued and homogeneous of degree 1 in τ;
— $a(z, \tau)$ is supported in a small conic neighborhood in (z, τ)-space;
— $a \in S^m$, that is $|\partial_\tau^\alpha \partial_z^\beta a| \leq C_{\alpha\beta}(1 + |\tau|)^{m-|\alpha|}$;
— ϕ is nondegenerate in the sense that $d_{x,\tau}\phi(x, \tau) \neq 0$ and $d_{x,\tau}(\partial\phi/\partial\tau_1), \ldots, d_{x,\tau}(\partial\phi/\partial\tau_N)$ are linearly independent on the set $C_\phi = \{(x, \tau) \in \mathbf{R}^n \times \mathbf{R}^N : d_\tau\phi(x, \tau) = 0\}$.

(The last condition vastly simplifies analysis of the critical points of ϕ. It turns out that the wave equation for long time can be solved using nondegenerate phase functions.)

Assume a' and ϕ' are defined on $\mathbf{R}^n \times \mathbf{R}^{N'}$ and satisfy analogous conditions; in particular $a' \in S^{m'}$.

(57) *Problem*. Decide when the distributions K and K' are approximately the same.

The first step in the classification is to determine the wave-front set of K. Stationary phase suggests that the important contributions to the integral (55) arise at $C_\phi = \{(z, \tau) \mid (\partial\phi/\partial\tau)(z, \tau) = 0\}$. Furthermore, one computes that $(z, \tau) \in C_\phi$ contributes $\iota_\phi(z, \tau) \equiv (z, (\partial\phi/\partial z)(z, \tau)) = (z, \xi)$ to $\mathrm{WF}(K)$. Thus, $\mathrm{WF}(K) = \iota_\phi(C_\phi) \equiv \Lambda_\phi$.

The nondegeneracy conditions on ϕ guarantee that Λ_ϕ is a manifold, homogeneous in the ξ-variable (i.e., $(z, \xi) \in \Lambda_\phi$ implies $(z, \delta\xi) \in \Lambda_\phi$, $\delta > 0$). Also Λ_ϕ has a very important special property: The symplectic form $\omega = \Sigma_k d\xi_k \wedge dz_k$ vanishes when restricted to Λ_ϕ. One says that Λ_ϕ is *Lagrangian*.

EXAMPLE. Suppose $z = (x, y)$, $\phi(z, \tau) = S(y, \tau) - \tau \cdot x$, where S is the generating function of a canonical transformation Φ. This is of course precisely the phase function considered before. Then $\Lambda_\phi = \{(x, \xi, y, -\eta) \mid \Phi(x, \xi) = (y, \eta)\}$. The vanishing of the symplectic form on Λ_ϕ just means that $\Sigma_k d\xi_k \wedge dx_k - \Sigma_k d\eta_k \wedge dy_k = 0$, i.e., Φ is canonical.

Now the classification problem (57) is answered by the following result [**34**].

THEOREM 8. *Let $\phi(z, \tau)$ and $\phi'(z, \tau')$ be phase functions defined on $\mathbf{R}^n \times \mathbf{R}^N$ and $\mathbf{R}^n \times \mathbf{R}^M$ respectively as above. Suppose $\Lambda_\phi = \Lambda_{\phi'}$. Then given any symbol $a \in S^m$, there is a symbol $a' \in S^{m'}$ with $m' = m + \frac{1}{2}(N - M)$ so that the Fourier integral distributions* (55), (56) *differ only by a C^∞ error. Moreover, a' can be computed modulo lower-order errors by the formula $(a \circ \iota_\phi^{-1}) \approx \theta \cdot (a' \circ \iota_{\phi'}^{-1})$ on Λ_ϕ. Here, θ is an elliptic symbol depending on ϕ, ϕ' but not on a.*

To prove this, one simply calculates the asymptotic behavior of the Fourier transforms of K, K' at infinity using stationary phase.

The above theorem lets us define a class of Fourier integral distributions associated to a Lagrangian manifold Λ, since locally one can easily find nondegenerate phase functions ϕ so that $\Lambda = \Lambda_\phi$. The choice of ϕ is irrelevent by Theorem 8, and one can then patch together local Fourier integral distributions into global ones by pseudodifferential partitions of unity. By calculating explicitly θ in Theorem 8, one can define intrinsically the principal symbol of a Fourier integral distribution as a section of a suitable bundle.

This means that we can associate Fourier integral operators to canonical transformations, even when there is no generating function. As expected, taking the product of two Fourier integral operators corresponds to composing the canonical transformations from which they arise. In particular, taking $\Phi_t =$ Hamiltonian flow as in the short-time case, we can show that the (half)-wave equation is solved globally in time by Fourier integral operators arising from the canonical transformation Φ_t.

These ideas have interesting applications. For instance, they can be used to relate the eigenvalues of the Laplacian to the lengths of closed geodesics on a manifold. (See Duistermaat and Guillemin [**15**].) We shall not pursue this here. We have meant these fragmentary remarks only as a lead-in to the following explanation, contributed by D. H. Phong, of Fourier integral operators with complex phase functions. These will be needed for the Bergman and Szegö kernels in Chapter 12.

8. Introduction to Fourier integral operators with complex phase. The Bergman and Szegö projections can be represented as Fourier integral operators, but with complex phase functions. This can be easily seen in the model case of the Mizohata operator

$$P = \partial/\partial t + t\,|D_x|\,, \tag{57a}$$

which can be viewed as an analogue of the $\bar{\partial}_b$ equation and will play an important role in the sequel. The null space N of P is nontrivial, but the orthogonal projection S onto N together with a right inverse Q with values in $N^\perp$ can be contructed as follows. Taking Fourier transforms with respect to x transforms P into $d/dt + t\,|\xi|$, which vanishes over $\phi_0(t,\xi) = e^{-(1/2)t^2|\xi|}/|\xi|^{1/4}$. On the other hand $Pf = \mu$ implies

$$(d/dt + t\,|\xi|)\tilde{f}(t,\xi) = \tilde{\mu}(t,\xi)$$

and thus

$$\tilde{f}(t,\xi) = \int_0^t e^{-(1/2)|\xi|(t^2-s^2)}\tilde{\mu}(s,\xi)\,ds + K\phi_0(t,\xi).$$

The condition $\langle \tilde{f}(t,\xi), \phi_0(\cdot,\xi)\rangle_{L^2(\mathbf{R})} = 0$ yields

$$K(\xi) = -|\xi|^{1/4}\int_{-\infty}^{\infty} e^{-|\xi|\alpha^2}\int_0^\alpha e^{(1/2)s^2|\xi|}\tilde{\mu}(s,\xi)\,ds\,d\alpha.$$

From this we obtain the following formulae for S and Q:

$$\begin{aligned}(Su)(t,x) &= \int e^{i\langle x,\xi\rangle}\left\langle \tilde{\mu}(\cdot,\xi), \phi_0(\cdot,\xi)\right\rangle_{L^2(\mathbf{R})}\phi_0(t,\xi)\,d\xi\\ &= \iint e^{i\langle x,\xi\rangle}e^{-(1/2)(s^2+t^2)|\xi|}\tilde{\mu}(s,\xi)\frac{1}{|\xi|^{1/2}}\,ds\,d\xi\\ &= \iiint e^{i\langle x-y,\xi\rangle}e^{-(1/2)(s^2+t^2)|\xi|}\frac{1}{|\xi|^{1/2}}u(s,y)\,dy\,ds\,d\xi;\end{aligned} \tag{58}$$

$$\begin{aligned}(Qu)(t,x) &= \int e^{i\langle x,\xi\rangle}\tilde{f}(t,\xi)\,d\xi\\ &= \iint e^{i\langle x,\xi\rangle}e^{it\tau}\sigma(t,\tau,\xi)\tilde{\mu}(\tau,\xi)\,d\tau\,d\xi,\end{aligned}$$

where

$$\sigma(t,\tau,\xi) = \left(\frac{|\xi|}{\pi}\right)^{1/2}\int_{-\infty}^{\infty} e^{-|\xi|\alpha^2}\int_\alpha^t e^{-(1/2)|\xi|(t^2-s^2)}e^{-i(t-s)\tau}\,ds\,d\alpha. \tag{59}$$

A direct calculation shows that

$$PQ = \mathrm{I}, \qquad QP = \mathrm{I} - S, \tag{60}$$

which reduces existence and regularity questions for P to the study of Q and S. Now Q and S are pseudodifferential operators of class $S_{1/2,1/2}$, and sharp estimates for Q can be obtained by imbedding Q in more restrictive classes (see Nagel-Stein **[49]**, in which the formula (59) first appears); however, one can also view S as a Fourier integral operator with a classical symbol, but a complex phase function

$$\phi(x, t; y, s; \xi) = \langle x - y, \xi \rangle + i(t^2 + s^2)|\xi|/2. \tag{61}$$

We now turn to a study of such integrals, following Melin-Sjöstrand **[46]**. The discussion will be local stressing only the main differences with the real case. Once more the key tool will be the method of stationary phase, suitably extended to complex functions.

Consider then the integral

$$I(x, \lambda) = \int e^{i\lambda\phi(x,y)}\mu(y, \lambda)\, dy, \qquad x \in \mathbf{R}^n, y \in \mathbf{R}^N,$$

where $\operatorname{Im}\phi \geqslant 0$, and $|D_x^\alpha D_\lambda^\beta \mu(y, \lambda)| \leqslant C_{\alpha\beta}\lambda^{m-|\beta|}$. The main contributions to the integral will come from the critical points of ϕ, which will be assumed to be nondegenerate. Thus assume $\mu(y, \lambda)$ is supported in a small neighborhood of 0, $d_y\phi(0,0) = 0$, and $d_{yy}^2\phi(0,0) \neq 0$. Unlike the real case, the set $C_{\phi,\mathbf{R}} = \{(x, y) \in \mathbf{R}^n \times \mathbf{R}^N;\ d_y\phi(x, y) = 0\}$ will not in general be a C^∞ manifold, nor if it is will it have the proper dimension. If ϕ is analytic we may view $C_{\phi,\mathbf{R}}$ as the set of real points of $C_\phi = \{(x, y) \in \mathbf{C}^n \times \mathbf{C}^N;\ \partial_y\phi(x, y) = 0\}$ which is C^∞. When ϕ is merely C^∞ we are naturally led to consider almost analytic extensions of ϕ.

Thus let ω, Ω be open sets in $\mathbf{R}^n$ and $\mathbf{C}^n$ respectively, with $\Omega \cap \mathbf{R}^n = \omega$, and let f be a C^∞ function on Ω. f is said to be almost analytic if $\bar\partial f$ vanishes to infinite order on ω. Any C^∞ function on ω admits an infinite number of almost analytic extensions which are all equivalent in the sense that the difference between any two vanishes to infinite order at the real points. We shall not distinguish between equivalent functions. If $\mu \in S^m$ is a symbol on a cone, then the extension can be chosen to satisfy analogous estimates. An almost analytic manifold $\Lambda \subset \Omega$ is a C^∞ manifold defined locally near real points by almost analytic functions $f_1(x) = \cdots = f_k(x) = 0$, with $\partial f_1(x), \ldots, \partial f_k(x)$ linearly independent over $\mathbf{C}$. Two almost analytic manifolds Λ_1 and Λ_2 will be said to be equivalent if $d(z, \Lambda_1) \leqslant C_N |\operatorname{Im} z|^N$ for $z \in \Lambda_2$ (here $d(z, \Lambda_1)$ denotes the distance from z to Λ_1). This is easily seen to be an equivalence relation which will allow us to identify manifolds in the same class. Almost analytic functions on Λ will be restrictions to Λ of almost analytic functions on Ω.

Returning now to the integral $I(x, \lambda)$, observe that the set

$$C_{\tilde\phi} = \left\{(x, y) \in \mathbf{C}^n \times \mathbf{C}^N;\ \partial_y\tilde\phi(x, y) = 0\right\}, \tag{62}$$

where $\tilde\phi$ denotes an almost analytic extension of ϕ, is an almost analytic manifold. The points of $C_{\tilde\phi}$ will still be called critical points of ϕ, despite the fact that this notion does not coincide with the usual one when x or y is complex. Near 0, parametrize $C_{\tilde{\phi}}$ by $x \to (x, Y(x))$. Then

THEOREM 9 (METHOD OF STATIONARY PHASE). *Locally near* 0, *we have*

$$\text{(63)}\quad \begin{aligned} &\text{(a)} \quad \operatorname{Im} \tilde{\phi}(x, Y(x)) \geqslant C\,|\operatorname{Im} Y(x)\,|^2. \\ &\text{(b)} \quad \textit{Let } \tilde{\mu}(y, \lambda) \textit{ be an almost analytic extension of } \mu \textit{ with} \\ &\qquad\ \textit{good estimates. Then} \end{aligned}$$

$$\text{(64)}\qquad I(x, \lambda) \sim \lambda^{-N/2} e^{i\lambda\tilde{\phi}(x, Y(x))} \sum_{k=0}^{\infty} \lambda^{-k} \big(R_k(x, D)\tilde{\mu}\big)(Y(x), \lambda).$$

Here $R_k(x, D)$ *are differential operators of order* $\leqslant 2k$, *with*

$$\text{(65)}\qquad R_0(x, D) = (2\pi)^{N/2}\Big[\det\Big(i^{-1}\operatorname{Hess}_{yy}\tilde{\phi}(x, Y(x))\Big)\Big]^{-1/2}.$$

(*The branch of the square root is chosen so that it equals* 1 *at* 1; *this can be achieved since the eigenvalues of* $i^{-1}\operatorname{Hess}_{yy}\phi(0,0)$ *do not lie on the negative real axis.*)

PROOF OF (a). Let $A = \operatorname{Hess}_{yy}\tilde{\phi}(x, Y(x))$. The almost analyticity of $\tilde{\phi}$ and Taylor's formula imply

$$\text{(66)}\qquad \begin{aligned} \tilde{\phi}(x, y) = \tilde{\phi}(x, Y(x)) &+ \tfrac{1}{2}\langle A(y - Y(x)), y - Y(x)\rangle \\ &+ O\big(|\operatorname{Im} Y(x)\,|^3 + |\,y - Y(x)\,|^3\big). \end{aligned}$$

Since $\operatorname{Im}\phi(x, y) \geqslant 0$ for real values of the argument, it suffices to exhibit y in the real ball centered at $\operatorname{Re} Y(x)$ and of radius $\operatorname{Im} Y(x)$ such that

$$\text{(67)}\qquad \operatorname{Im}\big\langle A(y - Y(x)), (y - Y(x))\big\rangle \leqslant -C\,|\operatorname{Im} Y(x)\,|^2.$$

Writing $y = Y(x) = i[\operatorname{Im} Y(x) + ik]$, $k \in \mathbf{R}^N$, (67) reduces to finding k such that $|\,k\,| < |\operatorname{Im} Y(x)\,|$ and

$$\operatorname{Im}\big\langle A(\operatorname{Im} Y(x) + ik), \operatorname{Im} Y(x) + ik\big\rangle \leqslant C\,|\operatorname{Im} Y(x)\,|^2.$$

Now for $\varepsilon > 0$ small enough the equation

$$\text{(68)}\qquad (I + Q^t)A(I + Q) = A + i\varepsilon I$$

admits a solution $Q(\varepsilon)$ since it does for $\varepsilon = 0$ and the differential of the mapping $Q \to (I + Q^t)A(I + Q)$ from the space of complex $N \times N$ matrices to the space of complex symmetric matrices admits a right inverse when A is invertible. Obviously $Q(\varepsilon) \to 0$ as $\varepsilon \to 0$. Letting $T(\varepsilon) = (\operatorname{Im} Q)(I + \operatorname{Re} Q)^{-1}$ and multiplying (68) by $(I + \operatorname{Re} Q)^{-1}$ yields

$$(I + iT^t)A(I + iT) = \big(I + \operatorname{Re} Q^t\big)^{-1}(A + i\varepsilon I)(I + \operatorname{Re} Q)^{-1}.$$

Set $k = T(\varepsilon)(\operatorname{Im} Y(x))$. In view of the fact that $\operatorname{Im} A \geqslant 0$ we have

$$\operatorname{Im}\big\langle A(\operatorname{Im} Y(x) + ik), \operatorname{Im} Y(x) + ik\big\rangle \geqslant \varepsilon\,\big|\big(I + \operatorname{Re} Q^t\big)^{-1}k\,\big|^2 \geqslant C\,|\operatorname{Im} Y(x)\,|^2,$$

as was to be shown. $\square$

PROOF OF (b). We begin by choosing for each x an appropriate system of coordinates for $\mathbf{C}^N$ near $Y(x)$. According to Taylor's formula

$$\text{(69)}\quad \tilde{\phi}(x, y) = \tilde{\phi}(x, Y(x)) + l(x, y) + \tfrac{1}{2}\big\langle R(x, y)(y - Y(x)), y - Y(x)\big\rangle.$$

Here the presence of the "linear term" $l(x, y)$ is caused by the fact that $Y(x)$ is not a critical point of ϕ in the usual sense when $Y(x)$ is complex. The terms l and R satisfy the estimates

$$|\bar\partial_y R(x, y)| + |l(x, y)| \leqslant C_K\big(|\operatorname{Im} y|^K + |\operatorname{Im} Y(x)|^K\big). \tag{70}$$

Let P be a matrix with $P^t R(0,0)P = i\mathrm{I}$. Now the map $Q \to iQ^tQ$ from $\mathrm{GL}(N, \mathbf{C})$ into the space of symmetric matrices is analytic with surjective differential at $Q = P^{-1}$, and thus for (x, y) small we may find $Q(x, y)$ analytic in (x, y) such that

$$iQ^t(x, y)Q(x, y) = R(x, y), \qquad Q(0,0) = P^{-1}$$

and

$$|\bar\partial_x Q(x, y)| \leqslant C_K\big(|\operatorname{Im} y|^K + |\operatorname{Im} Y(x)|^K\big).$$

For each x define now new coordinates in $\mathbf{C}^N$ by

$$y \to \zeta(x, y) = Q(x, y)(y - Y(x)).$$

Then (69) becomes

$$\tilde\phi(x, y) = \tilde\phi(x, Y(x)) + i\langle\zeta, \zeta\rangle/2 + l(x, y). \tag{71}$$

Set $\zeta = \xi + i\eta$. Observe that at $x = 0$, $Y(0) = 0$, $|\xi|^2 - |\eta|^2 = \operatorname{Im}\phi(0, y) \geqslant 0$ for $y \in \mathbf{R}^N$, so that in the new coordinates $\mathbf{R}^N$ is given by an equation $\eta = g(\xi, x)$. We write $y(\zeta)$ for the inverse map of $y \to \zeta(x, y)$ and introduce the following chains $\Gamma_{x,t}$ in $\mathbf{C}^N$:

$$\zeta_t(\xi) = \xi + itg(\xi, x), \qquad 0 \leqslant t \leqslant 1, \xi \in \mathbf{R}^N,$$

$$\Gamma_{x,t}\colon \xi \to y(\zeta_t(\xi)).$$

We now show that

$$\int_{\mathbf{R}^N} e^{i\lambda\phi(x,y)}\mu(y, \lambda)\, dy - \int_{\Gamma_{x,0}} e^{i\lambda\tilde\phi(x,y)}\tilde\mu(y, \lambda)\, dy_1 \wedge \cdots \wedge dy_N = O(\lambda^{-K}) \tag{72}$$

for any K. Indeed by Stokes' formula the left-hand side of (72) can be estimated by

$$\left|\int_{\substack{\cup\Gamma_{x,t}\\ 0\leqslant t\leqslant 1}} \bar\partial_y\big(e^{i\lambda\tilde\phi(x,y)}\tilde\mu(y, \lambda)\big)\, dy_1 \wedge \cdots \wedge dy_N\right|$$

$$\leqslant C \sup_{\substack{\xi \text{ near } 0\\ x \text{ near } 0\\ 0\leqslant t\leqslant 1}} \left|e^{i\lambda\tilde\phi(x,y(\zeta_t(\xi)))}\big[i\lambda\tilde\mu(y(\zeta_t), \lambda)\bar\partial_y\tilde\phi(x, y(\zeta_t)) + \bar\partial_y\tilde\mu(y(\zeta_t), \lambda)\big]\right|,$$

and thus (72) is a consequence of the almost analyticity of $\tilde u$, $\tilde\phi$, and the following inequalities:

$$(1 - t)(|\operatorname{Im} Y(x)| + |\xi|) \geqslant C|\operatorname{Im} y(\zeta_t)|, \tag{73}$$

$$|\operatorname{Im}\tilde\phi(x, y(\tilde\zeta_t))| \geqslant C(1 - t)\big(|\operatorname{Im} Y(x)|^2 + |\xi|^2\big). \tag{74}$$

To establish (73), write

$$|\operatorname{Im} y(\zeta_t(\xi))| = |\operatorname{Im} y(\zeta_t(\xi)) - \operatorname{Im} y(\zeta_1(\xi))| \leq C(1-t)|g(\xi, x)|$$
$$\leq C(1-t)(|g(0,x)| + |\xi|)$$

and observe that $|g(0,x)| \sim |\operatorname{Im} Y(x)|$; as for (74) it follows from (63), (73), the identity

$$\operatorname{Im}\phi(x, y(\tilde{\zeta}_t)) = t^2 \operatorname{Im}\phi(x, y(\tilde{\zeta}_1)) + (1-t^2)\left[\operatorname{Im}\phi(x, Y(x)) + \tfrac{1}{2}|\xi|^2\right]$$
$$+ (1-t^2)\operatorname{Im} l(x, y(\tilde{\zeta}_1)) + \left\{\operatorname{Im}\left(l(x, y(\tilde{\zeta}_t)) - l(x, y(\tilde{\zeta}_1))\right)\right\}$$

(which is a corollary of (71)), and from the estimates (70) which hold for l and its derivatives.

We are thus reduced to the study of the asymptotic expansion of the integral

$$(75) \int_{\Gamma_{x,0}} e^{i\lambda\tilde{\phi}(x,y)}\tilde{\mu}(y,\lambda)\, dy_1 \wedge \cdots \wedge dy_N$$
$$= \int_{\mathbf{R}^N} e^{i\lambda(\tilde{\phi}(x,Y(x)) + i|\xi|^2/2 + l(x,y(\xi)))}\tilde{\mu}(y(\xi),\lambda)\left(\det\left(\frac{\partial y}{\partial \xi}\right)\right) d\xi$$
$$= e^{i\lambda\tilde{\phi}(x,Y(x))}\int e^{-\lambda|\xi|^2/2}\tilde{\mu}(y(\xi),\lambda)\left(\det\left(\frac{\partial y}{\partial \xi}\right)\right) d\xi$$
$$+ \int\int_0^1 e^{i\lambda(\tilde{\phi}(x,Y(x)) + i|\xi|^2/2 + sl(x,y(\xi)))} l(x, y(\xi))\tilde{\mu}(y(\xi),\lambda)\left(\det\left(\frac{\partial y}{\partial \xi}\right)\right) d\xi\, ds$$

which becomes

$$(76) \qquad \int_{\Gamma_{x,0}} e^{i\lambda\tilde{\varphi}(x,y)}\tilde{\mu}(y,\lambda)\, dy_1 \wedge \cdots \wedge dy_N$$
$$\sim e^{i\lambda\tilde{\phi}(x,Y(x))}\int e^{-\lambda|\xi|^2/2}\tilde{\mu}(y(\xi),\lambda)\left(\det\left(\frac{\partial y}{\partial \xi}\right)\right) d\xi,$$

since the integrand in the expression between brackets in (75) is bounded by

$$C\lambda^{(\text{fixed power})}\left(|\operatorname{Im} Y(x)|^K + |\xi|^K\right)\left(e^{-c\lambda(|\operatorname{Im} Y(x)|^2 + |\xi|^2)}\right)$$

in view of (63), (70), and (73).

The phase in (76) being real, we can now proceed as in the classical case by applying Parseval's formula and expanding $(e^{-\lambda|\xi|^2/2})\hat{} = \lambda^{-(N/2)}e^{-|\xi|^2/2\lambda}$ in a Taylor series. The identity (65) follows from an inspection of the coefficients. The proof of Theorem 2 is complete. □

We now consider Fourier integral distributions. Let $\phi(x,\theta)$ be a C^∞ function on $\mathbf{R}^n \times \mathbf{R}^N$ satisfying the following conditions:

(i) $\phi(x, \lambda\theta) = \lambda\phi(x,\theta)$, $\lambda \in \mathbf{R}^+$;

(ii) $d_{x,\theta}\phi(x,\theta) \neq 0$;

(iii) $d_{x,\theta}(\partial\phi/\partial\theta_1), \ldots, d_{x,\theta}(\partial\phi/\partial\theta_N)$ are linearly independent over $\mathbf{C}$ on the set

$$C_{\phi,\mathbf{R}} = \left\{(x,\theta) \in \mathbf{R}^n \times \mathbf{R}^N \setminus 0;\ d_\theta\phi(x,\theta) = 0\right\};$$

(iv) $\operatorname{Im}\phi(x,\theta) \geq 0$;

and let $a(x, \theta)$ be a classical symbol of order m. The distribution

$$\langle I(a, \phi), u\rangle = \iint e^{i\phi(x,\theta)}a(x, \theta)u(x)\, dxd\theta, \qquad u \in C_0^\infty(\mathbf{R}^n),$$

can then be defined as an oscillatory integral, with wave front set contained in

$$\mathrm{WF}(I(a, \phi)) \subset \{(x, d_x\phi(x, \theta)); (x, \theta) \in C_{\phi,\mathbf{R}}\} = \Lambda_{\phi,\mathbf{R}}$$

(observe that $(d_x\phi)(x, \theta)$ is real when $(x, \theta) \in C_{\phi;\mathbf{R}}$). We next determine to what extent the class of distributions $I(a, \phi)$ for various symbols a depends on the phase function ϕ. The method of stationary phase outlined above suggests considering not the sets $C_{\phi,\mathbf{R}}$ and $\Lambda_{\phi,\mathbf{R}}$, but rather the corresponding almost analytic manifolds defined by almost analytic extensions of ϕ. Thus let $\tilde{\phi}$ be an almost analytic extension of ϕ and set

$$C_{\tilde{\phi}} = \{(\tilde{x}, \tilde{\theta}) \in \mathbf{C}^n \times (\mathbf{C}^N \setminus 0); d_{\tilde{\theta}}\tilde{\phi}(\tilde{x}, \tilde{\theta}) = 0\},$$

$$\Lambda_{\tilde{\phi}} = \text{Image of } C_{\tilde{\phi}} \text{ under the mapping } T_{\tilde{\phi}}\colon (x, \theta) \to (\tilde{x}, d_{\tilde{x}}\tilde{\phi}(\tilde{x}, \tilde{\theta})).$$

Then $C_{\tilde{\phi}}$ and $\Lambda_{\tilde{\phi}}$ are almost analytic manifolds, $C_{\tilde{\phi},\mathbf{R}} = C_{\tilde{\phi}} \cap (\mathbf{R}^n \times \mathbf{R}^N \setminus 0)$, $T_{\tilde{\phi}}(C_{\tilde{\phi},\mathbf{R}}) = \Lambda_{\tilde{\phi}} \cap (\mathbf{R}^n \times (\mathbf{R}^N \setminus 0))$, and $\Lambda_{\tilde{\phi}}$ is a positive conic Lagrangian manifold in the sense that near a given real point (x_0, ξ_0), there always exist real symplectic coordinates (x, ξ) with almost analytic extensions $(\tilde{x}, \tilde{\xi})$ so that Λ is given by the equations

$$\tilde{x} = \partial_{\tilde{\xi}}S(\tilde{\xi}) \tag{77}$$

for some homogeneous almost analytic function $S(\tilde{\xi})$ with positive imaginary part. It follows that in any other system of symplectic coordinates (y, η) with the projection $(\tilde{y}, \tilde{\eta}) \in \Lambda_{\tilde{\phi}} \to \tilde{y} \in \mathbf{C}^n$ having surjective differential $\Lambda_{\tilde{\phi}}$ can be represented in the same way as in (77) with a suitable choice of generating functions S.

THEOREM 10 (INVARIANCE UNDER CHANGE OF PHASE FUNCTIONS). *Let $\phi(x, \theta)$ and $\psi(x, \omega)$ be phase functions defined on $\mathbf{R}^n \times (\mathbf{R}^N \setminus 0)$ and $\mathbf{R}^n \times (\mathbf{R}^M \setminus 0)$ respectively. Assume that $d_\theta\phi(x_0, \theta_0) = 0$, $d_\omega\psi(x_0, \omega_0) = 0$ and $\Lambda_\phi \sim \Lambda_\psi$ near $(x_0, \xi_0) = (x_0, d_x\phi(x_0, \theta_0)) = (x_0, d_x\psi(x_0, \omega_0))$. Then any distribution $I(a, \phi)$ for $a \in S^m(\mathbf{R}^n \times (\mathbf{R}^N \setminus 0))$ with conic support in a small conic neighborhood of (x_0, θ_0) can also be represented as a distribution $I(b, \psi)$ for a suitable symbol b in $S^{m+(1/2)(N-M)}(\mathbf{R}^n \times \mathbf{R}^M \setminus 0)$. (Here we identify distributions which differ by C^∞ functions in a neighborhood of x_0.) If a is classical, then b can be chosen to be classical also.*

PROOF OF THEOREM 10. With appropriate coordinates on $\mathbf{R}^n$, we may represent Λ_ϕ near (x_0, ξ_0) by

$$\Lambda_\phi = \{(\tilde{x}, \tilde{\xi}) \in \mathbf{C}^n \times (\mathbf{C}^n \setminus 0); \tilde{x} = \tilde{x}(\tilde{\xi})\},$$

with $\tilde{x}(\tilde{\xi})$ almost analytic and positively homogeneous of degree 0. It is then easily seen that the critical points of the function $(\tilde{x}, \tilde{\theta}) \to \tilde{\phi}(\tilde{x}, \tilde{\theta}) - \langle \tilde{x}, \tilde{\xi}\rangle$ are exactly the pull backs of $(\tilde{x}(\tilde{\xi}), \tilde{\xi})$ under the map $T_{\tilde{\phi}}$, which we shall denote by

$(\tilde{x}(\tilde{\xi}), \tilde{\theta}(\tilde{\xi}))$. Now the Fourier transform of $I(a, \phi)$ at η is given by

$$(I(a, \phi))\hat{}(\eta) = \lambda^N \iint e^{i\lambda(\phi(x,\theta) - \langle x, \xi \rangle)} a(x, \lambda\theta)\, dx d\theta,$$

where we have written $\eta = \lambda\xi$, $|\xi| = 1$, $\lambda \in \mathbf{R}^+$. The integral over the region away from the critical points of $\phi(x, \theta) - \langle x, \xi \rangle$ is of order $C_N \lambda^{-N}$, and thus Theorem 9 applied to the integral near the critical point yields

$$(I(a, \phi))\hat{}(\eta) \sim \lambda^{N-(n/2)} e^{i\lambda(\tilde{\phi}(\tilde{x}(\xi), \tilde{\theta}(\xi)) - \langle \tilde{x}(\xi), \xi \rangle)} \times \sum_{k=0}^{\infty} \lambda^{-k} R_k(\xi, x, D_x, D_\theta)\big(\tilde{a}(\tilde{x}(\xi), \lambda\tilde{\theta}(\xi))\big),$$

where $\tilde{a}$ denotes an almost analytic extension of a. However $\tilde{\phi}(\tilde{x}, \tilde{\theta})$ is $O(|\operatorname{Im} \tilde{x}(\xi)|^K + |\operatorname{Im} \tilde{\theta}(\xi)|^K)$ since $\tilde{\phi}$ is homogeneous and $\tilde{\theta}$ is a critical point of $\tilde{\phi}$, while the arguments leading to (63) still yield

$$\operatorname{Im}\{\tilde{\phi}(\tilde{x}(\xi), \tilde{\theta}(\xi)) - \langle \tilde{x}(\tilde{\xi}), \xi \rangle\} \geqslant C(|\operatorname{Im} \tilde{x}(\xi)|^2 + |\operatorname{Im} \tilde{\theta}(\xi)|^2).$$

It thus follows that

$$(I(a, \phi))\hat{}(\eta) \sim \lambda^{N-(n/2)} e^{i\lambda\langle \tilde{x}(\xi), \xi \rangle} \sum_{k=0}^{\infty} \lambda^{-k} R_k(\xi, x, D_x, D_\theta)\big(\tilde{a}(\tilde{x}(\xi), \tilde{\theta}(\xi))\big).$$

Observing that the phase now no longer depends on ϕ and only on Λ_ϕ, we see that $(I(b, \phi))\hat{}(\eta)$ is given by the same kind of expression with the same phase. The proof of the theorem can be completed by showing that given any symbol $\alpha(x, \eta)$ we may choose $a(x, \theta)$ so that

$$\sum_{k=0}^{\infty} |\eta|^{-k} R_k(\eta/|\eta|, x, D_x, D_\theta)\big(\tilde{a}(\tilde{x}(\eta/|\eta|), \tilde{\theta}(\eta))\big) \sim \alpha(x, \eta).$$

This can be done by successive approximations since $R_0(\eta/|\eta|, x, D_x, D_\theta)$ is $\neq 0$ in view of (65). □

With Theorem 10 it is now possible to associate to a given closed conic positive almost analytic Lagrangian manifold Λ a class of distributions $I^m(\mathbf{R}^n, \Lambda)$. A distribution A will be said to be in $I^m(\mathbf{R}^n, \Lambda)$ if $\mathrm{WF}(A) \subseteq \Lambda_{\mathbf{R}}$ ($\equiv \Lambda \cap \mathbf{R}^n$), and for each $(x_0, \xi_0) \in \Lambda$ we have $(x_0, \xi_0) \notin \mathrm{WF}(A - I(a, \phi))$ for some phase function ϕ with $\Lambda = \Lambda_\phi$ near (x_0, ξ_0), and some classical symbol a of order $m + \frac{1}{4}(n - 2N)$. Theorem 10 then simply says that $I^m(\mathbf{R}^n, \Lambda)$ and $I^m(\mathbf{R}^n, \Lambda^\dagger)$ are identical modulo C^∞ functions when Λ and $\Lambda^\dagger$ are equivalent. The converse is also true when the symbols involved are required to be classical and not merely in $S^{m+(1/4)(n-2N)}$.

Fourier integral operators can next be defined through their distribution kernels. Let X and Y be open sets in $\mathbf{R}^{n_x}$ and $\mathbf{R}^{n_y}$ respectively. Given a submanifold C of $T^*(X \times Y)$, we shall denote by C' the manifold $\{((x, \xi), (y, \eta)) \in T^*(X \times Y); ((x, \xi), (y, -\eta)) \in C\}$. A positive canonical relation C is a submanifold of $(T^*(X \times Y) \setminus 0)^\sim$ such that $C_{\mathbf{R}} \equiv C \cap T^*(X \times Y) \subseteq (T^*(X) \setminus 0) \times (T^*(Y) \setminus 0)$ is a closed set, and C' is a positive conic almost analytic Lagrangian manifold. A Fourier integral operator A associated to a positive canonical relation C is an operator with distribution kernel K_A in $I^m(X \times Y; C')$. Such operators can be extended as operators from compactly supported distributions on Y to compactly supported distributions on X by neglecting regularizing kernels. Now let A_1 and A_2 be Fourier integral operators with kernels $K_{A_1} \in I^{m_1}(X \times Y, C_1')$, $K_{A_2} \in I^{m_2}(Y \times Z, C_2')$, where X, Y, Z are open sets in $\mathbf{R}^{n_x}, \mathbf{R}^{n_y}, \mathbf{R}^{n_z}$, and C_1, C_2 are positive canonical relations; then the composition $A_1 \circ A_2$: $\mathcal{E}'(Z) \to \mathcal{E}'(X)$ is a well-defined operator with singularities determined by

$$\mathrm{WF}(K_{A_1 \circ A_2}) \subseteq \mathrm{WF}(K_{A_1}) \circ \mathrm{WF}(K_{A_2}) \subseteq C_{1,\mathbf{R}} \circ C_{2,\mathbf{R}}.$$

To insure that $A_1 \circ A_2$ also be a Fourier integral operator, however, we will need some addition conditions. If

$$\Delta = T^*(X) \times (\operatorname{diag} T^*(Y) \times T^*(Y)) \times T^*(Z)$$

and $\tilde{\Delta}$ is its almost analytic complexification, we have

THEOREM 11 (COMPOSITION OF FOURIER INTEGRAL OPERATORS). *Assume that $C_1 \times C_2$ and $\tilde{\Delta}$ intersect transversally at $(C_{1,\mathbf{R}} \times C_{2,\mathbf{R}}) \cap \Delta$ and that the projection $C_{1,\mathbf{R}} \times C_{2,\mathbf{R}} \to (T^*(X) \setminus 0) \times (T^*(Z) \setminus 0)$ is one-to-one and proper. Then the mapping*

$$(C_1 \times C_2) \cap \tilde{\Delta} \to (T^*(X, Z) \setminus 0)^{\sim}, \qquad (x, \xi, y, \eta; y, \eta, z, \zeta) \to (x, \xi, z, \zeta)$$

is a local diffeomorphism near real points whose image is a conic positive canonical relation which will be denoted by $C_1 \circ C_2$ and satisfies $(C_1 \circ C_2)_{\mathbf{R}} = C_{1,\mathbf{R}} \circ C_{2,\mathbf{R}}$. Furthermore, the operator $A_1 \circ A_2$ is a Fourier integral operator associated to $C_1 \circ C_2$ of order $m_1 + m_2$.

PROOF OF THEOREM 11. We may represent $K_{A_1 \circ A_2}$ as

$$K_{A_1 \circ A_2}(x, z) = \iiint e^{i(\phi_1(x,y,\theta) + \phi_2(y,z,\sigma))} a_1(x, y, \theta) a_2(y, z, \sigma)\, dy d\theta d\sigma, \tag{78}$$

where ϕ_1, ϕ_2, a_1, a_2 are phase functions and symbols defining A_1 and A_2. Partial integration with respect to (x, θ) and (z, σ) respectively show that the integrals over the regions $|\sigma| < \varepsilon |\theta|$ and $|\theta| < \varepsilon |\sigma|$ are of class C^∞ in (x, z) so that we may restrict our attention to the region where $|\theta| \sim |\sigma|$. There however the integral in (78) may be rewritten as

$$\iiint e^{i\phi(x_1,\theta,\sigma,\omega)} a(x, z; \theta, \sigma, \omega)\, d\theta d\sigma d\omega, \tag{79}$$

where (θ, σ, ω) are the new frequency variables, and

$$\phi(x, z; \theta, \sigma, \omega) = \phi_1\big(x, \omega(|\theta|^2 + |\sigma|^2)^{-1/2}, \theta\big) + \phi_2\big(\omega(|\theta|^2 + |\sigma|^2)^{-1/2}, z, \sigma\big)$$

$$a(x, z; \theta, \sigma, \omega) = a_1\big(x, \omega(|\theta|^2 + |\sigma|^2)^{-1/2}, \theta\big) a_2\big(\omega(|\theta|^2 + |\sigma|^2)^{-1/2}, z, \sigma\big)$$
$$\times \big(|\theta|^2 + |\sigma|^2\big)^{-n_y/2}$$

are phase functions and symbols satisfying good estimates. That (79) is a Fourier integral distribution associated to $C_1 \circ C_2$ can now be verified by calculating the critical points of ϕ, using the transversality conditions, and applying Theorem 10. □

EXAMPLES. Returning to the operator (58) with the phase function (61) we readily see that C_ϕ is given by $x - y + i(t^2 + s^2)\xi/(2|\xi|) = 0$, and thus the canonical relation C corresponding to ϕ is the manifold of points $\{(x, \xi), (t, \tau); (y, \eta), (s, \rho)\}$ satisfying

$$\begin{aligned} \eta &= \xi, & \tau &= it|\xi|, \\ y &= x + i(t^2 + s^2)\xi/(2|\xi|), & \rho &= -is|\xi|. \end{aligned}$$

Let $V = \{(x, \xi, t, \tau) \in T^*(\mathbf{R}^{n+1}) \setminus 0;\ t = \tau = 0\}$ and $\tilde{V} = \{(x, \xi, t, \tau) \in (T^*(\mathbf{R}^{n+1}) \setminus 0)^\sim;\ i\tau + t|\xi| = 0\}$ be respectively the real and complex characteristic varieties of P; then it is not difficult to see that C is the only positive canonical relation which contains diag V and is contained in $\tilde{V} \times \bar{\tilde{V}}$.

Once the model case P is understood we can construct parametrices and projections for systems of equations with noninvolutive characteristics by bringing them back to P by a canonical transformation. That such a canonical relation does exist has been established in the analytic case by Sato-Kawai-Kashiwara **[55]** and in the C^∞ case by Duistermaat-Sjöstrand **[16]** and Bontet de Monvel **[6]**. Applying a canonical transformation ϕ corresponds to conjugating with a Fourier integral operator F whose canonical relation is just the graph of ϕ. The transversality conditions of Theorem 11 are then satisfied, and the theorem implies that FSF^* is a Fourier integral operator with complex phase associated to the canonical relation $\mathcal{C} = \phi \circ C \circ \phi^{-1}$.

The case of $\bar\partial_b$ on a strongly pseudoconvex boundary $b\Omega \subset \mathbf{C}^n$ is one such example. The above procedure leads then to a representation for the Szegö projection S as a Fourier integral operator whose canonical relation $\mathcal{C}$ is easily determined in terms of the characteristic variety of $\bar\partial_b$. Then Theorem 10 allows us to change phases in the formula for S. A defining phase function for $\mathcal{C}$ is calculated to be

$$(x, y, \tau) \to \tau\psi(x, y), \qquad \tau \in \mathbf{R}^+,$$

where $\psi(x, y)$ is an almost analytic extension of the defining function r of $b\Omega$ (i.e., $\psi(x, x) = r(x)/i$, $\psi(x, y) = -\overline{\psi(y, x)}$, $\bar\partial_x\psi$, $\bar\partial_x\psi$ vanish of infinite order for $y = x$), and thus K_S can be written as

$$\int_0^\infty e^{i\tau\psi(x,y)} a(x, y, \tau)\, d\tau.$$

Since $a(x, y, \tau)$ is a classical symbol in τ, Hadamard's principal value formulas immediately yield the well-known asymptotic expansion for the Szegö kernel. We return to these matters in Chapter 12.

CHAPTER 5. ELEMENTARY DIFFERENTIAL GEOMETRY AND THE HEAT EQUATION

The first three sections of this chapter concern the geometry of Riemannian and complex manifolds. Much of this is background material for the later chapters. The fourth and fifth sections introduce the Laplacian $\square = dd^* + d^*d$ on a Riemannian n-manifold M associated to the exterior derivative d and its adjoint d^* and then relate it to the cohomology of the manifold and to the Euler characteristic $\chi(M)$. In fact we prove

$$\begin{aligned}(1)\qquad &\operatorname{Trace}(e^{-t\square_{\text{even}}}) - \operatorname{Trace}(e^{-t\square_{\text{odd}}}) \\ &= \dim(\ker \square_{\text{even}}) - \dim(\ker \square_{\text{odd}}) = \chi(M),\end{aligned}$$

where $\square_{\text{even}}$ ($\square_{\text{odd}}$) is the Laplacian restricted to forms of even (odd) degree.

§6 is concerned with the scalar heat equation (i.e. the heat operator acts on functions rather than forms)

$$(2)\qquad \partial u/\partial t - \Delta_x u = 0, \qquad (x, t) \in M \times [0, \infty), \qquad u|_{t=0} = f.$$

The solution has the form $u(x, t) = \int_M K_t(x, y) f(y) d\operatorname{vol}(y)$, and we shall derive the asymptotics of the kernel $K_t(x, y)$. In particular, we shall see that

$$(3)\qquad K_t(x, x) \sim \sum_{l \geq 0} t^{-n/2+l} \gamma_l(x) \quad \text{for small } t,$$

where the γ_l are polynomials in the curvature tensor and its covariant derivatives.

Since the operator T_t: $f \to u(\cdot, t)$ is formally $e^{-t\Delta}$, we have from (3)

$$(4)\qquad \operatorname{Trace}(e^{-t\Delta}) \sim \sum_{l \geq 0} t^{-n/2+l} \int_M \gamma_l(x) d\operatorname{vol}(x).$$

On the other hand, diagonalizing Δ by its eigenfunctions ψ_k and eigenvalues λ_k, we see that

$$(5)\qquad \operatorname{Trace}(e^{-t\Delta}) = \sum_k e^{-t\lambda_k}.$$

So there is a connection between integrals of "curvatures" $\gamma_l(x)$ and the eigenvalues of the Laplacian.

It is a simple matter to extend these results to the Laplacian on forms. Then by comparing (1) and (4) we obtain (in principle) a formula for the Euler characteristic in terms of curvature, namely

$$(6)\qquad \chi(M) = \int_M P(R_{pqrs/\sigma}) d\operatorname{vol},$$

where P is a polynomial in the indeterminates $R_{pqrs/\sigma}$.

The remainder of the chapter contains the invariant theory necessary to discover the form of $P(\cdot)$. The classical invariant theory of Weyl is sufficient to show that the polynomials corresponding to the heat kernel $K_{t,\square_i}(x, x)$ on forms of degree i are composed of terms of the form

$$\operatorname{Trace}\left[R_{abcd/\sigma} \otimes \cdots \otimes R_{ijkl/\tau}\right].$$

This is discussed in §8. §9 contains a theorem of Gilkey in invariant theory, which is powerful enough to say exactly which polynomial P enters into (6). It turns out that P involves only the curvature and none of its covariant derivatives. P is called the Pfaffian, and (6) is the Chern-Gauss-Bonnet theorem.

1. Riemannian manifolds. Gauss considered the geometry of a surface in, for example, $\mathbf{R}^3$. Locally we we can write the surface as $z = u(x, y)$. In general we can translate and rotate the coordinates so that this takes the form

$$z = \lambda x^2 + \mu y^2 + O\left(|(x, y)|^3\right).$$

The product $K = \lambda\mu$ is called the Gaussian curvature.

EXAMPLE. Consider the sphere of radius A and center $(0, 0, +A)$ in $\mathbf{R}^3$,

$$x^2 + y^2 + (z - A)^2 = A^2.$$

We have

$$z = A\left(1 - \frac{x^2 + y^2}{A^2}\right)^{1/2} - A = \frac{-1}{A}x^2 - \frac{1}{A}y^2 + O\left(|(x, y)|^3\right),$$

so that $K = (-1/A)\cdot(-1/A) = 1/A^2$.

This definition of Gaussian curvature depends prima facie on the embedding of the surface in $\mathbf{R}^3$. Gauss discovered that in fact K depends only on the intrinsic geometry of the surface, not on the embedding. There are many different ways of defining the curvature of a manifold. In this section we give two: the first is very geometric but not obviously intrinsic; the second is obviously intrinsic but not as geometric.

Let

$$x^i = f^i(u^1, \ldots, u^m), \qquad i = 1, \ldots, n, \tag{7}$$

be the components of a parametrized surface $\Sigma \subset \mathbf{R}^n$ and consider a parametrized curve γ, $u^i = g^i(t)$ on the surface. Suppose we have a vector field

$$t \to v(t) \in \mathbf{R}^n$$

along γ such that the vector $v(t)$ lies in the tangent space of Σ at $\gamma(t)$; i.e.

$$v(t) \in \operatorname{span}\{\Sigma_1, \ldots, \Sigma_m\} \quad \text{at } \gamma(t),$$

where Σ_i are the coordinate tangent vectors $\Sigma_i \equiv (\partial f^1/\partial u^i, \ldots, \partial f^n/\partial u^i)$. The vector field $v(t)$ is said to be *parallel in the sense of Levi-Civita* if the derivative dv/dt along the curve is always normal to the surface, i.e.

$$\frac{dv}{dt}\cdot\Sigma_i = 0, \qquad i = 1, \ldots, m.$$

In this case two vector $v(t)$ and $v(t')$ are said to be obtained from one another by *parallel transport*.

EXAMPLE. Vectors in the hypersurface $\Sigma = \mathbf{R}^{n-1} \in \mathbf{R}^n$ that are parallel in the sense of Levi-Civita are parallel in the ordinary sense.

Suppose that X and Y are vector fields on Σ. Using parallel transport we will define a vector field $\nabla_X Y$, called the covariant derivative of Y with respect to X. If $p \in \Sigma$, let $\gamma(t)$ be the integral curve of X through $p = \gamma(0)$ (i.e. $\dot{\gamma}(t) = X(\gamma(t))$) and let Y_t be the parallel transport of $Y(\gamma(t))$ to p. Then

$$(\nabla_X Y)(p) \equiv \lim_{t\to 0} \frac{Y_t - Y_0}{t} \tag{8}$$

satisfies

$$\begin{array}{ll} \text{(i)} & \nabla_X(\alpha Y_1 + \beta Y_2) = \alpha \nabla_X Y_1 + \beta \nabla_X Y_2, \\ \text{(ii)} & \nabla_{\alpha X_1 + \beta X_2} Y = \alpha \nabla_{X_1} Y + \beta \nabla_{X_2} Y, \\ \text{(iii)} & \nabla_{fX} Y = f \nabla_X Y, \\ \text{(iv)} & \nabla_X(fY) = X(f)Y + f\nabla_X Y, \\ \text{(v)} & \nabla_X Y - \nabla_Y X = [X, Y] \text{ (the commutator)}, \\ \text{(vi)} & Z\langle X, Y\rangle = \langle \nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle, \end{array} \tag{9}$$

where α, β are numbers, f is a function, and $\langle \cdot, \cdot \rangle$ is the inner product on $\mathbf{R}^n$. We call

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]$$

the torsion; (v) says that the covariant derivative (8) has no torsion. A covariant derivative satisfying (vi) is called *Riemannian*.

These properties can be used intrinsically to define a covariant derivative on a general Riemannian manifold M; this approach is due to Kozul. First we recall a few facts. Let M be a manifold. If $u^1, \ldots, u^m$ are a fixed set of local coordinates in a coordinate neighborhood $U \subset M$, then a vector field X on U can be expressed uniquely as

$$X = \sum_{k=1}^{m} x^k U_k,$$

where x^k are real-valued functions on U and $U_k = \partial/\partial u^k$. If we evaluate this at a point,

$$X(p) = \sum_{k=1}^{n} x^k(p) U_k,$$

we get a tangent vector; the set $T_p(M)$ of all such tangent vectors is called the tangent space at $p \in M$; a metric is the assignment of a positive-definite (or at least nondegenerate) inner product $\langle \cdot, \cdot \rangle_p$ on each vector space $T_p(M)$ in a manner that varies smoothly with respect to p. If $X = \Sigma x^k U_k$, $Y = \Sigma y^k U_k$ are two vector fields, then for appropriate functions g_{ij} on U we can define

$$\langle X(p), Y(p)\rangle_p = \sum_{i,j=1}^{n} g_{ij}(p) x^i(p) y^j(p), \tag{10}$$

so that g_{ij} gives a metric on $U \subset M$. We abbreviate this as $ds^2 = \Sigma g_{ij}(p) du^i du^j$. If $t \to u(t) = (u^i(t))_{i=1,\ldots,n}$ is a path in the coordinate neighborhood, then the *length* of the path is defined as

$$L = \int ds = \int \sqrt{\sum_{ij} g_{ij}(u(t)) \frac{du^i}{dt} \frac{du^j}{dt}}\, dt.$$

A surface imbedded in $\mathbf{R}^m$ has an obvious induced metric $g_{ij} = \Sigma_l (\partial f^l / \partial u_i)(\partial f^l / \partial u_j)$, which gives the right formula for lengths of curves. The pair $(M, \langle \cdot, \cdot \rangle)$ is called a Riemannian manifold. It is convenient at this point to switch notation and let X_i denote the coordinate vector fields U_i.

By these remarks the covariant derivative is completely determined by the n^3 functions Γ_{ij}^k on U defined by

$$\nabla_{X_i} X_j = \sum_{k=1}^{n} \Gamma_{ij}^k X_k. \tag{11}$$

PROPOSITION 1. *Given a Riemannian manifold* $(M, \langle \cdot, \cdot \rangle)$ *we can uniquely define the* Γ_{ij}^k *so that the covariant derivative* (11) *satisfies the axioms* (9).

PROOF. Define

$$g_{jk} = \langle X_j, X_k \rangle \tag{12}$$

and require

$$\frac{\partial}{\partial x^i}(g_{jk}) = X_i(g_{jk}) = \langle \nabla_{X_i} X_j, X_k \rangle + \langle X_j, \nabla_{X_i} X_k \rangle,$$

so that (9.vi) holds. Now we cyclically permute i, j, k to get 3 equations in the 3 unknowns

$$\langle \nabla_{X_i} X_j, X_k \rangle, \quad \langle \nabla_{X_j} X_k, X_i \rangle, \quad \langle \nabla_{X_k} X_i, X_j \rangle;$$

note there are 3 rather than 6 unknowns since $\nabla_{X_i} X_j = \nabla_{X_j} X_i$ by (9.v). If we solve this system we get

$$\langle \nabla_{X_i} X_j, X_k \rangle = \frac{1}{2} \left(\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k} \right). \tag{13}$$

But by (10) and (11), $\langle \nabla_{X_i} X_j, X_k \rangle = \Sigma_{l=1}^n \Gamma_{ij}^l g_{lk}$; therefore

$$\Gamma_{ij}^l = \sum_{k=1}^{n} \frac{1}{2} g^{kl} \left(\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k} \right), \tag{14}$$

where g^{kl} is the inverse of the matrix g_{lk}. Now extend the definition (10) to expressions of the form

$$\nabla_{(\Sigma_i r^i X_i)} \left(\sum_j s^j X_j \right) \quad \text{for functions } r^i, s^j$$

using the properties (9). It is easy to check explicitly that ∇ satisfies all the required axioms. $\square$

Equations (13) and (14) are known as the first and second Christoffel identities and the connection Γ_{ij}^k is called the *Riemannian connection*.

EXAMPLE 1. Let $M = \mathbf{R}^n$ with the flat metric $\langle \cdot , \cdot \rangle$ (the standard inner product). For $R = \Sigma r^i X_i$, $S = \Sigma s^i X_i$, define

$$\nabla_R S = \sum_{i=1}^{n} (Rs^i) X_i, \tag{15.i}$$

where $Rs^i = \langle (r_1, \ldots, r_n), \nabla s^i \rangle$.

EXAMPLE 2. Let $\Sigma \subset \mathbf{R}^n$ be a hypersurface and define

$$\begin{aligned} \nabla^{\Sigma}_{R(p)} S = \text{projection of } &(\nabla_R S \text{ defined by (15.i)}) \\ &\text{onto the tangent space } T_p(\Sigma). \end{aligned} \tag{15.ii}$$

EXAMPLE 3. Let $\Sigma = \{(x_1, x_2) \in \mathbf{R}^2 \mid x_2 > 0\}$ with $ds^2 = (dx_1^2 + dx_2^2)/x_2^2$ (Poincaré half-plane). We leave it as an exercise to compute Γ^k_{ij}.

Let γ be a curve in M with a tangent vector field T. A vector field Y along γ is called *parallel in the sense of Kozul* if

$$\nabla_T Y = 0 \quad \text{on } \gamma.$$

EXAMPLE. We return to the parametrized surface $\Sigma \subset \mathbf{R}^m$ containing the curve γ,

$$\Sigma\colon x^i = f^i(u^1, \ldots, u^n), \qquad i = 1, \ldots, m,$$
$$\gamma\colon u^i = g^i(t), \qquad i = 1, \ldots, n,$$

and the vector field V with components $(v^1(t), \ldots, v^n(t))$ defined on γ and lying in the tangent space of Σ; γ has a field of tangent vectors $T = (dx^1/dt, \ldots, dx^n/dt)$. The vector field V is parallel along $\gamma \Leftrightarrow \nabla_T V = 0$ along $\gamma \Leftrightarrow$

$$\frac{dv^k}{dt} + \sum_{i,j=1}^{n} \Gamma^k_{ij} \frac{dx^i}{dt} v^j = 0, \qquad k = 1, \ldots, n, \text{ along } \gamma, \tag{16}$$

by (9) and (11). It is not hard to see from this that V is parallel in the sense of Levi-Civita along γ.

Recall that the Lie bracket $[X, Y]$ of two vector fields about p_0 has the following geometric interpretation. Fix a small $t > 0$ and flow along the integral curve of X through p_0 for t units of time to reach p_1; then flow along the integral curve through p_1 of Y for t units of time to reach p_2; then flow along the integral curve through p_2 of $(-X)$ to reach p_3; finally flow along the integral curve through p_3 of $(-Y)$ to reach $\gamma_0(t)$. Define $\gamma(t) = \gamma_0(\sqrt{t})$. Then

$$[X, Y]_p = \left.\frac{d\gamma}{dt}\right|_0.$$

This means that by flowing along the integral curve through $\gamma_0(t)$ of $-[X, Y]_p$ for t^2 units of time, the procedure above gives a closed loop up to second order in t. If we parallel transport a tangent vector Z around this loop, then Z is taken to

$$Z + t^2(\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z) + O(t^3).$$

This suggests that we define the *curvature* $R(X, Y)Z$ of the covariant derivative ∇ to be the vector field

$$R(X, Y)Z = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z, \tag{17}$$

for vector fields X, Y, Z.

PROPOSITION 2. *$R(X, Y)Z$ is a tensor, i.e.*

(i) *the value of $R(X, Y)Z$ at a point $p \in M$, depends only on the values X_p, Y_p, Z_p of the vector fields at this point,*

(ii) *the map*

$$(X_p, Y_p, Z_p) \to R(X_p, Y_p)Z_p$$

is trilinear.

PROOF. Property (ii) is obvious. To verify (i), replace X by fX; then

$$\nabla_X\nabla_Y Z \to f\,\nabla_X\nabla_Y Z, \qquad -\nabla_Y\nabla_X Z \to -(Yf)\nabla_X Z - f\,\nabla_Y\nabla_X Z,$$
$$-\nabla_{[X,Y]}Z \to -f\,\nabla_{[X,Y]}Z + (Yf)\nabla_X Z.$$

Adding these we get

$$R(fX, Y)Z = fR(X, Y)Z,$$

and similarly for $R(X, fY)Z$ and $R(X, Y)(fZ)$. Therefore if $X = \Sigma_i x^i U_i$, $Y = \Sigma_j y^j U_j$, $Z = \Sigma_k z^k U_k$, we have

$$R(X, Y)Z = \sum R\big(x^i U_i, y^j U_j\big)\big(z^k U_k\big) = \sum x^i y^j z^k R(U_i, U_j)U_k$$

and so

$$(R(X, Y)Z)_p = \sum x^i(p)y^j(p)z^k(p)\big(R(U_i, U_j)U_k\big)_p.$$

That is, $(R(X, Y)Z)_p$ depends only on the values of the vector fields X, Y, Z at the point p and not on the values at nearby points. □

We switch notation again and denote the standard coordinate vector fields by X_i. The *Riemann-Christoffel tensor* is defined by

$$R(X_i, X_j)X_k = \sum_{l=1}^{n} R^l_{ijk}X_l. \tag{18}$$

Since $[X_i, X_j] = 0$ for all i, j,

$$R(X_i, X_j)X_k = \nabla_{X_i}\nabla_{X_j}X_k - \nabla_{X_j}\nabla_{X_i}X_k,$$

and we see that the curvature $R(\cdot, \cdot)^{\cdot}$ measures the extent to which the second covariant derivatives $\nabla_{X_i}\nabla_{X_j}$ fail to be symmetric. Using (8) and (11) it is easy, but tedious, to calculate

$$R^l_{ijk} = \frac{\partial\Gamma^l_{ik}}{\partial x^j} - \frac{\partial\Gamma^l_{ij}}{\partial x^k} + \sum_m \left(\Gamma^m_{ik}\Gamma^l_{mj} - \Gamma^m_{ij}\Gamma^l_{mk}\right). \tag{19}$$

The *Ricci tensor* $\mathrm{Ric}(X, Y)$ is defined by

$$\mathrm{Trace}(Z \to R(Z, X)Y), \tag{20.i}$$

so its components are given by

(20.ii)
$$\mathrm{Ric}_{ik} = \sum_{j=1}^{n} R^j_{ijk}.$$

We define the *scalar curvature* to be the trace of Ric_{ik}, that is

(21)
$$R = \sum_{ij} g^{ij} R_{ij},$$

where as always (g^{ij}) is the inverse of (g_{ij}). Finally if we use the inner product to define $g_{im} = \langle X_i, X_m \rangle$, we can define the *Riemann curvature tensor*

(22)
$$R_{ijkl} = \sum_{m=1}^{n} g_{im} R^m_{jkl}.$$

EXAMPLE. We return to the sphere $\Sigma \subset \mathbf{R}^3$ of radius A defined by

$$Z = \frac{-1}{A} x^2 - \frac{1}{A} y^2 + O\left(|(x, y)|^3\right)$$

and its Gaussian curvature $1/A^2$. If we parametrize the sphere by

$$x^1 = A \sin u^1 \cos u^2, \qquad x^2 = A \sin u^1 \sin u^2, \qquad x^3 = A \cos u^1,$$

then the metric g_{ij} is given by (15.iii)

$$g_{11} = A^2, \qquad g_{12} = g_{21} = 0, \qquad g_{22} = A^2 \sin^2 u^1,$$

and after a long calculation using (14), (19), (20) and (21), we find $R = 1/A^2$. We have finally attached an intrinsic significance to the Gaussian curvature, independent of the embedding.

Before leaving the covariant derivative, we show how it can be extended to arbitrary tensors. Suppose we are given an arbitrary tensor, say a 100-tensor

$$A \colon (X_1, \ldots, X_{100}) \to \text{numbers}.$$

We want the *covariant derivative* of A at p to be a 101-tensor $A(X_1, \ldots, X_{100}, Y)$. First let $\gamma(t)$ be the integral curve through $p = \gamma(0)$ of Y and let X_i' be the parallel transport to $\gamma(t)$ of the tangent vector X_i at p. Define

(23.i)
$$\nabla A(X_1, \ldots, X_{100}, Y) = \lim_{t \to 0} \frac{1}{t}\left[A_{\gamma(t)}(X_1', \ldots, X_{100}') - A_p(X_1, \ldots, X_{100})\right].$$

Using an argument like that in Proposition 2, we can check that A is indeed a 101-tensor; it has components

$$(\nabla A)_{i_1 \cdots i_{100} j} = \frac{\partial A_{i_1 \cdots i_{100}}}{\partial x^j} - \sum_{\mu=1}^{100} \sum_{l=1}^{n} \Gamma^l_{i_\mu j} A_{i_1 \cdots i_{\mu-1} l i_{\mu+1} \cdots i_{100}}.$$

We now characterize those paths that locally minimize the distance D between two points on a Riemannian manifold (M, g_{ij}). Let $(x^1, \ldots, x^n)$ be local coordinates about the two points and let

$$x^i = r^i(t), \qquad 0 \leqslant t \leqslant 1,$$

be a path connecting them. We define

$$D(r) = \int_0^1 L(r(t), \dot{r}(t))\, dt \tag{24.i}$$

with

$$L(x, \dot{x}) = \left(\sum_{i,j=1}^{n} g_{ij}(x)\dot{x}^i\dot{x}^j \right)^{1/2}. \tag{24.ii}$$

Note that this is the usual definition of length on $\mathbf{R}^n$ with the flat metric. The paths that locally minimize the distance between the two points will be stationary points of the functional (24). To find these, let

$$x^i = s^i(t), \qquad 0 \leqslant t \leqslant 1,$$
$$s^i(0) = s^i(1) = 0$$

be a smooth path and consider the variation r_σ of r given by

$$x^i = r^i(t) + \sigma s^i(t)$$

and its length

$$D(r_\sigma) = \int_0^1 L(r_\sigma(t), \dot{r}_\sigma(t))\, dt.$$

By integrating by parts we get

$$\begin{aligned} \frac{\partial D(r_\sigma)}{\partial \sigma} &= \int_0^1 \sum_{j=1}^{n} \left[\frac{\partial L}{\partial x^j} s^j + \frac{\partial L}{\partial \dot{x}^j} \dot{s}^j \right] dt \\ &= \int_0^1 \sum_{j=1}^{n} \left[\frac{\partial L}{\partial x^j} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^j} \right) \right] s^j\, dt. \end{aligned}$$

At a stationary point $\partial D(r_\sigma)/\partial\sigma = 0$ for all variations $x^i = s^i(t)$; therefore we get the *Euler-Lagrange equations*

$$\frac{\partial L}{\partial x^j} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^j} \right). \tag{25}$$

Substituting (24.ii) yields

$$\frac{1}{2} L^{-1} \sum_{p,q=1}^{n} \frac{\partial g_{pq}}{\partial x^j} \dot{x}^p \dot{x}^q = \frac{d}{dt}\left(\frac{1}{2} L^{-1} 2 \sum_{q=1}^{n} g_{jq}\dot{x}^q \right).$$

When the curve is parametrized by arclength, L is constant and this reduces to

$$\sum_{p,q=1}^{n} \frac{\partial g_{pq}}{\partial x^j} \dot{x}^p \dot{x}^q = 2 \sum_{r=1}^{n} \frac{\partial g_{jq}}{\partial x^r} \dot{x}^r \dot{x}^q + \sum_{q=1}^{n} g_{jq}\ddot{x}^q;$$

using the expressions (14) for Γ^l_{ij} in terms of the derivatives of the metric gives

$$\ddot{x}^l + \sum_{i,j=1}^{n} \Gamma^l_{ij}\dot{x}^i\dot{x}^j = 0. \tag{26}$$

A curve $\gamma = \{x^i = r^i(t)\}$ satisfying (26) is called a *geodesic*. If we let the vector field V in (16) be the tangent field T of the curve γ, we see that (26) holds $\Leftrightarrow \nabla_T T = 0$ along γ.

We end this section by defining a coordinate system around $p \in M$ in which geodesics through p are straight lines through the origin. Let X be a tangent vector in $T_p(M)$. As long as $|X|$ is small, we can find a geodesic $\gamma = \{x^i = r^i(t)\}$ as a solution of (26) satisfying

$$\gamma(0) = p, \qquad \left.\frac{d\gamma}{dt}\right|_{t=0} = X.$$

The map

$$\exp\colon T_p(M) \to M, \qquad X \to \gamma(1) \tag{27}$$

is a diffeomorphism between a neighborhood of 0 in $T_p(M)$ and a neighborhood of p in M. These local coordinates in M are called *normal coordinates*. The manifold M is called *complete* if exp is defined for all points p and all vectors X. This means that the geodesics continue for all time.

The following observation will be useful later.

PROPOSITION 3. *In a normal coordinate system, the coefficients in the Taylor expansion of $g_{ij}(x)$ about the origin are polynomials in the curvature R_{ijkl} and its covariant derivatives at* 0.

PROOF. At the origin in a normal coordinate system, one has $g_{ij}(0) = \delta_{ij}$, $(\partial g_{ij}/\partial x_k)(0) = 0$, so the derivatives of order 0 and 1 are already known. A long computation with (19), (14) and the definition of normal coordinates shows that

$$R_{ijkl,\mu_1\cdots\mu_s}\big|_0 = (\text{const})\left.\left(\frac{\partial^{s+2} g_{ik}}{\partial x_j \partial x_l \partial x_{\mu_1}\cdots\partial x_{\mu_s}}\right)\right|_0 + \begin{pmatrix}\text{polynomial in lower order}\\ \text{derivatives of the } g_{ab}\end{pmatrix}.$$

From this it is clear by induction that all derivatives of g_{ik} at 0 may be expressed as polynomials in the $R_{ijkl,\mu_1\cdots\mu_s}$. $\square$

2. Complex manifolds. We begin with some preliminaries about complex vector spaces. Suppose that we are given a real vector space V and an automorphism $J\colon V \to V$ satisfying $J^2 = -\mathrm{id}$. Then V becomes a complex vector space if we define $\sqrt{-1}\,v = J(v)$, $v \in V$. Conversely given a complex vector space, we can naturally define a pair (V, J) in the obvious way. We will be concerned below with the case in which V is the tangent space of a manifold at a point.

Let V^* denote the dual space of V consisting of all real-valued linear functions on V; then $V^* \otimes \mathbf{C}$ is the space of all complex-valued $\mathbf{R}$-linear functions on V. When V is the tangent space, $V^* \otimes \mathbf{C}$ becomes the space of complex-valued

1-forms. We say

$$(31)\qquad \begin{aligned} &\phi \in V^* \otimes \mathbf{C} \text{ is of type } (1,0) \quad \text{if } \phi(JZ) = \sqrt{-1}\,Z,\ Z \in V. \\ &\phi \in V^* \otimes \mathbf{C} \text{ is of type } (0,1) \quad \text{if } \phi(JZ) = -\sqrt{-1}\,Z,\ Z \in V. \end{aligned}$$

It is easy to check that

$$(32)\qquad V^* \otimes \mathbf{C} = V_{10} \oplus V_{01},$$

where V_{10} (V_{01}) is the space of all (1,0)-forms ((0,1)-forms). Whenever a vector space V splits, the exterior powers $\Lambda^r V$ inherit the splitting. In our case the splitting (32) induces a splitting

$$(33)\qquad \Lambda^r(V^* \otimes \mathbf{C}) = \bigoplus_{p+q=r} \Lambda^p(V_{10}) \wedge \Lambda^q(V_{01}).$$

An element in $\Lambda^p(V_{10}) \wedge \Lambda^q(V_{01})$ is called a *(p,q)-covector*. Finally we can dualize this construction and define *(p,q)-multivectors*.

Let $U \subset \mathbf{C}^n$ be an open set and $f\colon U \to \mathbf{C}^n$ a C^1 map. We say

$$(34)\qquad \begin{aligned} f \text{ is holomorphic} &\Leftrightarrow (df)_z\colon \mathbf{C}^n \to \mathbf{C}^n \text{ is complex linear, } z \in U \\ &\Leftrightarrow J(df)_z = (df)_z J. \end{aligned}$$

A *complex manifold M* is a manifold with an atlas of coordinate charts $\{(U_\alpha, \phi_\alpha)\colon U_\alpha \to \mathbf{C}^n)\}$ such that the coordinate changes $\phi_\beta \circ \phi_\alpha^{-1}$ are holomorphic, when defined. Now for every point $p \in \mathbf{C}^n$, multiplication by $\sqrt{-1}$ maps a tangent vector at p into a tangent vector at p. By (34) the coordinate changes $\phi_\beta \circ \phi_\alpha^{-1}$ preserve the notion of multiplication by $\sqrt{-1}$; therefore we get a map

$$(35)\qquad J\colon T_p(M) \to T_p(M);$$

in fact, it is not hard to check that J is well defined globally. It is called the *almost complex structure*. If $(z_1 = x_1 + \sqrt{-1}\,y_1, \ldots, z_n = x_n + \sqrt{-1}\,y_n)$ are local coordinates for M and $(\bar z_1 = x_1 - \sqrt{-1}\,y_1, \ldots, \bar z_n = x_n - \sqrt{-1}\,y_n)$ are the conjugate variables, then

$$\frac{\partial}{\partial x_1}, \frac{\partial}{\partial y_1}, \ldots, \frac{\partial}{\partial x_n}, \frac{\partial}{\partial y_n} \text{ is a basis of } T_p(M),$$
$$\frac{\partial}{\partial z_1}, \frac{\partial}{\partial \bar z_1}, \ldots, \frac{\partial}{\partial z_n}, \frac{\partial}{\partial \bar z_n} \text{ is a basis of } T_p(M),$$

with

$$J\colon \frac{\partial}{\partial x_i} \to \frac{\partial}{\partial y_i}, \qquad \frac{\partial}{\partial y_i} \to \frac{-\partial}{\partial x_i},$$
$$J\colon \frac{\partial}{\partial z_i} \to \frac{\partial}{\partial z_i}, \qquad \frac{\partial}{\partial \bar z_i} \to \frac{-\partial}{\partial \bar z_i};$$

the $\partial/\partial z_i$ are (1,0)-vectors, while the $\partial/\partial \bar z_i$ are (0,1)-vectors; finally a (p,q)-form may be written locally as

$$(36)\qquad \sum \phi_{i_1 \cdots i_p j_1 \cdots j_q}\, dz_{i_1} \wedge \cdots \wedge dz_{i_p} \wedge d\bar z_{j_1} \wedge \cdots \wedge d\bar z_{j_q}.$$

We will denote the space of all C^∞ (p,q)-forms on M by $\Lambda^{p,q}(M)$.

Consider a Riemannian metric $g(\cdot\,,\,\cdot)$ on a complex manifold M satisfying

$$g(J_pX, J_pY) = g(X, Y), \tag{37}$$

for all $p \in M$, and $X, Y \in T_p(M)$. If we set $\omega(X, Y) = g(X, JY)$, then

$$h(X, Y) = g(X, Y) + i\omega(X, Y) \tag{38}$$

is a complex-valued sesquilinear form called the *Hermitian metric*, which may be extended to the complexification

$$T_p^c(M) \equiv T_p(M) \otimes \mathbf{C}$$

of the tangent space. Since

$$\omega(X, Y) = g(X, JY) = g(JY, X) = g(J^2Y, JX) = -\omega(Y, X),$$

$\omega(\cdot\,,\,\cdot)$ is an exterior 2-form. It is called the *Kähler form*. If $(z_1,\ldots,z_n)$ are local complex coordinates, then

$$Z_\alpha \equiv \frac{\partial}{\partial z_\alpha} \qquad (\text{type}(1,0)),$$

$$\bar{Z}_\alpha \equiv Z_{\bar{\alpha}} \equiv \frac{\partial}{\partial \bar{z}_\alpha} \qquad (\text{type}(0,1)),\ \alpha = 1,\ldots,n,$$

are a basis for the complex tangent spaces $T_p^c(M)$; we abbreviate this to

$$Z_i, \qquad i = 1,\ldots,n,\bar{1},\ldots,\bar{n}.$$

If we set

$$h_{ij} = h(z_i, z_j),$$

then by (37), $h_{\alpha\beta} = h_{\bar{\alpha}\bar{\beta}} = 0$, while $h_{\alpha\bar{\beta}} = \overline{h_{\bar{\beta}\alpha}}$. For this reason the metric (38) is often written

$$ds^2 = \sum_{\alpha,\beta=1}^{n} h_{\alpha\bar{\beta}}\, dz_\alpha\, d\bar{z}_\beta. \tag{41}$$

Corresponding to ds^2 we have the Kähler form $\omega = \Sigma_{\alpha\beta} h_{\alpha\bar{\beta}}\, dz_\alpha \wedge d\bar{z}_\beta$. This is a globally defined form of type $(1,1)$. For the rest of this section we will use the convention

$$\begin{array}{ll} \text{Roman indices } i, j, k, & 1,\ldots,n,\bar{1},\ldots,\bar{n}, \\ \text{Greek indices } \alpha, \beta, \gamma, & 1,\ldots,n, \\ \bar{\bar{j}} = j. & \end{array}$$

Suppose now that M is equipped with a connection allowing us to parallel transport vectors Z. In general parallel transport does not preserve complex structure; that is

$$\text{parallel transport of } JZ \neq J\,(\text{parallel transport of } Z).$$

But if the Riemannian connection ∇ induced by the metric $g(\cdot\,,\,\cdot)$ according to Proposition 1 satisfies

$$\nabla_X(JY) = J(\nabla_X Y), \tag{42}$$

for all vector fields X, Y, then parallel transport does preserve the complex structure. In this case the Hermitian metric defined by (38) is called *Kähler*.

PROPOSITION 4. *The following are equivalent*:

(i) $\nabla_X(JY) = J(\nabla_X Y)$;

(ii) $d\omega = 0$;

(iii) *there exists a smooth real-valued function* $u(z)$ *such that*

$$h_{\alpha\bar\beta} = \frac{\partial^2 u(z)}{\partial z_\alpha \partial \bar z_\beta}. \tag{43}$$

PROOF. See [38, vol. 2, Chapter 9].

Note that (43) implies

$$\frac{\partial h_{\alpha\bar\beta}}{\partial z_{\bar\gamma}} = \frac{\partial h_{\alpha\bar\gamma}}{\partial z_{\bar\beta}}.$$

EXAMPLE 1. If $M = \mathbf{C}^n$, then

$$g = \sum_{k=1}^{n} |dz_k|^2 = \sum_{k=1}^{n} (dx_k)^2 + (dy_k)^2$$

and

$$h_{\alpha\bar\beta}(z) = \frac{1}{2}\frac{\partial^2}{\partial z_\alpha \partial \bar z_\beta} |z|^2,$$

so that $\mathbf{C}^n$ is Kähler.

EXAMPLE 2. If $M = \mathbf{CP}^n$ (complex projective n-space), then

$$h_{\alpha\bar\beta}(z) = \frac{1}{2}\frac{\partial^2}{\partial z_\alpha \partial \bar z_\beta} \log |z|^2$$

induces a metric called the Fubini-Study metric; therefore $\mathbf{CP}^n$ is also Kähler.

In later chapters we shall study carefully certain Kähler metrics on domains in $\mathbf{C}^n$.

If $h_{\alpha\bar\beta}$ is a Kähler metric on M, then using the defining property (42), we find

$$\begin{aligned} &\nabla_{Z_\alpha} \bar Z_\beta = 0, \qquad \nabla_{Z_\alpha} Z_\beta \text{ is of type } (1,0), \\ &\nabla_{\bar Z_\alpha} Z_\beta = 0, \qquad \nabla_{\bar Z_\alpha} \bar Z_\beta \text{ is of type } (0,1), \end{aligned} \tag{44}$$

when Z_α, $\bar Z_\alpha$ are the coordinate vector fields defined by (39). Because of these formulas, equations involving a Kähler metric are often simpler then their counterparts for a general Hermitian metric.

We illustrate this by calculating the components Γ_{ij}^k of the Riemannian connection of a Kähler manifold M with metric $h_{\alpha\bar\beta}$. The real part of the Hermitian metric $h_{\alpha\bar\beta}$ is a Riemannian metric (compare equation (38)) and by Proposition 1, it is associated with the Riemannian connection ∇ such that

$$\nabla_{Z_i} Z_j = \Gamma_{ij}^k Z_k.$$

Note that the formulas (44) and (9) imply

$$\Gamma^{\gamma}_{\alpha\beta} = \Gamma^{\gamma}_{\beta\alpha}, \qquad \Gamma^{\bar\gamma}_{\bar\alpha\bar\beta} = \Gamma^{\bar\gamma}_{\bar\beta\bar\alpha}, \qquad \text{other } \Gamma^k_{ij} = 0.$$

Now

$$\langle \nabla_{Z_\alpha} Z_\beta, \bar Z_\gamma \rangle = \sum_\delta \Gamma^{\delta}_{\alpha\beta} h_{\delta\bar\gamma}.$$

On the other hand by (9.vi) and (44), we have

$$\langle \nabla_{Z_\alpha} Z_\beta, \bar Z_\gamma \rangle = Z_\alpha \langle Z_\beta, \bar Z_\gamma \rangle = \frac{\partial h_{\beta\bar\gamma}}{\partial z_\alpha}.$$

Comparing these last two equations we see

$$\Gamma^{\gamma}_{\alpha\beta} = \sum_\delta h^{\gamma\bar\delta} \frac{\partial h_{\alpha\bar\delta}}{\partial z_\beta}, \tag{45.i}$$

where $h^{\alpha\bar\beta}$ is the inverse of the matrix $h_{\alpha\bar\beta}$. Similarly we find

$$\Gamma^{\bar\gamma}_{\bar\alpha\bar\beta} = \sum_\delta h^{\bar\gamma\delta} \frac{\partial h_{\bar\alpha\delta}}{\partial \bar z_\beta}. \tag{45.ii}$$

Note how much simpler this expression is compared to expression (14) for the Riemannian case.

The formulas for the Riemannian curvature and Ricci curvature are also much simpler. Instead of the formulas (44), we use the second-order version

$$\nabla_{Z_\alpha}\nabla_{Z_\beta} Z_\gamma = \nabla_{Z_\beta}\nabla_{Z_\alpha} Z_\gamma \qquad \text{(take the inner product of both sides with } \bar Z_\delta \text{ and apply (45.i))},$$

$$\begin{aligned} &\nabla_{\bar Z_\alpha}\nabla_{\bar Z_\beta} Z_\gamma = 0, \\ &\nabla_{\bar Z_\alpha}\nabla_{Z_\beta} Z_\gamma = \nabla_{\bar Z_\alpha}\Big(\sum_\delta \Gamma^{\delta}_{\beta\gamma} Z_\delta\Big) = \sum_\delta \Gamma^{\delta}_{\beta\gamma} Z_\delta, \\ &\nabla_{Z_\beta}\nabla_{\bar Z_\alpha} Z_\gamma = 0; \end{aligned} \tag{46}$$

it turns out that

$$R_{\alpha\bar\beta\gamma\bar\delta} = \sum_\varepsilon \left(\frac{\partial}{\partial \bar z_\beta} \Gamma^{\varepsilon}_{\alpha\gamma} \right) h_{\varepsilon\bar\delta}. \tag{47}$$

When the metric is expressed as

$$h_{\alpha\bar\beta}(z) = \frac{\partial^2}{\partial z_\alpha \partial \bar z_\beta} u(z),$$

we can rewrite this as

$$R_{\alpha\bar\beta\gamma\bar\delta} = u_{\alpha\bar\beta\gamma\bar\delta} - \sum_{\sigma,\tau} u^{\sigma\bar\tau} u_{\alpha\gamma\bar\tau} u_{\bar\beta\bar\delta\sigma}, \tag{48}$$

where $u^{\sigma\bar\tau}$ is the inverse of the matrix $u_{\sigma\bar\tau}$ and

$$u_{\bar\alpha\bar\beta\gamma} \equiv \frac{\partial^3 u}{\partial \bar z_\alpha \partial \bar z_\beta \partial z_\gamma}.$$

Finally using the identity

$$\frac{d}{dx}\log\det(A_{\alpha\bar\beta}) = \operatorname{Trace}\left(\frac{dA}{dx}A^{-1}\right),$$

we find

$$\operatorname{Ric}_{\alpha\bar\beta} = c\frac{\partial^2}{\partial z_\alpha \partial \bar z_\beta}(\log\det h_{\alpha\bar\beta}). \tag{49}$$

We end this section by discussing the exterior differentiation of (p, q)-forms. For $f \in \Lambda^{0,0}(M)$, define

$$\text{(50.i)}\quad \partial f = \sum_{\alpha=1}^{n}\frac{\partial f}{\partial z_\alpha}dz_\alpha \in \Lambda^{1,0}(M), \qquad \bar\partial f = \sum_{\alpha=1}^{n}\frac{\partial f}{\partial \bar z_\alpha}d\bar z_\alpha \in \Lambda^{0,1}(M)$$

and extend these in the standard way:

$$\text{(50.ii)}\quad \begin{aligned} \partial(f\,dz_I\wedge d\bar z_J) &= (\partial f)\wedge dz_I \wedge d\bar z_J \in \Lambda^{p+1,q}(M),\\ \bar\partial(f\,dz_I\wedge d\bar z_J) &= (\bar\partial f)\wedge dz_I \wedge d\bar z_J \in \Lambda^{p,q+1}(M),\end{aligned}$$

for I, J multi-indices with $|I| = p$, $|J| = q$. Finally extend ∂, $\bar\partial$ to all of $\Lambda^{p,q}(M)$ by linearity.

3. Bundles over Riemannian manifolds. In this section we give two additional definitions of connection, including the definition of a connection in a principal fibre bundle, a concept which we will need in Chapter 10.

Let ∇ be a connection on a Riemannian n-manifold $(M, \langle\cdot,\cdot\rangle)$ and let $U \subset M$ be an open set. We call the vector fields $X_1, \ldots, X_n$ an orthonormal frame on U if

$$\text{(51)}\quad \begin{aligned} &\text{(i)}\ \ X_1(p), \ldots, X_n(p) \text{ is a basis of } T_p(M), \text{ all } p \in U,\\ &\text{(ii)}\ \ \langle X_i(p), X_j(p)\rangle_p = \delta_{ij}, \text{ all } p \in U.\end{aligned}$$

Note that we do not necessarily assume that the X_i are coordinate vector fields. In particular, they need not commute. Let ω_i be the dual 1-forms defined by

$$\omega_i(X_j) = \delta_{ij} \tag{52.i}$$

and define the 1-forms ω_{ij} by

$$\nabla_Y X_i = \sum_{j=1}^{n}\omega_{ij}(Y)X_j. \tag{52.ii}$$

The 1-forms ω_i, ω_{ij} are called the *connection forms*.

Let $R(\cdot,\cdot)\cdot$ be the curvature of ∇ (equation (17)) and define the 2-forms Ω_{ij} by

$$R(Y, Z)X_i = \sum_{j=1}^{n}\Omega_{ij}(Y, Z)X_j,$$

i.e. $\Omega_{ij}(Y, Z) = \omega_i([\nabla_Y\nabla_Z - \nabla_Z\nabla_Y - \nabla_{[Z,Y]}]X_j)$. It is easy to check that

$$\Omega_{ij} = d\omega_{ij} + \sum_{k=1}^{n} \omega_{ij} \wedge \omega_{kj}. \tag{53}$$

The 2-forms Ω_{ij} are called the *curvature forms.*

We can reverse this process: given $(M, \langle\cdot, \cdot\rangle)$ and an orthonormal frame $\{X_i\}$, the *Cartan connection* ∇ is defined by (52.ii) and its curvature by (53). We define the torsion $T(\cdot, \cdot)$ of ∇ by

$$T(Y, Z) = \nabla_Y Z - \nabla_Z Y - [Y, Z] \tag{54.i}$$

and the torsion 2-forms by

$$T(Y, Z) = \sum_{i=1}^{n} T_i(Y, Z)X_i. \tag{54.ii}$$

PROPOSITION 5. (i) $d\omega_i = -\sum_{j=1}^{n} \omega_{ij} \wedge \omega_j + T_i$.
(ii) $d\omega_{ij} = -\sum_{k=1}^{n} \omega_{ik} \wedge \omega_{kj} + \Omega_{ij}$.

PROOF. We only prove (i); (ii) follows by a similar argument. Before we begin note that for a 1-form ω and for a 2-form $\omega \wedge \eta$, we have

$$d\omega(Y, Z) = Y\omega(Z) - Z\omega(Y) - \omega([Y, Z]), \tag{55.i}$$

$$(\omega \wedge \eta)(Y, Z) = \omega(Y)\eta(Z) - \eta(Y)\omega(Z). \tag{55.ii}$$

Now from the definition (54) of the torsion,

$$\begin{aligned}\sum_i T_i(Y, Z)X_i &= \nabla_Y Z - \nabla_Z Y - [Y, Z] \\ &= \nabla_Y\Big(\sum_j \omega_j(Z)X_j\Big) - \nabla_Z\Big(\sum_j \omega_j(Y)X_j\Big) - \sum_j \omega_j([Y, Z])X_j \\ &= \sum_j \{Y\omega_j(Z) - Z\omega_j(Y) - \omega_j[Y, Z]\}X_j \\ &\quad + \sum_j \{\omega_j(Z)\omega_{ij}(Y) - \omega_j(Y)\omega_{ij}(Z)\}X_j.\end{aligned}$$

Therefore comparing components and using (55.ii) we have

$$T_i(Y, Z) - \Big(\sum_j \omega_{ij} \wedge \omega_j\Big)(Y, Z) = Y\omega_i(Z) - Z\omega_i(Y) - \omega_i[Y, Z],$$

which by (55.i) becomes

$$d\omega_i(Y, Z) = \Big(-\sum_j \omega_{ij} \wedge \omega_j\Big)(Y, Z) - T_i(Y, Z).$$

REMARK. Equations (i) and (ii) are called the *first and second structural equations.*

There is a closely related construction in which G is a Lie group and $X_1, \ldots, X_n$ are left-invariant vector fields on G which span the tangent space $\tilde{G} = T_{\mathrm{id}}(G)$. We can identify $\tilde{G}$ with the Lie algebra of G. We define the structure constants c_{ij}^k of $\tilde{G}$ by

$$[X_i, X_j] = \sum_{k=1}^{n} c_{ij}^k X_k; \tag{56}$$

they are clearly antisymmetric in i and j. Let $\omega_1, \ldots, \omega_n$ be the dual 1-forms to $X_1, \ldots, X_n$ (equation 52.i). Now

$$d\omega_i(X_j, X_k) = -\omega_i([X_j, X_k]) = -c_{jk}^i,$$

$$-\frac{1}{2}\Big(\sum_{r,s} c_{rs}^i \omega_r \wedge \omega_s\Big)(X_j, X_k) = c_{jk}^i$$

and so,

$$d\omega_i + \sum c_{jk}^i \omega_j \wedge \omega_k = 0, \tag{57}$$

which is the structure equation for the Lie group G.

If we let

$$TM = \bigcup_{p \in M} T_p(M)$$

and let $\pi\colon TM \to M$ be the natural projection, then a vector field may be thought of as a map $s\colon U \subset M \to TM$ satisfying $\pi S = \mathrm{id}$ (i.e. a *section*). We can think of an orthonormal frame $\{X_1, \ldots, X_n\}$ on U in a similar manner. Let

$$E = \left\{ \begin{matrix} (m; e_1, \ldots, e_n)\colon e_1, \ldots, e_n \text{ is an oriented orthonormal} \\ \text{basis of } T_m(M) \end{matrix} \right\}$$

and let $\pi\colon E \to M$ be the natural projection. An orthonormal frame $X_1, \ldots, X_n$ over $U \subset M$ is just a section of $E \to M$.

We now describe some of the structure of $E \to M$. For $g \in \mathrm{SO}(n)$, let

$$R_g\colon E \to E,$$

$$b = (m; e_1, \ldots, e_n) \mapsto bg = \Big(m; \sum g_{i1} e_i, \ldots, \sum g_{in} e_i\Big)$$

then $(bg_1)g_2 = b(g_1 g_2)$, i.e. this is a right action of $\mathrm{SO}(n)$ on E. For a fixed $b_0 \in E$, define

$$f_{b_0}\colon \mathrm{SO}(n) \to E, \qquad g \mapsto bg.$$

We call

$$E_m \equiv \pi^{-1}(m) \subset E$$

the (vertical) fibre. The map f_{b_0} is a diffeomorphism onto the fibre E_m; that is, after choosing a frame b_0, we can identify $\mathrm{SO}(n)$ and the fibre. The derivative of this map is

$$f'_{b_0}\colon T(\mathrm{SO}(n)) \to T(E)$$

in which the image consists of vectors X satisfying

$$\pi_*(X) = 0;$$

these are called vertical vectors. We let

$$V_b = \left\{ X \in T_b(E) \colon \pi_* X = 0 \right\}.$$

These vectors are tangent to the fibre. We can identify $T(\mathrm{SO}(n))$ with the Lie algebra $\mathrm{so}(n)$ of $\mathrm{SO}(n)$. A basis for $\mathrm{SO}(n)$ is given by the matrices e_{ij} containing a 1 in the (i, j)-slot, a -1 in the (j, i)-slot, and 0 elsewhere. Then $X_{ij} \equiv f'_{b_0}(e_{ij})$ are vertical vector fields on E, which are in fact independent of the b_0 used to identify the fibre with $\mathrm{SO}(n)$. Now the X_{ij} do not give a framing of E; indeed, we must somehow choose n additional vector fields $X_1, \ldots, X_n$ to frame E. These are called horizontal vector fields. Unlike the construction of the X_{ij}, these cannot be defined in a canonical manner. Once they are chosen, set

$$H_b = \operatorname{span}\{ X_1(b), \ldots, X_n(b) \};$$

then we expect

(i) $T_b(E) = V_b \oplus H_b$,
(ii) $V_b \cap H_b = \{0\}$,
(iii) $H_{bg} = (R_g)_* H_b$.

The assignment $b \mapsto H_b \subset T_b(E)$ is called a connection.

If Y is a vector field on $U \subset M$, then the assignment $b \to H_b$ gives a *unique* horizontal vector field $\overline{Y}$ on $\overline{U} \equiv \pi^{-1}(U)$ satisfying

$$Y_b \in H_b, \quad \pi_*\left(\overline{X}_b\right) = X_{\pi(b)} \quad \text{for all } b \in \overline{U};$$

it is called the *horizontal lift* of Y. Now if γ is a curve on M with a tangent vector field T, then T lifts to a horizontal vector field $\overline{T}$ (after extending T to some neighborhood of U) and the integral curves of $\overline{T}$ define a curve $\bar{\gamma}$ in E called the lift of γ. A point $b(t) = (m(t);\ e_1(t), \ldots, e_n(t))$ on $\bar{\gamma}$ can be thought of as the parallel transport of the frame $(e_1(0), \ldots, e_n(0))$ at $m = m(0)$ to $m(t)$. This is why the map $b \mapsto H_b$ is called a connection.

In Chapter 10 we will need this construction in greater generality. We define a *principal fibre bundle* with structure group G as a smooth map of manifolds

(58) $$\pi \colon P \to M$$

together with an action of G on P satisfying

(i) $R\colon G \times P \to P$, $(g, p) \mapsto R_g(p) = pg$, with $pg = p$ for all $p \Leftrightarrow g = \mathrm{id}$.
(ii) $M = P/B$.
(iii) If $\{U_\alpha\}$ is a cover of M, then P is locally a product $U \times G$, i.e. there are maps $\psi_U\colon \pi^{-1}(U) \to U \times G, p \mapsto (s_U(p), t_U(p))$ which make the diagram

$$\begin{tikzcd} \pi^{-1}(U) \arrow[r, "\psi_u"] \arrow[dr, "\pi"'] & U \times G \arrow[d, "\text{proj}"] \\ & U \end{tikzcd}$$

commute and satisfy

(59) $$s_U(pg) = (s_U(p))g \quad \text{for all } p \in \pi^{-1}(U).$$

EXAMPLE. We have just shown that the frame bundle $E \to M$ described above is a principal fiber bundle. Let

$$V_p = \left\{ X \in T_p(P) \colon \pi_*(X) = 0 \right\} \tag{60}$$

be the *vertical space*. A *connection* is an assignment $p \mapsto H_p \subset T_p(P)$ of *horizontal spaces* satisfying

$$\begin{aligned} &\text{(i) } H_p \oplus V_p = T_p(P), H_p \cap V_p = \{0\}, \\ &\text{(ii) } H_{pg} = (R_g)_* H_p. \end{aligned} \tag{61}$$

This is almost our final definition of connection. We have had several, including

Coordinates	vector field X_i	$\nabla_{X_i} X_j = \Sigma \Gamma_{ij}^k X_k$ Γ_{ij}^k Christoffel symbols
Frames	orthonormal frame X_i	$\nabla_X X_i = \Sigma \omega_{ij}(X) X_j$ ω_{ij} connection forms
Bundles	section of $P \to E$	$T_p(P) = H_p \oplus V_p$ H_p horizontal space

Observe that on the principal bundle E a connection induces a natural $\mathfrak{T}$-valued one-form, i.e. a natural map from tangent vectors $X \in T_p(E)$ into $\mathfrak{T}$, the Lie algebra of G. In fact, the connection splits X uniquely into horizontal and vertical components, $X = X_h + X_v$ with X_v in the tangent space of the fibre. Since the tangent space of a fibre is naturally identified with the Lie algebra $\mathfrak{T}$, the linear map $X \to X_v \in \mathfrak{T}$ is a Lie-algebra-valued 1-form, which we call ω. Note that the horizontal subspace of $T_p(E)$ is the kernel of ω.

We close the discussion with a simple example showing how connections on frame bundles relate to the construction of canonical forms on Lie groups (equation (57)). The example is as follows.

In $\mathbf{R}^n$ with the standard metric $ds^2 = \Sigma_k dx_k^2$, fix an origin O and an orthonormal frame (e_α). Any Euclidean motion of $\mathbf{R}^n$ carries O to a new origin and (e_α) to a new frame; the new origin and frame uniquely specify the transformation. Therefore the bundle of orthonormal frames on $\mathbf{R}^n$ may be identified with the group E of Euclidean motions. The fibres in E are the cosets of $\mathrm{SO}(n) \subseteq E$, while $\mathbf{R}^n \simeq E/\mathrm{SO}(n)$.

On the Lie algebra level, we have $\mathfrak{E} \simeq \mathrm{so}(n) \oplus \mathbf{R}^n$ as a direct sum of vector spaces, so we have a basis $X_{ij} \in \mathrm{so}(n)$, $X_i \in \mathbf{R}^n$ for $\mathfrak{E}$. The X's may be thought of as left-invariant vector fields on E. In particular, the X_{ij} span the tangent space of a fibre, while the X_i may be considered horizontal. This gives rise to a notion of parallel transport in E which agrees with the obvious parallel transport of frames in $\mathbf{R}^n$. Passing from X_{ij}, X_i to the dual 1-forms ω_{ij}, ω_i, we can check that the structural equations (57) for the Lie group E take the form

$$d\omega_i = -\sum_j \omega_{ij} \wedge \omega_j, \qquad d\omega_{ij} = -\sum_k \omega_{ij} \wedge \omega_{kj}. \tag{62}$$

Comparing with Proposition 5, we see that our flat connection on $\mathbf{R}^n$ has zero curvature and torsion.

This simple example can be used to construct the Levi-Civita connection and curvature on a Riemannian manifold M with general (nonflat) metric. For, any connection in the bundle E_M of orthonormal frames induces the 1-forms ω_i, ω_{ij} on E_M, but we no longer have (62). Rather, the best we can do is to introduce the error terms T_i, Ω_{ij} as in Proposition 5, and try to pick the connection to make them as "close to zero" as possible. Computation shows that the T_i vanish for exactly one connection, which we will call the Levi-Civita connection. Then the Ω_{ij} with respect to the Levi-Civita connection are the natural obstructions to the flatness of M or of the bundle E_M. We call Ω_{ij} the curvature forms on M.

All of this seems unnecessarily abstract when applied to familiar elementary Riemannian geometry. However, using different groups in place of the Euclidean motions and $SO(n)$, we shall have a framework for defining curvatures at boundary points of strictly pseudoconvex domains in $\mathbf{C}^n$. This will be carried out in Chapter 11.

4. The Laplace operator. The Laplace operator Δ on $\mathbf{R}^n$ satisfies

$$\int_{\mathbf{R}^n} \langle \nabla u, \nabla v \rangle \, d\mathrm{vol} = -\int_{\mathbf{R}^n} v \Delta u \, d\mathrm{vol},$$

for u, v compactly supported C^2 functions. On a compact Riemannian manifold M with metric tensor g_{ij} the left-hand side still makes sense; therefore, we define Δ on M by

$$\int_M \left(\sum_{i,j=1}^n g_{ij} \frac{\partial u}{\partial x^i} \frac{\partial u}{\partial x^j} \right) d\mathrm{vol} = -\int_M v \Delta u \, d\mathrm{vol},$$

where

$$d\mathrm{vol} = \sqrt{g}\, dx^1 \cdots dx^n, \qquad \sqrt{g} = \left(\det(g_{ij})\right)^{1/2}$$

is the volume form. A short calculation shows

$$\Delta u = \sum_{i,j=1}^n \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^i} \left(\sqrt{g}\, g^{ij} \frac{\partial}{\partial x^j} \right) u, \tag{63}$$

where $g^{ij} g_{ij} = \delta_{ij}$.

Throughout the remainder of this chapter we let $(M, \langle \cdot, \cdot \rangle)$ be a compact Riemannian n-manifold and let $\Lambda^p(M)$ denote the space of C^∞ p-forms on M. We motivate the definition of the adjoint d^* and the Laplacian $\Box$ by considering the map

$$d^*\colon \Lambda^1(\mathbf{R}^n) \to \Lambda^0(\mathbf{R}^n), \qquad \omega \mapsto \operatorname{div}(\omega^\#),$$

where $\operatorname{div}(\cdot)$ is the divergence of a vector field and $\omega^\#$ is the dual vector field associated to the 1-forms ω by the Euclidean structure. Using some identities from vector calculus, we find

$$\Delta u = d^* du, \qquad \langle (du)^*, v \rangle = \langle u, d^*(v^\#) \rangle,$$

where $\langle\cdot,\cdot\rangle$ is the standard inner product on $\mathbf{R}^n$ and $u, v \in \Lambda^0(M)$. We now give the formal definition. Recall that if V has an inner product, then $\Lambda^p V$ inherits an inner product. In particular we can define an inner product

$$\langle\alpha,\beta\rangle_m, \qquad \alpha,\beta \in \Lambda^p(M),$$

for each $m \in M$ and by integration an inner product on $\Lambda^p(M)$

$$\langle\alpha,\beta\rangle = \int_M \langle\alpha,\beta\rangle_m \, d\mathrm{vol}(m). \tag{64}$$

Then $\Lambda^p(M)$ becomes a pre-Hilbert space with respect to the inner product (64); we let $L_p^2(M)$ denote its completion and let

$$d^*\colon L_p^2(M) \to L_p^2(M) \tag{65}$$

be the Hilbert space adjoint of d. Finally we define the Laplace-Beltrami operator

$$\square = dd^* + d^*d. \tag{66}$$

Harmonic forms (functions) are those annihilated by $\square(\Delta)$.

Let M be a compact manifold. The cohomology of the sequence

$$\Lambda^0(M) \xrightarrow{d} \Lambda^1(M) \xrightarrow{d} \cdots \xrightarrow{d} \Lambda^p(M) \xrightarrow{d} \cdots$$

is isomorphic by DeRham's theorem to the cohomology of the manifold

$$H^p(M) \approx \ker(d\colon \Lambda^p \to \Lambda^{p+1})/d(\Lambda^{p-1}).$$

Let us examine this sequence in more detail. Given $\omega \in \Lambda^p(M)$ with $d\omega = 0$, consider the equation

$$d\lambda = \omega, \qquad \lambda \in \Lambda^{p-1}(M). \tag{67.i}$$

Note that λ is not unique. Indeed, since $d^2 = 0$, $\lambda + d\mu$, for any $\mu \in \Lambda^{p-2}(M)$ is also a solution. Among all such solutions λ of this equation, we can single out one by requiring $\lambda \perp \operatorname{Ran} d$ (with respect to the inner product (63)); since $0 = \langle\lambda, d\mu\rangle = \langle d^*\lambda, \mu\rangle$, for $\mu \in \Lambda^{p-2}(M)$, this is equivalent to

$$d^*\lambda = 0. \tag{67.ii}$$

Applying $\square$ to the solution λ of the system (67) yields

$$\square\lambda = (d^*d + dd^*)\lambda = d^*\omega + d(0) = d^*\omega.$$

In other words, the system (67) leads naturally to the equation

$$\square\lambda = \alpha, \qquad \lambda, \alpha \in \Lambda^p(M). \tag{68}$$

If we introduce local coordinates we can check that $\square$ is a second-order elliptic operator in the sense that it satisfies hypotheses (i) and (ii) of Theorem 7 in Chapter 3. We can also check that it is selfadjoint.

THEOREM 1. (i) $L_p^2(M) = \operatorname{Ker}\square \oplus \operatorname{Ran}\square$.
(ii) $\{$*harmonic p-forms*$\} \subseteq \Lambda^p(M)$.
(iii) $H^p(M) \approx \{$*harmonic p-forms*$\}$.
(iv) $\dim H^p(M) < \infty$.

PROOF. As we remarked, Theorem 7 of Chapter 3 applies. Therefore $L_p^2(M) = \operatorname{Coker} \Box \oplus \operatorname{Ran} \Box$, with $\dim(\operatorname{Coker} \Box) < \infty$ and $\dim(\operatorname{Ker} \Box) < \infty$; moreover $\operatorname{Ker} \Box$ consists of smooth functions. This gives (ii). Since $\Box$ is selfadjoint, $\operatorname{Coker} \Box = \operatorname{Ker} \Box^* = \operatorname{Ker} \Box$ and we get (i).

Since $\Box = (d + d^*)(d + d^*)$, $\Box \omega = 0$ implies $(d + d^*)\omega = 0$. Therefore $0 = \langle (d + d^*)\omega, d\omega \rangle = \|d\omega\|^2 + \langle \omega, dd\omega \rangle$ and $0 = \langle (d + d^*)\omega, d^*\omega \rangle = \|d^*\omega\|^2 + \langle dd\omega, \omega \rangle$, so that $d\omega = 0$ and $d^*\omega = 0$; in particular, harmonic forms are closed. Therefore the map

$$j\colon \{\text{harmonic } p\text{-forms}\} \to H^p(M), \qquad \omega \mapsto \omega + \operatorname{Ran} d$$

is well defined. To show that j is injective, take $\omega \in \operatorname{Ran} d \cap \operatorname{Ker} \Box$. Then $\omega = d\lambda$, for some λ with $\Box(d\lambda) = 0$; that is, $dd^*d\lambda = 0$. Now $0 = \langle dd^*d\lambda, d\lambda \rangle = \|d^*d\lambda\|^2$, so $d^*d\lambda = 0$. Similarly $0 = \langle d^*d\lambda, \lambda \rangle = \|d\lambda\|^2$, so $d\lambda = 0$. This gives $\omega = 0$ and j is injective as desired. We now show j is surjective. By (i) any closed form β may be written as $\beta = \omega + \Box \alpha$ with $\Box \omega = 0$. Now $d\omega = 0$ and $d\beta = 0$, so $0 = d\Box\alpha = d(dd^* + d^*d)\alpha = dd^*d\alpha$. Therefore $0 = \langle dd^*d\alpha, d\alpha \rangle = \|d^*d\alpha\|^2$, which implies $\Box\alpha = d(d^*\alpha) \in \operatorname{Range} d$. So $\beta = \omega + d(\text{something})$, with ω harmonic, and j is surjective. The proof of (iii) is complete. Part (iv) is an immediate consequence of the finiteness of $\{\text{harmonic } p\text{-forms}\}$ and (iii). $\Box$

COROLLARY. *The Euler characteristic*

$$\chi(M) = \sum_{i=0}^{n} (-1)^i \dim H^i(M)$$

is finite.

5. The Euler characteristic. In this section we show how we can express the Euler characteristic $\chi(M)$ in terms of the action of $\Box$ on the space of forms. Let

$$\Lambda_{\text{even}}(M) = \bigoplus_{p \text{ even}} \Lambda^p(M), \qquad \Lambda_{\text{odd}}(M) = \bigoplus_{p \text{ odd}} \Lambda^p(M).$$

Consider the operator

$$L = d + d^*\colon \Lambda_{\text{even}}(M) \to \Lambda_{\text{odd}}(M).$$

Since

$$L^*L = \Box|_{\Lambda_{\text{even}}(M)}, \qquad LL^* = \Box|_{\Lambda_{\text{odd}}(M)}, \tag{69}$$

by Theorem 1 we get

$$\dim \operatorname{Ker}(L^*L) = \sum_{p \text{ even}} \dim H^p(M),$$

$$\dim \operatorname{Ker}(LL^*) = \sum_{p \text{ odd}} \dim H^p(M),$$

and so

$$\dim \operatorname{Ker}(L^*L) - \dim \operatorname{Ker}(LL^*) = \sum_{p=1}^{n} (-1)^p \dim H^p(M) = \chi(M). \tag{70}$$

If $0 \leqslant \lambda_1 \leqslant \lambda_2 < \cdots$ are the eigenvalues, listed according to their multiplicities, of $\square_{\text{even}}$ acting on $\lambda_{\text{even}}(M)$, then

(71.i) $$\operatorname{Trace}(e^{-t\square_{\text{even}}}) = \sum e^{-t\lambda_j}$$

and similarly

(71.ii) $$\operatorname{Trace}(e^{-t\square_{\text{odd}}}) = \sum e^{-t\mu_j},$$

where $0 \leqslant \mu_1 \leqslant \mu_2 \leqslant \cdots$ are the eigenvalues of $\square_{\text{odd}}$. The relation between the λ's and the μ's is given by

LEMMA 1. *Assume that L^*L has eigenvalues λ_k, that LL^* has eigenvalues μ_k and that the spectra of both operators are discrete. Then if those λ_k and μ_k which equal* 0 *are deleted, the list of the λ_k's and the μ_k's coincide.*

PROOF. Let $\lambda \neq 0$ be an eigenvalue of L^*L; then

$$L^*L\phi = \lambda\phi,$$

for some nonzero ϕ. This implies $L\phi \neq 0$ and

$$LL^*(L\phi) = \lambda(L\phi);$$

i.e. λ is an eigenvalue of LL^* as required. The same argument shows that the multiplicity of λ for LL^* is at least as great as for L^*L. Since the argument is symmetric, we are done. $\square$

Because the Laplacian acting on a compact manifold has a discrete spectrum, we can apply the lemma with $L = d + d^*$. In this case the number of 0 eigenvalues of L^*L (respectively LL^*) is equal to $\dim(\operatorname{Ker} \square_{\text{even}})$ (respectively $\dim(\operatorname{Ker} \square_{\text{odd}})$). If we take note of equations (69) and (70), this gives

PROPOSITION 6.

$$\operatorname{Trace}(e^{-t\square_{\text{even}}}) - \operatorname{Trace}(e^{-t\square_{\text{odd}}}) = \dim(\operatorname{Ker} \square_{\text{even}}) - \dim(\operatorname{Ker} \square_{\text{odd}}) = \chi(M).$$

6. Asymptotics of the heat kernel. To simplify the notation, we work primarily with functions and the Laplacian Δ, rather than forms and the Laplacian $\square$. In this section we will construct an asymptotic approximation to the fundamental solution of the heat equation

(72) $$(\partial/\partial t - \Delta_x)u(t, x, y) = 0, \qquad u(t, x, y)|_{t=0} = \delta_y(x)$$

on $M \times (0, \infty)$ (the *heat kernel*).

We begin by considering $M = \mathbf{R}^n$ and

$$u(t, x, y) = (4\pi t)^{-n/2} e^{-|x-y|^2/4t}.$$

Setting $y = 0$ and $a = -n/2$, we find

(73) $$\Delta u = (4\pi t)^a e^{-|x|^2/4t}\left\{\frac{|\nabla |x|^2|^2}{16t^2} - \frac{\Delta |x|^2}{4t}\right\},$$
$$\frac{\partial u}{\partial t} = (4\pi t)^a e^{-|x|^2/4t}\left\{\frac{|x|^2}{4t^2} - \frac{a}{t}\right\},$$

and since $|\nabla |x|^2|^2 = 4|x|^2$, $\Delta |x|^2 = 2n$, we see that u is a solution of (72) with $y = 0$ and that

$$K_t(x, y) = u(t, x, y)$$

is the required heat kernel. Note that this is an exact expression for the heat kernel and not an asymptotic one.

This suggests that for a general Riemannian manifold $(M, \langle\cdot, \cdot\rangle)$ we look for a solution of the heat equation with an asymptotic expansion of the form

$$(74) \qquad u(t, x, y) \sim (4\pi t)^{-n/2} e^{-\operatorname{dist}^2(x, y)/4t}\Big\{\sum_{i\geq 0} \phi_i(x, y)t^i\Big\}, \qquad t \to 0^+ .$$

Here dist(x, y) is the *geodesic distance* induced by the metric $\langle\cdot, \cdot\rangle$, i.e. the infimum of the length of all geodesics connecting x and y. This expansion is due to Minakshisundaram and Pleijel [**47**].

We proceed now to derive (74). Let y^i be a system of normal coordinates about the point x (equation (27)). Recall that in this coordinate system geodesics $t \to y^i(t)$ passing through x are simply straight lines; therefore the geodesic distance dist(x, y) is simply the length of the line connecting x and y. Put $r = \operatorname{dist}(x, y)$ (we use this notation throughout the remainder of the section) and consider a function Ψ on M that depends only upon the geodesic distance r. A short calculation (see [**3**, p. 134]) shows that

$$\Delta(\Psi) = \frac{d^2\Psi}{dr^2} + \frac{n-1}{r}\frac{d\Psi}{dr} + \frac{d\log\sqrt{g}}{dr}\frac{d\Psi}{dr}.$$

This in turn gives

$$(75) \quad \Delta(\Phi\Psi) = \Phi\left(\frac{d^2\Psi}{dr^2} + \frac{n-1}{r}\frac{d\Psi}{dr} + \frac{d\log\sqrt{g}}{dr}\frac{d\Psi}{dr}\right) + 2\frac{d\Phi}{dr}\frac{d\Psi}{dr} + \Psi\Delta\Phi,$$

for Φ an arbitrary function.

Put

$$u_N(t, x, y) = (4\pi t)^{-n/2} e^{-r^2/4t}\{\phi_0 + \phi_1 t + \cdots + \phi_N t^N\},$$
$$\Psi = (4\pi t)^{-n/2} e^{-r^2/4t}, \qquad \Phi = \{\phi_0 + \phi_1 t + \cdots + \phi_N t^N\}$$

and substitute this into (75):

$$\Delta(\Phi\Psi) = \Psi\Bigg\{\left(\frac{r^2}{4t^2} - \frac{1}{2t} - \frac{n-1}{2t} - \frac{d\log\sqrt{g}}{dr}\frac{r}{2t}\right)$$
$$- \frac{r}{t}\left(\frac{d\phi_0}{dr} + t\frac{d\phi_1}{dr} + \cdots + t^N\frac{d\phi_N}{dr}\right) + \left(\Delta\phi_0 + t\Delta\phi_1 + \cdots + t^N\Delta\phi_N\right)\Bigg\};$$

this implies

$$(76)\quad \left(\frac{\partial}{\partial t}-\Delta_x\right)u_N(t,x,y)$$
$$=\Psi\sum_{i=0}^{N}\left\{\left(\frac{r^2}{4t^2}+\frac{i}{2}-\frac{n}{2t}-\frac{r^2}{4t^2}+\frac{1}{2t}+\frac{n-1}{2t}+\frac{d\log\sqrt{g}}{dr}\frac{r}{2t}\right)\right.$$
$$\left.\cdot t^i\phi_1+t^{i-1}r\frac{d\phi_i}{dr}-t^i\Delta_x\phi_i\right\}.$$

To make this 0, we set the coefficient of Ψt^{i-1} equal to 0; i.e. we require $\phi_1\equiv 0$,

$$(77)\qquad r\frac{d\phi_i}{dr}+\frac{r}{2}\frac{d\log\sqrt{g}}{dr}\phi_i+i\phi_i=\Delta\phi_{i-1},\qquad i=0,1,\ldots,N.$$

We claim that there are C^∞ solutions to these equations. Indeed, if we write (77) as

$$r\frac{d}{dr}\left(r^ig^{1/4}\phi_i\right)=r^ig^{1/4}\Delta_x\phi_i,$$

then we get such solutions by setting

$$\phi_0(x,y)=\left(\frac{g(x)}{g(y)}\right)^{1/4}$$

and

$$\phi_i(x,y)=\frac{\phi_0(x,y)}{r^i(x,y)}\int_x^y\frac{r^{i-1}(x,z)\Delta_z\phi_{i-1}(x,z)}{\phi_0(x,z)}\,dr(x,z),\qquad i=1,\ldots,N,$$

where $r(x,y)=\operatorname{dist}(x,y)$. We can check by induction on i that ϕ_i is smooth. This is clear for ϕ_0. If ϕ_{i-1} is smooth, then in a normal coordinate system about x, $\phi_i(x,y)$ has the form

$$\Phi(y)=\frac{1}{|y|^i}\int_0^{|y|}r^{i-1}g\left(r\frac{y}{|y|}\right)dr=\int_0^1t^{i-1}g(ty)\,dt\quad\text{with } g\in C^\infty.$$

Evidently $\Phi\in C^\infty$ also, so ϕ_i is smooth. Note that with the ϕ_i defined this way,

$$(78)\qquad (\partial/\partial t-\Delta_x)u_N(t,x,y)=(4\pi t)^{-n/2}e^{-r^2/4t}\Delta\phi_Nt^N.$$

This completes the construction of the asymptotic solution (74).

Now put

$$(79.\text{i})\qquad K_t^N(x,y)=\theta(\operatorname{dist}(x,y),t)u_N(t,x,y),$$

where $\theta(\cdot,\cdot)\in C_0^\infty$ is a cutoff function with

$$(79.\text{ii})\qquad \theta(x,t)=\begin{cases}1 & \text{near the origin,}\\ 0 & \text{outside a small neighborhood of the origin.}\end{cases}$$

We define a *parametrix* $K(t, x, y)$ for the (scalar) heat operator by the properties

$$(80)\quad \begin{array}{rl} \text{(i)} & K(t, x, y) \in C^\infty(M \times M \times (0, \infty)), \\ \text{(ii)} & (\partial/\partial t - \Delta_x)K(t, x, y) \text{ and its derivatives up to high order} \\ & \text{tend to zero uniformly as } t \to 0^+, \\ \text{(iii)} & \lim_{t \to 0^+} K(t, \cdot, y) = \delta_y(\cdot). \end{array}$$

A parametrix $K^p(t, x, y)$ for the heat operator $\partial/\partial t - \Box_p$ acting on $\Lambda^p(M)$ is defined similarly.

THEOREM 2. (i) *For large N, a parametrix for the scalar heat operator is given by*

$$K_t^N(x, y) = \theta(r, t)(4\pi t)^{-n/2} e^{-r^2/4t}\{\phi_0(x, y) + t\phi_1(x, y) + \cdots + t^N\phi_N(x, y)\}$$

where $r = \text{dist}(x, y)$.

(ii) *On the diagonal $x = y$, the $\phi_i(x, y)$ are real-valued polynomial functions of the derivatives $\partial^\alpha g_{ij}$ of the metric tensor in normal coordinates about x.*

(iii) *There is a similar expansion for a parametrix*

$$K^{p,N}(x, y) = \theta(r, t)(4\pi t)^{-n/2} e^{-r^2/4t}\{\phi_0^p(x, y) + \cdots + \phi_N^p(x, y)t^N\}.$$

Here $\phi_i(x, y)$ is an endomorphism from $\Lambda^p T^*M|_y$ into $\Lambda^p T^*M|_x$ with $T^*M =$ cotangent bundle, and ϕ_i is smooth in $(x, y) \in M \times M$. In local coordinates, this makes ϕ_i a matrix of smooth functions of x, y.

(iv) *On the diagonal $x = y$ in a local coordinate system, the entries of the matrix ϕ_i^p are polynomial functions of the derivatives $\partial^\alpha g_{ij}$ of the components of the metric tensor.*

In order to use this we need the following lemma about the regularity of solutions to the heat equation

$$(81)\qquad (\partial/\partial t - \Delta)u = g, \qquad u|_{t=0} = 0$$

on $M \times (0, \infty)$.

LEMMA 2. *If the datum g satisfies*

(i) $\Delta^N g \in L^2$, *for some large N,*

(ii) *g vanishes to order N^1 in t at $t = 0$;*

then the solution u satisfies

(iii) $\Delta^P u \in L^2$, *for some large P, depending upon N, N^1,*

(iv) *u vanishes to order P^1, depending upon N, N^1.*

PROOF. By expanding the data and solution in the eigenfunctions of Δ, we can reduce the lemma to familiar questions about ordinary differential equations in t. $\Box$

The lemma tells us that if the data is very smooth and vanishes to high order, then the same is true of the solution. Consequently, our parametrix (which is annihilated to high order by the heat operator) must agree to high order with the exact solution of the heat equation.

PROOF OF THEOREM 2. We prove only the statements about the scalar heat operator; the corresponding statements for the heat operator $\partial/\partial t - \square_p$ follow by a similar argument.

PROOF OF (i). We must check that $K_t^N(x, y)$ satisfies the properties (80.i), (80.ii), (80.iii) of a parametrix. Property (80.i) is obvious; (80.ii) follows at once from the formal calculations used to construct the $\phi_i(x, y)$; and (80.iii) follows from a short calculation which we omit.

PROOF OF (ii). We work in normal coordinates about $x = 0$. Say that $f(y)$ is *polynomially determined* if all derivatives $\partial^\beta f(0)$ are polynomials in $\partial^\alpha g_{ij}(0)$. Since $g_{ij}(0) = \delta_{ij}$, manipulations with Taylor series show that (g^{ij}) and $g^{\pm 1/2}$ are polynomially determined. Thus $\phi_0(0, y)$ is polynomially determined. Again by manipulating Taylor series, we can check by induction on i that $\phi_i(0, y)$ is polynomially determined. In particular, $\phi_i(0, 0)$ is a polynomial in $\partial^\alpha g_{ij}(0)$, and this is the assertion of (ii). Note that we used only $g_{ij}(0) = \delta_{ij}$, not the full force of the normal coordinate system. This will be needed in Gilkey's invariant theory (Theorem 9). $\square$

The properties (80) that define the parametrix of the heat operator are useful for many purposes, but very often a stronger notion is needed. A *fundamental solution* $H(t, x, y)$ of the scalar heat equation is defined by

$$(81)\qquad \begin{aligned} &\text{(i)} && H(t, x, y) \text{ is } C^1 \text{ in } t \text{ and } C^2 \\ & && \text{in } x \text{ and } y \ (t > 0), \\ &\text{(ii)} && (\partial/\partial t - \Delta_x) H(t, x, y) = 0, \\ &\text{(iii)} && \lim_{t \to 0^+} H(t, \cdot, y) = \delta_y(\cdot). \end{aligned}$$

There is an analogous definition for the fundamental solution $H^p(t, x, y)$ of the heat operator $\partial/\partial t - \square_p$ acting on $\Lambda^p(M)$. Again we note the following as a consequence of Lemma 2: If $H(t, x, y)$ is the fundamental solution and $K(t, x, y)$ is a parametrix for the heat equation, then $H(t, x, y) - K(t, x, y)$ has many derivatives on $[0, \infty) \times M \times M$ and vanishes to high order at $t \to 0^+$. So K is an excellent approximation to H.

THEOREM 3. (i) *Let f_i be an orthonormal basis of eigenfunctions on M and let λ_i be the corresponding eigenvalues. Then the fundamental solution of the scalar heat equation may be written*

$$H_t(x, y) = \sum e^{-\lambda_i t} f_i(x) \overline{f_i(y)}.$$

In particular

$$\sum e^{-\lambda_i t} = \int_M H_t(x, x)\, d\mathrm{vol}$$

is well defined.

(ii) *As $t \to 0^+$, $H_t(x, y) \sim K_t(x, y)$.*

(iii) $H_t^p(x, y) = \sum e^{-\lambda_i t} f_i^p(x) \otimes f_i^p(y)$, $H_t^p(x, y) \sim K_t^p(x, y)$, *as $t \to 0^+$, where f_i^p are the normalized eigenfunctions of $\square_p$ on M with eigenvalues λ_i.*

PROOF. The proof is straightforward. See [3] for details. $\square$

7. The Chern-Gauss-Bonnet theorem. In this section we pull together the last several sections and show what the asymptotics of the heat kernel have to do with the Euler characteristic. We begin by reviewing what we know. By Theorem 3 the heat operator $\partial/\partial t - \square_p$ has a fundamental solution

$$(82)\quad H_t^p(x, y) \sim \theta(r, t)(4\pi t)^{-n/2} e^{-r^2/4t}\{\phi_0^p(x, y) + t\phi_1^p(x, y) + t^2\phi_2^p(x, y) + \cdots\},$$

where θ is the cutoff function (79.ii) and $r = \operatorname{dist}(x, y)$. By Theorem 2 the coefficients of the ϕ_i^p on the diagonal $x = y$ are polynomials in the derivatives $\partial^\alpha g_{ij}$ of the components of the metric tensor g_{ij}. Assume now that we introduce normal coordinates; then by Proposition 3 the $\partial^\alpha g_{ij}$ are polynomials in the covariant derivatives $R_{pqrs/\sigma}$ of the Riemann curvature tensor R_{pqrs} and, hence, so are the ϕ_i^p.

On the other hand note that because $H_t^p(x, y)$ is a fundamental solution of $\partial/\partial t - \square_p$ it satisfies

$$(83)\qquad \operatorname{Trace}(e^{-t\square_p}) = \int_M H_t^p(x, x)\, d\operatorname{vol}(x).$$

Now by Proposition 7,

$$(84)\qquad \chi(M) = \operatorname{Trace}(e^{-t\square_{\text{even}}}) - \operatorname{Trace}(e^{-t\square_{\text{odd}}}).$$

After comparing (82), (83) and (84), we see that

$$(85.\text{i})\qquad \chi(M) = \int_M P(R_{pqrs/\sigma})\, d\operatorname{vol},$$

where P is some polynomial in the $R_{pqrs/\sigma}$.

The rest of this section is a heuristic description of how we can use invariant theory to discover the form of $P(\cdot)$. Later, after we have developed the necessary invariant theory, we will be able to show that $P(\cdot)$ is the Pfaffian

$$(85.\text{ii})\quad \operatorname{Pff}(R_{ijkl})\, dx^1 \wedge \cdots \wedge dx^n = c \sum \operatorname{sgn}(i)\operatorname{sgn}(j) R_{i_1 i_2 j_1 j_2}\, dx_{j_1} \wedge dx_{j_2} \wedge \cdots \wedge R_{i_{2k-1} i_{2k} j_{2k-1} j_{2k}}\, dx_{j_{2k-1}} \wedge dx_{j_{2k}},$$

where $\operatorname{sgn}(i) = \operatorname{sgn}(i_1, i_2, \ldots, i_{2k-1}, i_{2k})$. The relation (85) for the Euler characteristic is the Chern-Gauss-Bonnet theorem.

Consider the parametrix for the scalar heat equation

$$(86)\qquad K_t(x, x) = \sum_{i \geqslant 0} (4\pi t)^{-n/2} \phi_i(x, x) t^i.$$

Let $(M, \langle\cdot,\cdot\rangle)$ be a Riemannian manifold and let M' be the same space but equipped with the metric

$$\langle\cdot,\cdot\rangle' = \lambda\langle\cdot,\cdot\rangle,$$

obtained by dilating $\langle\cdot,\cdot\rangle$. Let $\{e_i\}$ be an orthonormal frame on M. Then $\{e_i' = \lambda^{-1/2}e_i\}$ is an orthonormal frame on M' and $\{e_i^{*\prime} = \lambda^{1/2}e_i^*\}$ is the corresponding coframe. We have

$$R'_{ijkl} = \langle R'(e_i', e_j')e_k', e_l'\rangle = \lambda^{-1}\langle R(e_i, e_j)e_k, e_l\rangle = \lambda^{-1}R_{ijkl}$$

and similarly

$$R'_{ijkl/\sigma_1\cdots\sigma_m} = \lambda^{-m/2-1}R_{ijkl/\sigma_1\cdots\sigma_m}, \tag{87}$$

$$\phi_l'(x, y) = \lambda^{-1}\phi_i(x, y). \tag{88}$$

Now $\phi_l(x, y)$ is a polynomial in the $R_{ijkl/\sigma_1\cdots\sigma_m}$; that is, a sum of monomials of the form

$$R_{ijkl/\sigma_1\cdots\sigma_m} \otimes \cdots \otimes R_{i'j'k'l'/\sigma_1'\cdots\sigma_m'} \qquad (p \text{ terms}). \tag{89}$$

In order for $\phi_l(x, y)$ to transform correctly after a scale change

$$e_i \mapsto e_i' = \lambda^{-1/2}e_i,$$

we see, by comparing (87) and (88), that each monomial term (88) that occurs in $\phi_l(x, y)$ must satisfy

$$\sum_{\text{factors } \sigma} 2 + |\sigma| = 2l, \tag{90}$$

where the sum is over the p factors $R_{ijkl/\sigma_1\cdots\sigma_m}$, and where $|\sigma|$ denotes the length of σ, i.e. the number of indices.

Where $l = 0$, this formula tells us that $\phi_0(x, x)$ must be a constant, say c_0; (86) becomes

$$K_t(x, x) = (4\pi t)^{-n/2}\{C_0 + \phi_1(x, x)t + \cdots\}.$$

Now, orthogonal transformations preserve normal coordinates and hence the invariants $\phi_l(x, x)$ are defined by $O(n)$-invariant polynomials in $R_{ijkl/\sigma}$. In the next section we prove the following

Fact. The $O(n)$-invariants ϕ_l must be a sum of terms of the form

$$\operatorname{Trace}\left(R_{ijkl/\sigma} \otimes \cdots \otimes R_{i'j'k'l'/\sigma'}\right).$$

When $l = 1$, this fact together with the weight formula (90) implies that each monomial contains just one factor R_{ijkl}. Therefore

$$\phi_1 = (\text{const})R,$$

where $R = \sum g^{ik}g^{jl}R_{ijkl}$ is the scalar curvature (21). The parametrix now takes the form

$$K_t(x, x) = (4\pi t)^{-n/2}\{C_0 + C_1Rt + \cdots\}.$$

When $l = 2$, we need to consider $\Sigma(2 + |\sigma|) = 4$. Each monomial must involve one factor with $|\sigma| = 2$ or two factors with $|\sigma| = 0$. It turns out that only the term (const)ΔR arises from $R_{ijkl/\sigma_1\sigma_2}$ and only R^2, $\|R_{ijkl}\|^2$ and $\|\text{Ric}\|^2$ (see (20), (21)) arise from R_{ijkl}, $R_{i'j'k'l'}$. This means that

$$\phi_2 = C_2 R^2 + C_3 \|\mathrm{Ric}\|^2 + C_4 \|R_{ijkl}\|^2 + C_5 \Delta R$$

and

$$K_t(x, x)$$
$$= (4\pi t)^{-n/2}\left\{C_0 + C_1 R t + \left[C_2 R^2 + C_3 \|\mathrm{Ric}\|^2 + C_4 \|R_{ijkl}\|^2 C_5 \Delta R\right] t^2 + \cdots\right\}.$$

We can determine the constants C_i by making explicit comptuations for simple manifolds. Note that although the C_i do not depend on the particular manifold, they do depend on the dimension of the manifold.

This type of information is all we can determine from a single heat kernel $K_t(x, x)$. On the other hand, for the alternating sum of heat kernels associated with the operator

$$e^{-t\Box_{\text{even}}} - e^{-t\Box_{\text{odd}}}$$

some remarkable cancellations occur. It turns out that the resulting polynomial depends only upon the Riemann curvature tensor R_{ijkl} and not on its covariant derivatives $R_{ijkl/\sigma}$. When n is odd, this polynomial is identically 0; when n is even the polynomial is the Pfaffian, which we now define.

Let A_{ij} be a matrix on $\mathbf{R}^{2k}$ with $A_{ij} = -A_{ji}$. It can be shown that $\det(A_{ij})$ is the square of a polynomial. The Pfaffian $\mathrm{Pf}(A_{ij})$ associated with A_{ij} is defined by

$$\det(A_{ij}) = \left[\mathrm{Pf}(A_{ij})\right]^2.$$

When A_{ij} is a two-tensor on $\mathbf{R}^{2k}$, we define the associated Pfaffian by

$$\mathrm{Pf}(A_{ij}) = C_n \sum \mathrm{sgn}(i_1 \cdots i_{2k}) A_{i_1 i_2} A_{i_3 i_4} \cdots A_{i_{2k-1} i_{2k}},$$

where $\mathrm{sgn}(\cdot)$ is the sign of the permuation if all the i_j are different and 0 otherwise. Finally we let

$$R_{ij} = \sum_{k,l} R_{ijkl}\, dx^k \wedge dx^l.$$

Since 2-forms commute and $R_{ij} = -R_{ji}$, $\mathrm{Pf}(R_{ij})$ is a well-defined n-form, and we define the Pfaffian of the Riemann curvature tensor by

$$\text{(91.i)} \qquad \mathrm{Pf}(R_{ij}) = \mathrm{Pff}(R_{ijkl})\, dx^1 \wedge \cdots \wedge dx^n,$$

or, equivalently

$$\text{(91.ii)} \quad \mathrm{Pff}(R_{ijkl})\, dx^1 \wedge \cdots \wedge dx^n = c \sum \mathrm{sgn}(i)\,\mathrm{sgn}(j) R_{i_1 i_2 j_1 j_2}\, dx^{j_1} \wedge dx^{j_2}$$
$$\wedge \cdots \wedge R_{i_{2k-1} i_{2k} j_{2k-1} j_{2k}}\, dx^{j_{2k-1}} \wedge dx^{j_{2k}},$$

where $\mathrm{sgn}(i) = \mathrm{sgn}(i_1 \cdots i_{2k})$.

8. Hermann Weyl's invariant theory. First we prove an easy theorem about the action of $\mathrm{SO}(n)$ on n-vectors $v^1, \ldots, v^N$. The problem is to find those polynomials in the components of the n-vectors that are invariant under the action of $\mathrm{SO}(n)$.

There are two obvious examples:

(92)
(i) the inner product $\langle v^j, v^k \rangle = v_1^j v_1^k + \cdots + v_n^j v_n^k$;

(ii) the bracket factor $\left[v^{j_1},\ldots,v^{j_n}\right] = \det \begin{pmatrix} v_1^{j_1} & \cdots & v_1^{j_n} \\ \vdots & & \vdots \\ v_n^{j_1} & \cdots & v_n^{j_n} \end{pmatrix}$.

In fact we have

THEOREM 4. *Examples* (i) *and* (ii) *generate the invariant polynomials.*

Next we consider the more complicated problem of determining the SO(n)-invariant polynomials in the components of a collection of tensors of fixed rank, $\{A_{ij\cdots p}\}$.

THEOREM 5. (i) *We can write down an infinite list of polynomials generating the invariant polynomials as a vector space.*

(ii) *Finitely many terms from this list generate all the invariant polynomials as a ring.*

The third problem we will consider also concerns the set of tensors of fixed rank, $\{A_{ij\cdots p}\}$. The problem is to find all SO(n)-invariant linear maps

$$l\colon \{A_{ij\cdots p}\} \to \mathbf{C}.$$

There are two obvious examples:

(93)
(i) the map formed by pairing indices and contracting to scalars;

(ii) the map formed by first tensoring $A_{ij\cdots p}$ with an alternating n-tensor Ω to get $A \otimes \Omega$ and then pairing indices and contracting to scalars,

and a theorem.

THEOREM 6. *Linear combinations of examples* (i) *and* (ii) *are the only possible* SO(n)*-invariant linear maps.*

PROOF OF THEOREM 4. We will derive this theorem as a consequence of Theorem 6. Let $P(v^1,\ldots,v^N)$ be an SO(n)-invariant polynomial, i.e.,

(94)
(i) $P(gv^1,\ldots,gv^N) = P(v^1,\ldots,v^N)$ for all $g \in \mathrm{SO}(n)$,

(ii) P is a polynomial in the components $v_1^i,\ldots,v_n^i$ of the n-vectors v^i.

We may assume that P is homogeneous of degree m_k in the vector v^k, that is

$$P(v^1,\ldots,\lambda v^k,\ldots,v^N) = \lambda^{m_k} P(v^1,\ldots,v^k,\ldots,v^N). \tag{95}$$

Consider a map ϕ

$$\phi\colon (v^1,\ldots,v^N) \to v^1 \otimes \cdots \otimes v^1 \otimes \cdots \otimes v^N \otimes \cdots \otimes v^N,$$

where the vector v^i is repeated m_i times. Let $\mathcal{E}$ equal the set of all $(m_1 + \cdots + m_N)$-tensors obtained this way, and let $\mathcal{E}^+$ be the vector subspace of $V = \{(m_1 + \cdots + m_N)\text{-tensors}\}$ generated by $\mathcal{E}$. We can always find a linear map $l\colon \mathcal{E}^+ \to \mathbf{C}$ such that $P = l \circ \phi$.

Note that l is an $\mathrm{SO}(n)$-invariant map and $\mathcal{E}^+$ is an $\mathrm{SO}(n)$-invariant subspace of V. Recall that whenever a compact group G acts on a vector space V, we can define an invariant inner product by integrating any inner product over G. We have such an inner product because $\mathrm{SO}(n)$ is compact; therefore, we can define the complementary subspace $\mathcal{E}^-$ to $\mathcal{E}^+$ with respect to this invariant inner product and split V, $V = \mathcal{E}^+ \oplus \mathcal{E}^-$. If we define l to be 0 on $\mathcal{E}^-$, then l becomes an $\mathrm{SO}(n)$-invariant map on V. By Theorem 6, $P = l \circ \phi$ is a linear combination of

(i) terms that arise from

$$\underbrace{v^1 \otimes v^1} \otimes \cdots \otimes \underbrace{v^2 \otimes \cdots \otimes v^N}$$

by pairing as indicated and then contracting indices; this gives terms such as

$$\langle v^1, v^1\rangle \cdot \langle v^2, v^N\rangle \cdot \cdots ;$$

(ii) terms that arise by tensoring with an antisymmetric n-tensor Ω

$$v^1 \otimes v^1 \otimes \cdots \otimes v^N \otimes \Omega$$

and then pairing and contracting indices; these are products of terms like those from (i) with

$$\operatorname{Trace}\left[v^1 \otimes \cdots \otimes v^N \otimes \Omega\right],$$

which is the bracket factor. □

PROOF OF THEOREM 5. Part (i) follows from the same argument used to prove Theorem 4. Let $P(A, B,\ldots,C)$ be an $\mathrm{SO}(n)$-invariant polynomial acting on tensors of a fixed rank, i.e. an invariant polynomial in the components of these tensors. We may assume that P is homogeneous of degree $m_1, m_2,\ldots,m_N$ in the tensors $A, B,\ldots, C$. The map ϕ is defined by

$$\phi\colon (A, B,\ldots,C) \to A \otimes \cdots \otimes A \otimes \cdots \otimes C \otimes \cdots \otimes C,$$

where the tensors are repeated the number of times given by their degrees. Again $\mathcal{E} = \operatorname{Ran}\phi$ generates an $\mathrm{SO}(n)$-invariant subspace $\mathcal{E}^+$ of a finite-dimensional tensor space V and V splits with respect to an $\mathrm{SO}(n)$-invariant inner product, $V = \mathcal{E}^+ \oplus \mathcal{E}^-$. Finally, we can define an $\mathrm{SO}(n)$-invariant linear map l such that

$$P \circ \phi = l|_{\mathcal{E}}.$$

By Theorem 6 we know that all invariant polynomials in $A, B,\ldots,C$ are linear combinations of

$$\operatorname{Trace}[A \otimes \cdots \otimes A \otimes \cdots \otimes C], \qquad \operatorname{Trace}[A \otimes \cdots \otimes A \otimes \cdots \otimes C \otimes \Omega].$$

This finishes part (i); part (ii) is an immediate corollary of the next theorem. □

THEOREM 7 (HILBERT). *If G is a compact group acting linearly on a vector space V, then the ring of G-invariant polynomials P on V is finitely generated.*

PROOF. Let $\mathcal{I}$ be the ideal of polynomials generated by the G-invariant polynomials homogeneous of degree > 0. Then $\mathcal{I}$ is finitely generated; let $P_1, \ldots, P_k$ be its generators. We can assume that they are G-invariant and homogeneous of degree d_l. Let $D = \max\{d_1, \ldots, d_k\}$ and expand this list of generators to a larger finite list which generates the vector space of all G-invariant, homogeneous polynomials of degree d with $0 \leqslant d \leqslant D$. We claim that this finite list generates the ring of invariant polynomials. To prove the claim, let Q be an invariant homogeneous polynomial of degree greater than D. Then

$$Q = S_1 P_1 + \cdots + S_k P_k,$$

where $P_1, \ldots, P_k$ are G-invariant. Under the action of G the equation becomes

$$Q = S_1^g P_1 + \cdots + S_k^g P_k.$$

Since G is compact, we can average this equation over the group to obtain

$$Q = \hat{S}_1 P_1 + \cdots + \hat{S}_k P_k,$$

where the $\hat{S}_l$ are invariant, since they arise by averaging. Now we can assume that

$$\deg \hat{S}_l = \deg Q - \deg P_k,$$

because we can just delete the other terms. This allows us to complete the proof by induction on $\deg Q$; we just note that $\hat{S}_l$ are G-invariant homogeneous polynomials with lower degree than Q. $\square$

PROOF OF THEOREM 6. Recall that the theorem concerns the classification of $\mathrm{SO}(n)$-invariant linear maps l: $\{A_{ij\cdots p}\} \to \mathbf{C}$. Since the kernel of such a map is an $\mathrm{SO}(n)$-invariant subspace, one could consider the more general question of what are all the $\mathrm{SO}(n)$-invariant subspaces of the space of tensors of fixed length N.

This problem amounts to writing down explicitly a decomposition of the space V of tensors into irreducible representations. A typical irreducible subspace of V consists of all tensors of the form

$$\delta_{i_1 i_2} \delta_{i_3 i_4} \cdots \delta_{i_{2s-1} i_{2s}} A_{j_1 \cdots j_m}, \quad \text{where } \left(A_{j_1 \cdots j_m}\right) \text{ has all its traces} = 0$$

and is (roughly) symmetric in certain indices and antisymmetric in others. The pieces are described precisely in terms of combinatorial objects called Young's tableaux. Irreducibility can be proved using Weyl's character formula. A beautiful explanation of these ideas is given in Weyl [**67**].

But the proof of Theorem 6 is much less complicated. We shall use induction on the dimension, and make the transition from $(n-1)$ to n dimensions by exploiting the simple case in which the tensors $(A_{j_1 \cdots j_N})$ are all symmetric. The first step is to prove

LEMMA 3. *Assume $n > 1$. If l is an $\mathrm{SO}(n)$-invariant linear functional on the space of symmetric N-tensors $(A_{j_1 \cdots j_N})$, then l is a constant multiple of*

$$\left(A_{j_1 \cdots j_N}\right) \mapsto \operatorname{Trace}\left(A_{j_1 \cdots j_N}\right) = \sum_{i_1 \cdots i_{N/2}} A_{i_1 i_1 i_2 i_2 \cdots i_{N/2} i_{N/2}}.$$

In particular $l = 0$ if N is odd.

PROOF. We may identify $(A_{j_1\cdots j_N})$ with the polynomial

$$P(x) = \sum_{j_1\cdots j_N} A_{j_1\cdots j_N} x_{j_1}\cdots x_{j_N}.$$

Thus l is an $\mathrm{SO}(n)$-invariant functional on the space of polynomials on $\mathbf{R}^n$ homogeneous of degree N. Since l is rotation-invariant, we may average over $\mathrm{SO}(n)$ to write $l(P) = l(\int_{\mathrm{SO}(n)} P^g\, dg)$. Now $\int_{\mathrm{SO}(n)} P^g\, dg$ is rotation-invariant and homogeneous of degree N; hence it is a constant multiple of $|x|^N$ and so must vanish for N odd. For N even we see that l is determined uniquely by its action on the single polynomial $|x|^N$, so the space of all possible l is one-dimensional. □

We will now prove Theorem 6 using induction and the result above concerning symmetric tensors. If $n = 1$, then $(A_{i_1\cdots i_N}) = A_{1\cdots 1}$, so the tensor space is one dimensional and l is a constant multiple of

$$A \to A_{1\cdots 1} = \begin{cases} \operatorname{Trace} A, & N \text{ even},\\ \operatorname{Trace}(A \otimes \Omega), & N \text{ odd}. \end{cases}$$

So assume Theorem 6 in $(n-1)$ dimensions, and let l be an $\mathrm{SO}(n)$-invariant functional on the space T_N^n of N-tensors on $\mathbf{R}^n$. Consider the action of $\mathrm{SO}(n-1) \subseteq \mathrm{SO}(n)$, which fixes the nth unit vector e_n. Now, given a tensor $(A_{j_1\cdots j_N}) \in T_N^n$ and a subset $E \subseteq \{1,\ldots,N\}$, the components $(A_{j_1\cdots j_N})$, $j_s = n$ for $s \in E$, $j_s \neq n$ for $s \notin E$ form an $(N - |E|)$ tensor on $\mathbf{R}^{n-1}$. In this way we can write $T_N^n \cong \sum_{E\subseteq\{1\cdots N\}} \oplus T_{N-|E|}^{n-1}$, and the isomorphism commutes with the action of $\mathrm{SO}(n-1)$.

Note that the pieces of a given tensor $(A_{j_1\cdots j_N})$ all have the form $\operatorname{Trace}(A_{j_1\cdots j_N} \otimes e_n \otimes e_n \otimes \cdots \otimes e_n)$, where the trace contracts to an $(N - |E|)$-tensor. Since l is $\mathrm{SO}(n-1)$-invariant on a direct sum of $T_{N-|E|}^{n-1}$, Theorem 6 in $(n-1)$ dimensions express l as a linear combination of terms of the form

(i) $$\operatorname{Trace}' \operatorname{Trace}[A \otimes e_n \otimes \cdots \otimes e_n]$$

and

(ii) $$\operatorname{Trace}' \operatorname{Trace}[A \otimes \Omega' \otimes e_n \otimes \cdots \otimes e_n].$$

Here Trace acts on tensors on $\mathbf{R}^n$, and contracts to an $(N - |E|)$-tensor, while Trace′ acts on tensors on $\mathbf{R}^{n-1}$ and contracts to a scalar. Also, Ω' is an alternating tensor on $\mathbf{R}^{n-1}$.

Now, $\Omega' = \operatorname{Trace}(\Omega \otimes e_n)$, and for a 2-tensor (A_{ij}) we have

$$\operatorname{Trace}' A = \sum_{i=1}^{n-1} A_{ii} = \operatorname{Trace} A - A_{nn} = \operatorname{Trace} A - \operatorname{Trace}(A \otimes e_n \otimes e_n).$$

Analogous formulas hold for higher-rank tensors, and it follows that (i), (ii) are linear combinations of terms

(i′) $$\operatorname{Trace}[A \otimes e_n \otimes \cdots \otimes e_n],$$

(ii′) $$\operatorname{Trace}[A \otimes \Omega \otimes e_n \otimes \cdots \otimes e_n].$$

This time the traces contract to scalars.

Set $B = \mathrm{Trace}[A \otimes \Omega]$, so that B is a tensor and the term (ii′) may be written as $\mathrm{Trace}[B \otimes e_n \otimes \cdots \otimes e_n]$.

So far, we have used only the $\mathrm{SO}(n-1)$-invariance of l; now it is time to use $\mathrm{SO}(n)$-invariance. In view of the full $\mathrm{SO}(n)$-invariance, e_n in (i′), (ii′) may be replaced by any other unit vector v. Thus, l is a linear combination of terms

$$\mathrm{Trace}[A \otimes v \otimes \cdots \otimes v], \qquad \mathrm{Trace}[B \otimes v \otimes \cdots \otimes v],$$

where the coefficients of these terms are independent of $v \in S^{n-1}$. Averaging over all $v \in S^{n-1}$, we have l expressed as a linear combination of terms

$$l'(A) = \int_{S^{n-1}} \mathrm{Trace}[A \otimes v \otimes \cdots \otimes v]\, dv,$$

and

$$l'(B) = \int_{S^{n-1}} \mathrm{Trace}[B \otimes v \otimes \cdots \otimes v]\, dv.$$

A glance at the definition shows that l' is an $\mathrm{SO}(n)$-invariant functional, and that $l'(A) = l'(\hat{A})$, where

$$\hat{A}_{j_1 \cdots j_N} = \frac{1}{N!} \sum_{\pi \in S_N} A_{j_{\pi(1)} \cdots j_{\pi(N)}}$$

is the symmetrization of A. So the lemma on symmetric tensors shows that l' is a constant multiple of $\mathrm{Trace}\,\hat{A}$, which is a combination of traces of A. Applying the same reasoning to $l'(B)$, we have shown that l is a linear combination of terms of the form $\mathrm{Trace}\,A$, $\mathrm{Trace}\,B$. Finally, since $B = \mathrm{Trace}[A \otimes \Omega]$, we see that l is a linear combination of terms $\mathrm{Trace}\,A$, $\mathrm{Trace}[A \otimes \Omega]$. The inductive step is complete, and Theorem 6 is proved. □

We need some notation in order to state the next theorem. Let

$$T^k = \{\text{tensors of rank } k \text{ on } \mathbf{R}^n\}, \qquad \mathbf{T} = T^{k_1} \otimes \cdots \otimes T^{k_p},$$

where $k_1, \ldots, k_p$ are a collection of indices. The theorem concerns $\mathrm{SO}(n)$-invariant sets $\mathfrak{S} \subseteq \mathbf{T}$. For example, a point of $\mathfrak{S}$ could consist of a curvature tensor and all its covariant derivatives up to order 6, so that $\mathfrak{S} \subset T^4 \oplus T^5 \oplus \cdots \oplus T^{10}$.

THEOREM 8 (WEYL). *Let P be a polynomial on $\mathbf{T}$ whose restriction to $\mathfrak{S}$ is $\mathrm{SO}(n)$-invariant. Then $P = P_1 + P_2$, where $P_1 \equiv 0$ on $\mathfrak{S}$, P_2 is $\mathrm{SO}(n)$-invariant on $\mathbf{T}$, and P_2 is generated from elements on a certain finite list.*

PROOF. The last statement follows from Theorem 5. We assume that P is homogeneous in $A_{i_1 \cdots i_{k_l}}$ of degree m_{k_l}. Let

$$\mathfrak{T} = T^{k_1} \otimes \cdots \otimes T^{k_1} \otimes \cdots \otimes T^{k_p} \otimes \cdots \otimes T^{k_p},$$

where the T^{k_l} factor is included m_{k_l} times and define ϕ and l as in Theorem 4

$$\mathfrak{S} \subseteq \mathbf{T} \xrightarrow{\phi} \mathfrak{T} \xrightarrow{l} \mathbf{C}.$$

Let $\mathfrak{S}^+ = \mathrm{span}(\phi(\mathfrak{S}))$, so that

$$P|_{\mathfrak{S}} = l \circ \phi,$$

where l is an SO(n)-invariant linear map on $\mathfrak{S}^+$. If $\mathfrak{S}^-$ is the SO(n)-invariant complementary subspace to $\mathfrak{S}^+$, then $\mathfrak{T} = \mathfrak{S}^+ \oplus \mathfrak{S}^-$ and we can define $l^\#$ by

$$l^\# |_{\mathfrak{S}^+} = l, \qquad l^\# |_{\mathfrak{S}^-} = 0,$$

so that $l^\#$ is an SO(n)-invariant linear map on $\mathfrak{T}$. Putting

$$P_2 = l^\# \circ \phi, \qquad P_1 = P - P_2,$$

finishes the proof, since P_2 is SO(n)-invariant and $P_2 |_{\mathfrak{S}} = P |_{\mathfrak{S}}$. $\square$

We turn now to Weyl's invariant theory for SO(n, m), the special orthogonal group preserving the quadratic form $x_1^2 + \cdots + x_n^2 - x_{n+1}^2 - \cdots - x_{n+m}^2$. Since SO($n, m$) is not compact, all the proofs above involving the existence of SO(n, m)-invariant complementary subspaces have to be modified. This is achieved through Weyl's "*unitarian trick*", which reduces questions concerning SO(n, m) to those concerning SO($n + m$).

PROPOSITION 7. *Let $\mathfrak{S} \subset T^s$ be an* SO(n, m)*-invariant subspace. Then there exists an* SO(n, m)*-invariant complementary subspace.*

PROOF. Introduce the norms and map indicated:

$$\begin{array}{ccc} \mathrm{SO}(n+m) & & \mathrm{SO}(n,m) \\ \|v\|^2 = v_1^2 + \cdots + v_{n+m}^2 & & \|v\|^2 = v_1^2 + \cdots + v_n^2 - v_{n+1}^2 - \cdots - v_{n+m}^2 \\ \alpha_0\colon (v_1 \cdots v_{n+m}) & \longrightarrow & (v_1, \ldots, v_n, iv_{n+1}, \ldots, iv_{n+m}), \end{array}$$

and define the map $\alpha\colon T^s \to T^s$ by

$$A_{j_1 \cdots j_s} \to \sigma(j_1) \cdots \sigma(j_s) A_{j_1 \cdots j_s},$$

where

$$\sigma(j) = \begin{cases} 1, & 1 \leqslant j \leqslant n, \\ i, & n+1 \leqslant j \leqslant n+m. \end{cases}$$

Denote the Lie algebras of SO(n, m) and SO($n + m$) by so(n, m) and so($n + m$), and let α_0' be the map induced on the Lie algebras by α_0. Using the diagram

$$\begin{array}{ccc} \mathrm{so}(n+m) & \text{acts on} & T^s \\ \downarrow \alpha_0' & & \downarrow \alpha \\ \mathrm{so}(n,m) & \text{acts on} & T^s \end{array}$$

and noting that the complexifications of so(n, m) and so($n + m$) are isomorphic as algebras over $\mathbf{R}$, we can now prove results for SO(n, m) using the analogous results for SO($n + m$). In particular a subspace of T^s is SO(n, m)-invariant if and only if it is annihilated by the complexification of so(n, m); so there is an SO(n, m)-invariant-complement to $\mathfrak{S}$ as required, finishing the proof. $\square$

This finishes our tour of Weyl's invariant theory. We now see what we can say about the heat kernel (86),

$$K_t(x, x) = \sum_{i \geqslant 0} (4\pi t)^{-n/2} \phi_i(x, x) t^i.$$

By Weyl's theorem (Theorem 8), each ϕ_l is a sum of terms of the form

$$\operatorname{Trace}\left[R_{ijkl/\sigma}\otimes\cdots\otimes R_{abcd/\tau}\right].$$

Because of the $O(n)$-invariance of the ϕ_i, terms of the form

$$\operatorname{Trace}\left[R_{ijkl/\sigma}\otimes\cdots\otimes R_{abcd/\tau}\otimes\Omega\right]$$

never appear. By the homogeneity formula (90) of the last section, each term must satisfy

$$(2+|\sigma|)+\cdots+(2+|\tau|)=2l.$$

Recall that by Theorem 2, the heat kernel for the operator $\partial/\partial t-\square_p$ takes the form

$$K_p^t(x,y)=(4\pi t)^{-n/2}e^{-\operatorname{dist}^2(x,y)/4t}\left\{\sum_{i\geqslant 0}\phi_i^p(x,y)t^i\right\}.$$

The expressions

$$S_\alpha=\partial_y^\alpha\phi_i^p(x,y)\,|_{y=x}$$

are tensor-valued polynomials in the $R_{ijkl/\sigma}$. It is not hard to check that all of our invariant theory goes through for tensor-valued polynomials (except that the statement of Theorem 7 has to be changed a little); therefore S_α is a linear combination of terms of the form

$$\operatorname{Trace}\left[R_{ijkl/\sigma}\otimes\cdots\otimes R_{abcd/\tau}\right],$$

where now the trace is taken by contracting down until we are left with s indices, where s is the number of indices in the multi-index $\alpha=(\alpha_1,\ldots,\alpha_s)$.

If we work hard enough, we can obtain this result without using invariant theory. This is because at each stage of the computation of the $\phi_l(x,y)$, we remain within the class of functions of this form. More precisely, each function $\psi(x,y)$ appearing in the computation of $\phi_l(x,y)$ has the property that in normal coordinates $(\partial_y^\alpha\psi(x,y)\,|_{y=x})_{|\alpha|=p}$ is a linear combination of p-tensors of the form $\operatorname{Trace}[R_{ijkl/\sigma}\otimes\cdots\otimes R_{abcd/\tau}]$. Later, when we consider analogous proofs of the asymptotic expansion of the Bergman kernel, the computations will take us out of the appropriate class of functions, and invariant theory will be needed to complete the proof.

9. Gilkey's invariant theory. With Gilkey's invariant theory we can explicitly calculate the polynomial involved in computing the Euler characteristic and thus prove

THEOREM 9 (CHERN-GAUSS-BONNET).

$$\chi(M)=C_n\int_M \operatorname{Pff}(R_{ijkl})\,d\mathrm{vol},$$

where

$$\operatorname{Pff}(R_{ijkl})\, dx_1 \wedge \cdots \wedge dx_n = c \sum \operatorname{sgn}(i) \operatorname{sgn}(j) R_{i_1 i_2 j_1 j_2}\, dx_{j_1} \wedge dx_{j_2} \wedge \cdots \wedge R_{i_{n-1} i_n j_{n-1} j_n}\, dx_{j_{n-1}} \wedge dx_{j_n}$$

for n even, and $\operatorname{sgn}(i) = \operatorname{sgn}(i_1, \ldots, i_n)$.

PROOF. We claim

$$\chi(M) = \int_M P(\partial^\alpha g_{ij})\, d\,\mathrm{vol},$$

where

(96)
(i) P is a polynomial in $(\partial^\alpha g_{ij})$, as long as $g_{ij}(x) = \delta_{ij}$ and $\partial_{x_k} g_{ij}(x) = 0$ (see the proof of Theorem 2, part (ii)),
(ii) P is $O(n)$-invariant,
(iii) P satisfies the weight condition (90) with $l = n$,
(iv) for manifolds $M^n = M^{n-1} \times S^1$ with the product metric on M^n, $P \equiv 0$.

Indeed, we already have proved the first three properties. To prove (iv), recall how P was obtained. On $\Lambda^p(M)$ forms may, or may not, contain $d\theta$:

$$f\, dx_{i_1} \wedge \cdots \wedge dx_{i_p}, \qquad g\, dx_{i_1} \wedge \cdots \wedge dx_{i_{p-1}} \wedge d\theta,$$

and so $(\partial/\partial t + \Box) = 0$ splits into two problems involving

$$\Box(f\, dx_{i_1} \wedge \cdots \wedge dx_{i_p}), \qquad (\Box(g\, dx_{i_1} \wedge \cdots \wedge dx_{i_{p-1}})) \wedge d\theta.$$

The heat kernels of both are the same, but the degrees are shifted by 1, causing the alternating sum to add to 0; thus, $P \equiv 0$. The proof is completed once the following theorem is proved. $\Box$

THEOREM 10 (GILKEY). *Properties* (96) *imply*

$$P = C_n \operatorname{Pff}(R_{ijkl}).$$

PROOF. First it is easy to verify that Pff satisfies (96.i)–(96.iv); all that remains is to show that the space of such P satisfying (96.i)–(96.iv) is 1 dimensional (0 dimensional, if n is odd). We begin by proving through four observations that (96.ii), (96.iii) and (96.iv) imply $P = P(R_{ijkl})$, with no covariant derivatives involved.

Observation 1. $\partial^\alpha g_{ij}$ are independent variables for $i \leqslant j$, $|\alpha| \geqslant 2$.

Observation 2. Suppose $M = \tilde{M} \times S^1$ with coordinates $(x_1, \ldots, x_n)$, where $(x_2, \ldots, x_n)$ are coordinates for $\tilde{M}$, while (x_1) coordinates S^1. We claim that M has the product metric iff the equation

$$\partial^\alpha g_{ij} = 0 \quad \text{for } |\alpha| \geqslant 2$$

holds, whenever $i = 1$, or if any index in α contains a 1. To see this, note that for the product metric, $g_{1k} = g_{k1} \equiv 1$ or 0; and so $\partial^\alpha g_{1k} = \partial^\alpha g_{k1} = 0$, for any $|\alpha| \geqslant 1$, and, in particular, for $|\alpha| \geqslant 2$. By rotation invariance we conclude that $\partial_{x_1} g_{ij} = 0$. Conversely, assume that $\partial^\alpha g_{ij} = 0$, if $i, j = 1$ or if 1 occurs in α, for $|\alpha| \geqslant 2$. Then $g_{ij}(x_1, \ldots, x_n) = g_{ij}(x_2, \ldots, x_n)$ and so the metric satisfies

$$ds^2 = dx_1^2 + \sum_{i,j>1} g_{ij}\, dx_i\, dx_j,$$

as required.

Observation 3. By Observations 1 and 2, property (96.iv) forces each monomial in $P(\partial^\alpha g_{ij})$ to involve the index 1, either in α or in i, j.

Observation 4. Each monomial in $P(\partial^\alpha g_{ij})$ contains at least two occurrences of the index 1, although maybe not in the same factor $\partial^\alpha g_{ij}$; this is forced by $O(n)$-invariance and is checked by looking at reflections. Since there is nothing special about the index 1, each monomial in $\partial^\alpha g_{ij}$ is of the form

$$\partial^\alpha g_{ab} \cdots \partial^\gamma g_{kl}$$

where each of $1, 2, \ldots, n$ appears at least two times. If s is the number of factors in this monomial, then the weight property (90) implies

(i) $$(|\alpha| + |\beta| + \cdots + |\gamma|) = n$$

while also

(ii) $$(|\alpha| + |\beta| + \cdots + |\gamma|) + 2s \geqslant 2n,$$

since each index $1, \ldots, n$ appears at least twice. Since $|\alpha|, |\beta|, \ldots, |\gamma| \geqslant 2$, (i) yields $2s \leqslant n$, so (ii) becomes $(|\alpha| + |\beta| + \cdots + |\gamma|) \geqslant 2n - 2s \geqslant n$. If either inequality is strict, we get a contradiction with (i). Therefore $s = n/2$ and $|\alpha| = |\beta| = \cdots = |\gamma| = 2$. So our monomial has degree $n/2$ and involves only second derivatives of the metric, as claimed. In other words, $P = P(\partial^\alpha g_{ij})$ ($|\alpha| = 2$) with P homogeneous of degree $n/2$.

To prove Gilkey's theorem, we shall produce a linear functional on the space of P satisfying (96) which vanishes only at zero. This implies that the space (96) is one dimensional. Our functional will be of the simple form

$$P \to \text{coefficient of a monomial},$$

so it is sufficient to pick out a monomial whose coefficient cannot vanish if $P \neq 0$.

We choose the monomial $g_{11/22} g_{33/44} \cdots g_{n-1,n-1/nn}$ and calculate its coefficient. Here $g_{ij/kl} = \partial^2 g_{ij}/\partial x_k \partial x_l$. Let P satisfy (96). We conclude that

1. P contains a monomial of the form $g_{1j/kl} \cdot$ (other terms).

Since P is $O(n)$-invariant, it is invariant under permutations of the indices $\{1, \ldots, n\}$, so this is obvious.

2. P contains a monomial of the form $g_{11/kl} \cdot$ (other terms). If this were false then $j \neq 1$ in step 1. By permuting indices we may take $j = 2$, so P contains a monomial $g_{12/kl} \cdot$ (other terms). One checks easily that a generic rotation in the $x_1 - x_2$ plane produces in P a monomial $g_{11/pq} \cdot$ (other terms).

3. P contains a monomial of the form $g_{11/22} \cdot$ (other terms). If $k = l$ in step 2, this is obvious by permuting indices. Otherwise, we may assume $k = 2$, $l = 3$, so

that P contains a monomial $g_{11/23}\cdot$(other terms). Now performing a generic rotation in the $x_2 - x_3$ plane, we see that P contains a monomial $g_{11/22}\cdot$(other terms), as claimed.

4. Note that the other terms in step 3 cannot contain any more 1's or 2's. Now repeat steps 2 and 3 over and over, to obtain in P a monomial $g_{11/22} \cdots g_{2k-1,2k-1/2k,2k}\cdot$(other terms) for ever larger k. Finally, we obtain

5. P contains the monomial $g_{11/22}g_{33/44}\cdots g_{n-1,n-1/nn}$. The proof of Gilkey's theorem is complete. $\square$

Instead of using invariant theory, one can prove the Gauss-Bonnet theorem by very careful study of the construction of the heat kernels. This was the original method used by Patodi [52] to relate the heat equation to the Gauss-Bonnet theorem. Gilkey's theorem appears in [27]; this also contains a list of earlier references and a discussion of other index theorems.

CHAPTER 6. AN OVERVIEW OF TOPICS IN SEVERAL COMPLEX VARIABLES

1. Introduction. Now that the preliminaries are finished, we will briefly discuss the topics in several complex variables that will be covered. We begin with a table comparing the topics for Riemannian manifolds that have been discussed with the analogous questions in several complex variables.

Riemannian manifold M:	*Strongly pseudoconvex domain D* :
linear analysis: Laplacian, wave equation	$\bar\partial$, $\bar\partial_b$, $\square$, $\square_b$
geometry: normal coordinates, $\nabla_X Y$, parallel transport, geodesics, arclength, connection in principal bundle	Moser normal form, geometry of Poincaré metric, Monge-Ampère equation, Cheng-Yau theorem, formal analysis at ∂D, Cartan-Tanaka-Chern invariants, chains, parallel transport of frames along chains, parametrization
refined linear analysis: heat equation, invariance theory, Chern-Gauss-Bonnet theorem	Bergman-Szegö kernels, invariance theory for nonsemisimple groups, ?

We began our study of Riemannian geometry by first considering the simplest case, $\mathbf{R}^n$, with the Lie group of Euclidean motions. Similarly, we will begin the study of strongly pseudoconvex domains by considering the unit ball in $\mathbf{C}^n$ and the Siegel Domain and their Lie groups of linear fractional transformations, $\mathrm{SU}(n, 1)$ and the Heisenberg group.

The study of strongly pseudoconvex domains leads us to the topic of reflections, a subject which has no analogy in the Riemannian case. Using reflections, we can give a practical classification (in other words, computable in a reasonable fashion) for certain domains with analytic boundary.

For weakly pseudoconvex domains matters are much less settled. We have theorems establishing the interior regularity of the Poincaré metric, and, if we assume that the boundary is analytic, we have theorems establishing the C^∞-regularity and subellipticity of $\bar\partial$; but the analogies for the other entries in the table are still unknown.

Now we will fill in a few of the details of this big picture.

2. Linear analysis: $\bar\partial$, $\Box$, $\bar\partial_b$, $\Box_b$. We will consider functions, (0, 1)-forms, and (0, 2)-forms on a bounded, strongly ψ-convex domain D. Recall, for example, that a (0, 2)-form is locally of the form

$$f_{jk}\, d\bar z_j \wedge d\bar z_k.$$

We want to solve the equation

$$\bar\partial u = \alpha \tag{1.a}$$

on D, where α is a (0, 1)-form satisfying $\bar\partial\alpha = 0$. There are many solutions, since given any solution, we can produce others by adding holomorphic functions to the original solution. If we introduce the Hilbert space $L^2(D)$, we can define a good solution u by requiring that

$$u \perp \{\text{holomorphic functions}\}. \tag{1.b}$$

We will need a general fact about a first order system of differential operators $\mathfrak{L}$ on $\bar D$; for $u, v \in C^\infty(\bar D)$ and $\mathfrak{L}^*$ the formal adjoint of $\mathfrak{L}$, we have

$$\int_D (\mathfrak{L}u)\bar v\, d\mathrm{vol} = \int_D u\,\overline{(\mathfrak{L}^*v)}\, d\mathrm{vol} + \int_{\partial D} u\,\overline{(A^\# v)}\, d\mathrm{vol},$$

where $A^\#$ is a 0th order system of operators on the boundary. The domain of the adjoint $\mathfrak{L}^*$ consists of those $v \in C^\infty(\bar D)$ such that the boundary term equals 0.

In our case $\mathfrak{L} = \bar\partial$ and the domain of the adjoint is

$$\mathrm{Dom}(\bar\partial^*) = \{v \in C^\infty(\bar D)\colon A^\# v = 0 \text{ on } \partial D\}.$$

Note that with $u = \bar\partial^*\omega$, then for F holomorphic,

$$\langle \bar\partial^*\omega, F\rangle = \langle \omega, \bar\partial F\rangle = 0,$$

whenever the (0, 1)-form $\omega \in \mathrm{Dom}(\bar\partial^*)$; in other words, $u \perp$ {holomorphic functions}. This means that if we solve the equation

$$\bar\partial\bar\partial^*\omega = \alpha, \qquad \omega \in \mathrm{Dom}(\bar\partial^*),$$

then we can solve our original problem (1) by putting $u = \bar\partial^*\omega$.

In fact problem (1) is equivalent to the case $\bar\partial\alpha = 0$ of the system

$$\begin{gathered}\Box\omega \equiv (\bar\partial\bar\partial^* + \bar\partial^*\bar\partial)\omega = \alpha,\\ \omega \in \mathrm{Dom}(\bar\partial^*), \qquad \bar\partial\omega \in \mathrm{Dom}(\bar\partial^*).\end{gathered} \tag{2}$$

To see this, note that $0 = \bar\partial\alpha = \bar\partial(\bar\partial\bar\partial^* + \bar\partial^*\bar\partial)\omega = \bar\partial\bar\partial^*\bar\partial\omega$, and so

$$0 = \langle\bar\partial\omega, \bar\partial\bar\partial^*\bar\partial\omega\rangle = \langle\bar\partial^*\bar\partial\omega, \bar\partial^*\bar\partial\omega\rangle,$$

showing that $\bar\partial^*\bar\partial\omega = 0$. Comparing this with equation (2) gives $\bar\partial\bar\partial^*\omega = \alpha$, which by our earlier calculation is equivalent to (1). Problem (2) for general α is called the "$\bar\partial$-*Neumann problem.*"

We turn now to the boundary manifold $\partial D = M$ and its tangent bundle TM. Define $\mathfrak{T} \subseteq TM$ as the maximal subspace such that $i\mathfrak{T} = \mathfrak{T}$. If we fix a map

$$J\colon \{\text{real vectors in } \mathfrak{T}\} \to \{\text{real vectors in } \mathfrak{T}\}$$

such that $J^2 = -1$ (J corresponds to multiplications by $\sqrt{-1}$ in the ambient $\mathbf{C}^n$), then we can split $\mathbf{C} \otimes \mathfrak{T} = \mathfrak{T}^{(0,1)} \oplus \mathfrak{T}^{(1,0)}$ by defining

$$\mathfrak{T}^{(1,0)} = \{Z\colon Z = (X + iJX), \text{ some } X\},$$
$$\mathfrak{T}^{(0,1)} = \{\bar Z\colon \bar Z = (X - iJX), \text{ some } X\}.$$

As a brief review, recall that for boundary forms

type	*description*
(0, 0)	function
(0, 1)	linear form mapping $\bar Z \to \mathbf{C}$
(0, 2)	alternating 2-form mapping $(\bar Z, \bar W) \to \mathbf{C}$.

Now define $\bar\partial_b$ on (0, 0)- and (0, 1)-forms by

$$(\bar\partial_b f)(\bar Z) = \bar Z f, \qquad (\bar\partial_b\lambda)(\bar Z, \bar W) = c\{\bar Z\lambda(\bar W) - \bar W\lambda(\bar Z) - \lambda[\bar Z, \bar W]\}$$

so that $(\bar\partial_b)^2 = 0$. The analogy for problem (1) is

$$\bar\partial_b u = \alpha \quad \text{on } M$$

where we are given a (0, 1)-form α satisfying $\bar\partial_b\alpha = 0$. As above, this is equivalent to the problem

$$\square_b\,\omega \equiv (\bar\partial_b\bar\partial_b^* + \bar\partial_b^*\bar\partial_b)\omega = \alpha \quad \text{on } M.$$

This time there are no boundary conditions, for M has no boundary.

The codimension one subspace $\mathfrak{T} \subset T(\partial D)_p$ and the map $J\colon \mathfrak{T} \to \mathfrak{T}$ contain all the information for complex analysis and geometry on ∂D. More generally, we define a *C-R* (*Cauchy-Riemann*) *manifold* to be a $(2n - 1)$-manifold M together with a subbundle $\mathfrak{T} \subseteq TM$ with codimension one fibres, and a smoothly varying automorphism $J\colon \mathfrak{T} \to \mathfrak{T}$ with $J^2 = -\mathrm{Id}$. As in the familiar case $M = \partial D$, we can split $\mathbf{C} \otimes \mathfrak{T}$ into $\mathfrak{T}^{(1,0)} \oplus \mathfrak{T}^{(0,1)}$. For $M = \partial D$, the commutator of two vector fields in $\mathfrak{T}^{(1,0)}$ agains lies in $\mathfrak{T}^{(1,0)}$. A general *C-R* manifold with this property is called *integrable*. There is also a natural definition of strict pseudoconvexity for a *C-R* manifold. Thus, biholomorphic geometry may be thought of more generally as the study of strictly pseudoconvex integrable *C-R* manifolds.

In fact, a recent remarkable theorem of Kuranishi **[42]** asserts that every such manifold of real dimension $\geqslant 9$ arises locally as the boundary of a domain in $\mathbf{C}^n$.

This breaks down in dimension 3 (Nirenberg [50]); the intermediate dimensions are not yet settled. Kuranishi's proof is based on careful study of a highly degenerate boundary problem for the $\square_b$-operator—see also Boutet de Monvel [4] for an earlier result.

3. The unit ball and the Siegel domain. In this section we consider two simple models for strictly pseudoconvex domains: the unit ball and the Siegel domain. The *unit ball* is defined by

$$B = \left\{ z \in \mathbf{C}^n \colon u(z) = 1 - \sum_1^n |z_k|^2 > 0 \right\}$$

and its linear fractional transformations are given by $\mathrm{SU}(n,1)$. To realize the action of $\mathrm{SU}(n,1)$, let $z_k = \zeta_k/\zeta_0$, for $(\zeta_0,\ldots,\zeta_n) \in \mathbf{C}^{n+1}$, so that the inverse image of the ball is the subset of $\mathbf{C}^{n+1}$ given by

$$\left\{ \zeta \colon |\zeta_0|^2 - \sum_1^n |\zeta_k|^2 > 0 \right\}.$$

Since $\mathrm{SU}(n,1)$ is the group of linear transformations that preserve this quadratic form, an action of this group on the ball is induced by its action on $\mathbf{C}^{n+1}$.

The *Siegel domain* D is the unbounded version of the unit ball; it is defined by

$$D = \left\{ z \in \mathbf{C}^n \colon \operatorname{Re} z_1 > |z'|^2 \right\},$$

where $z' = (z_2,\ldots,z_n)$. Using the same coordinates as above, we can identify D with the subset of $\mathbf{C}^{n+1}$

$$\left\{ \zeta \colon \frac{1}{2}\zeta_1\bar{\zeta}_0 + \frac{1}{2}\zeta_0\bar{\zeta}_1 - \sum_2^n |\zeta_k|^2 > 0 \right\}$$

so that once again the group $\mathrm{SU}(n,1)$ of linear transformations preserving this quadratic form acts by linear fractional transformations on D. Note that there is a linear fractional transformation of $\mathbf{C}^n$ bringing the unit ball to D.

The boundary of D has the structure of a nilpotent Lie group, with the group multiplication law

$$(z_1, z') \cdot (w_1, w') = (z_1 + w_1 + 2z' \cdot w', z' + w').$$

This is the *Heisenberg group* N. N acts on ∂D by moving the origin around; moreover, $N \subset \mathrm{SU}(n,1)$. We also have a subgroup $H^+ \subset \mathrm{SU}(n,1)$, leaving the origin in ∂D fixed:

$$H^+ = \{ T \in \mathrm{SU}(n,1) \colon Te = \lambda e \},$$

where $e = (1,0,\ldots,0)$. Note that H^+ contains the subgroup of Heisenberg dilations

$$(z_1, z') \to (\delta^2 z_1, \delta z')$$

for $\delta > 0$. This later group is just a copy of $\mathbf{R}^\times$, the multiplicative group of positive reals. H^+ also contains the subgroup H defined by

$$H = \{ T \in \mathrm{SU}(n,1) \colon Te = e \}.$$

To clarify the structure of H, introduce the inversion operator i on the Siegel domain D; i is a linear fractional transformation of D which interchanges 0 and ∞. Explicitly, i is the linear fractional transformation induced by the matrix

$$\left(\begin{array}{cc|c} 0 & 1 & \\ 1 & 0 & 0 \\ \hline 0 & & \mathrm{I} \end{array}\right) \in U(n,1);$$

note that $i^2 = \text{Identity}$. Since the Heisenberg translations $(= N)$ preserve ∞ and move 0, $i(N)i$ fixes 0; also, its derivative at $0 = \mathrm{I}$. This means that $H \supset U(n-1)$ and $H \supset iNi$; in other words $H \supset N$. In terms of Lie algebras we have

$$\mathrm{su}(n,1) = n \oplus h^+, \qquad h^+ = \mathbf{R}^\times \oplus h, \qquad h = u(n-1) \oplus n.$$

A natural family of curves in the unit sphere is preserved under linear fractional transformations; this is the family of circles called chains. A *chain* is the intersection of the sphere with a complex line, not necessarily through the origin. An example of a chain in the Siegel domain is given by the line

$$\{z\colon \operatorname{Re} z_1 = 0, z' = 0\}.$$

The unit ball carries an SU$(n, 1)$-invariant metric

$$ds^2 = \sum_{j,k} \frac{\partial^2 \log(1 - |z|^2)}{\partial z_j \partial \bar{z}_k} dz_j \, d\bar{z}_k$$

and the associated volume element

$$d\,\mathrm{vol} = C_n \frac{s dz_1 \wedge \cdots \wedge dz_n \wedge d\bar{z}_1 \wedge \cdots \wedge d\bar{z}_n}{\left(1 - |z|^2\right)^{n+1}}.$$

With this metric the Ricci curvature is constant and negative. Finally, we claim that the Bergman kernel $K(z, w)$ and the Szegö kernel $S(z, w)$ for the unit ball B are

$$K(z,w) = \frac{n!}{\pi^n} \frac{1}{(1 - z \cdot \bar{w})^{n+1}}, \qquad S(z,w) = \frac{(n-1)!}{2\pi^n} \frac{1}{(1 - z \cdot \bar{w})^n}.$$

Indeed, recall the Bergman kernel arises from the orthogonal projection

$$\Pi\colon L^2(B) \to H(B), \qquad H(B) = \text{holomorphic functions}$$

via

$$\Pi f(z) = \int_B K(z,w) f(w) \, dw.$$

For $f \in H$ we want to show that for K given as above, $\Pi f = f$. Now

$$\begin{aligned} \Pi f(0) &= \int_B K(0,w) f(w) \, dw = \int_B \frac{n!}{\pi^n} \frac{1}{(1 - 0 \cdot \bar{w})^{n+1}} f(w) \, dw \\ &= \text{Average}_B(f) = f(0). \end{aligned}$$

Using this computation and the transitive action of $SU(n,1)$ on B, we can show that $(\Pi f)(z) = f(z)$, for $z \in B$, $f \in H$. This shows that $\Pi|_H =$ identity. Since $K(z,w)$ is holomorphic in z, it is easy to see that image $\Pi \subset H$. Finally because of the symmetric form of the kernel $K(z,w)$, a simple calculation shows that Π is selfadjoint. Therefore Π is a projection as required. The Szegö kernel is handled similarly.

From these simple models we can see that the function $u(z) = 1 - |z|^2$ appears many times. To get analogous formulas for general strictly pseudoconvex domains D, we will use the geometry of the Poincaré metric and the Monge-Ampère equation.

4. Geometry of strictly pseudoconvex domains: Moser normal form. Let D be a strictly pseudoconvex domain with real analytic boundary. This last assumption will be written as ∂D is C^ω. The boundary ∂D is said to be in *Moser normal form* near $0 \in \partial D$, if locally it can be represented by a convergent power series

$$\operatorname{Im}(z_n) = |z'|^2 + \sum_{\substack{|\alpha|,|\beta|>2\\ l>0}} A^l_{\alpha\bar\beta}\,(\operatorname{Re} z_n)^l z'^\alpha \bar z'^\beta,$$

where $z' = (z_1, \dots, z_{n-1})$, and where certain sums of the $A^l_{\alpha\bar\beta}$'s vanish. Note that when ∂D is in Moser normal form, the straight line

$$\{z\colon z' = 0 \text{ and } \operatorname{Im} z_n = 0\}$$

lies on ∂D.

THEOREM. *If D is a strictly pseudoconvex domain with C^ω boundary and $p \in D$, then there exists a biholomorphic map on a neighborhood of p such that*

$$p \to 0, \qquad \partial D \to \partial\tilde D,$$

where $\partial\tilde D$ is locally in Moser normal form.

PROOF (SKETCH). In this proof we will assign weights to terms using the table

$\operatorname{Re} z_n$	weight 2
$\operatorname{Im} z_n$	2
$z_1, \dots, z_{n-1}$	1
$\bar z_1, \dots, \bar z_{n-1}$	1

Since we can assume that ∂D is tangent to $\operatorname{Im} z_n = 0$ at $p = 0$, we can represent ∂D locally by

$$\begin{aligned}\operatorname{Im} z_n = &\sum_{1\le j,k\le n-1} \left(\lambda_{jk} z_j z_k + \bar\lambda_{jk} \bar z_j \bar z_k\right)\\ &+ \sum_{1\le j,k\le n-1} g_{j\bar k} z_j \bar z_k + (\text{terms of higher weight}),\end{aligned}$$

which under the change of coordinates

$$\begin{aligned}&z \to \hat{z}, \text{ where}\\ &\hat{z}_1 = z_1, \ldots, \hat{z}_{n-1} = z_{n-1}, \qquad \hat{z}_n = z_1 - 2 \sum_{1 \leqslant j,k \leqslant n} \lambda_{jk} z_j z_k,\end{aligned}$$

becomes

$$\operatorname{Im} \hat{z}_n = \sum_{1 \leqslant j,k \leqslant n} g_{j\bar{k}} z_j \bar{z}_k + (\text{terms of higher weight}).$$

Since the strict pseudoconvexity implies that the quadratic form is positive definite, we can rotate and dilate coordinates 1 through $n-1$ so that ∂D is given by

$$\operatorname{Im} z_n = |z'|^2 + (\text{terms of higher weight}).$$

To complete the proof we will use the family of maps

$$\mathcal{C} = \left\{\begin{array}{l}\text{biholomorphic maps } \phi\colon z \to w \text{ such that } \phi(0) = 0\\ \text{and locally } \phi \text{ preserves surfaces of the form}\\ \operatorname{Im} z_n = |z'|^2 + (\text{terms of weight} \geqslant 3)\end{array}\right\},$$

and the subfamily

$$\mathcal{C}_0 = \left\{\begin{array}{l}\phi \in \mathcal{C}\colon \phi'(0) = \mathrm{I}, \ \dfrac{\partial^2 w'}{\partial z_j \partial z_k}\bigg|_0 = 0, \text{ and}\\[2ex] \operatorname{Re}\left(\dfrac{\partial^2 w_n}{\partial z_n^2}\right) = 0, \text{ where } \phi\colon (z', z_n) \to (w', w_n) \text{ and}\\[2ex] 1 < j, k < n-1\end{array}\right\}.$$

We can simplify an element $\phi \in \mathcal{C}$ by composing it successively with linear fractional transformations from $\mathbf{R}^\times$, from $U(n-1)$, and from iNi.

Here we have in mind the lattice of subgroups

$$\begin{tikzcd}[column sep=small] & \mathrm{SU}(n-1,1) \arrow[dl] \arrow[dr] & & \\ N & & H^+ \arrow[dl] \arrow[dr] & \\ & H \arrow[dl] \arrow[dr] & & R^\times \\ U(n-1) & & iNi & \end{tikzcd}$$

The result is

$$\mathcal{C} = H^+ \mathcal{C}^0 = \mathcal{C}^0 H^+ . \tag{3}$$

In fact we have the

THEOREM (MOSER). *Locally, any* $\partial D \in C^\omega$ *given by*

$$\operatorname{Im} z_n = |z'|^2 + (\textit{terms of weight} \geqslant 3)$$

may be carried to

$$\operatorname{Im} z_n = |z'|^2 + \sum_{\substack{|\alpha|,|\beta|>2 \\ l \geq 0}} A^l_{\alpha\beta}(\operatorname{Re} z_n)^l z'^\alpha \bar{z}'^\beta$$

by exactly one member of $\mathcal{Q}^0$.

Using formal power series arguments, we can prove uniqueness; convergence is established using another argument.

Note that the decomposition (3) induces an action of H^+ on normal forms: take a normal form, apply an element of H^+, and then apply the unique element of $\mathcal{Q}^0$ which returns the map to normal form.

Let ϕ: $\partial D \to \partial \tilde{D}$ be a map that sends $D \in C^\omega$ to Moser normal form. We know that such maps exist for strictly pseudoconvex domains by the theorem above. Recall that the straight line

$$\{\tilde{z} \colon \tilde{z}' = 0 \text{ and } \operatorname{Im} \tilde{z}_n = 0\}$$

lies on $\partial\tilde{D}$. A *chain* is defined to be the inverse image of this line under ϕ. There is a whole family of chains through a point on ∂D; these chains turn out to be the solution curves to certain second order ODE's. In other words, chains are partly the analogues of geodesics. On the sphere all chains are circles. More precisely, they are the intersections of the sphere with complex lines.

Moser normal forms are one way of getting local invariants for strictly pseudoconvex domains. Later we will see how local invariants can reduce the question of whether a domain is locally equivalent to the unit ball to the question of whether the defining function for the boundary satisfies a finite number of nonlinear PDE's.

The theorem above produces Moser normal forms for embedded submanifolds. We will now sketch an intrinsic formulation for abstract Cauchy-Riemann manifolds; this is a second means of producing local invariants. Given an integrable *C-R* manifold *M*, we will define an $\mathbf{R}^\times$-bundle *E* over *M* and a frame bundle $Y \to E$. The bundle *Y* carries a natural Cartan connection ω satisfying

$$\begin{aligned} d\omega + [\omega, \omega] &= 0 \quad \text{in the flat model (Siegel domain)}, \\ &= \Pi \quad \text{a matrix of 2-forms, in general.} \end{aligned}$$

Chains are defined by projecting parallel transport down to *M*. For embedded surfaces the coefficients of Π are the $A^l_{\alpha\bar{\beta}}$ for $l = 0$ and $|\alpha|, |\beta| \leq 3$. Here, we will define these bundles only for the flat model; the general case is studied in Chapter 10.

As a first step, we define the bundles $\tilde{Y} \to \tilde{E}$ and $\tilde{E} \to M$. Define

$$M = \{\operatorname{Re} z_1 = |z'|^2\},$$

$$\tilde{E} = \left\{\xi \colon \|\xi\|^2 \equiv \frac{1}{2}\xi_0\bar{\xi}_1 + \frac{1}{2}\xi_1\bar{\xi}_0 - \sum_2^n \xi_k\bar{\xi}_k = 0\right\},$$

where $z' = (z_2, \dots, z_n)$, $\xi = (\xi_0, \dots, \xi_n)$ and $z_j = \xi_j/\xi_0$ for $j \geq 1$, so that $\tilde{E} \to M$ is a $\mathbf{C}^\times$-bundle over M. Next consider frames $F = \{(e_0, e_1, \dots, e_n)\}$ satisfying

$$e_i \in \mathbf{C}^{n+1},$$
$$\| e_0 \|^2 = \| e_1 \|^2 = 0, \qquad \langle e_\alpha, \bar{e}_\beta \rangle = \delta_{\alpha,\beta} \quad \text{for } \alpha, \beta > 2,$$
$$\langle e_0, \bar{e}_\alpha \rangle = \langle e_1, \bar{e}_\alpha \rangle = 0, \qquad \langle e_0, \bar{e}_1 \rangle = 1.$$

The frame F is called a Q-frame. Putting $\tilde{Y} = \{Q\text{-frames}\}$ and defining a projection $(e_0, e_1, \dots, e_n) \to e_0$ turns $\tilde{Y} \to \tilde{E}$ into a frame bundle. To define the connection note that $\tilde{Y}$ and $U(n, 1)$ can be identified. Then the tangent space of $\tilde{Y}$ is $u(n, 1)$ and the Lie algebra $h \subset u(n, 1)$ of H can be used to define the "horizontal subspace" of the connection; in other words the connection is defined by splitting the tangent space

$$u(n, 1) = n \oplus h,$$

where n is the Lie algebra of the Heisenberg translations N. Observe next that there is a natural action of the circle on $\tilde{E}$ and $\tilde{Y}$,

$$(e_0, e_1, \dots, e_n) \to (| z | e_0, | z | e_1, \dots, | z | e_n),$$

for z with $| z | = 1$. The bundles E and Y are defined by modding out with respect to this action. The splitting becomes

$$\mathrm{su}(n, 1) = n \oplus h^+,$$

and we see that, on the Lie algebra level, the Heisenberg translations N provide a natural Cartan connection satisfying the structural equations.

A third way of defining local invariants for the boundary of a strictly pseudoconvex domain D involves the Monge-Ampère equation; this is the subject of the next section.

5. The Poincaré metric and the complex Monge-Ampère equation. A *Poincaré metric* for a strictly pseudoconvex domain D is a Hermitian metric

$$ds^2 = \sum_{j,k} \frac{\partial^2 \log \psi}{\partial z_j \partial \bar{z}_k} dz_j \, d\bar{z}_k$$

with constant negative curvature

$$\mathrm{Ric}_{j\bar{k}} = -C g_{j\bar{k}}.$$

The existence of a Poincaré metric can be reduced to finding a solution u of the *complex Monge-Ampère equation*

$$\text{(4)} \qquad \mathrm{Det}(\partial\bar{\partial} \log u) = c u^{-(n+1)} \, dV$$
$$= \text{volume form for the Poincaré metric,}$$

where dV is the form

$$dV = dz_1 \wedge \cdots \wedge dz_n \wedge d\bar{z}_1 \wedge \cdots \wedge d\bar{z}_n.$$

For example the solution of the equation for the unit ball is $u(z) = 1 - | z |^2$. In general solving the equation consists of two parts.

The first part is the formal study of the equation

$$u \sim \phi + O([\delta(z)]^n),$$

where ϕ is smooth and calculated from the Taylor expansion of the boundary, and $\delta(z) = \operatorname{dist}(z, \partial D)$. We find

$$u \sim \phi + \sum_{\substack{k,l \\ k \geqslant (n+1)l}} \phi_{kl} \delta^k(z) \log \delta^l(z).$$

The second part is an existence theorem.

THEOREM (CHENG-YAU). *There exists a unique solution u of* (4) *with $u \in C^\infty(D)$ and $u \in C^{n+1/2-\varepsilon}(\bar{D})$.*

Returning to the example of the unit ball for a moment, we see that along with the Poincaré metric, we have the related metric on $\mathbf{C}^{n+1}$ given by

$$ds^2 = |d\zeta_0|^2 - \sum_{1 \leqslant k \leqslant n} |d\zeta_k|^2,$$

where $z = \zeta^1/\zeta_0$ and $\zeta' = (\zeta_1, \ldots, \zeta_n)$. If u is a solution of the Monge-Ampère equation (4), then the analogous metric on $D \times \mathbf{C}^\times$ is given by

$$ds^2 = \sum_{j,k>0} \frac{\partial^2}{\partial z_j \partial \bar{z}_k} \{|z_0|^2 u(z)\} \, dz_j \, d\bar{z}_k.$$

Although the metric ds^2 is nondegenerate, it is indefinite; one finds that the restriction $\widetilde{ds}^2$ of ds^2 to $\partial D \times \mathbf{C}^\times$ is degenerate. We can define a nondegenerate metric $\approx{ds}^2$

$$\widetilde{ds}^2 = \zeta^2(\overset{\approx}{ds}{}^2),$$

where $(z, \theta, \zeta) \in \partial D \times S^1 \times \mathbf{R}^\times$ and $z_0 = \zeta e^{i\theta}$. In fact $\overset{\approx}{ds}{}^2$ is a Lorentz metric on $\partial D \times S^1$ and the conformal class of this metric is a local invariant of ∂D.

Rays of light are invariant curves on $\partial D \times S^1$ and their projections onto ∂D define the *chains* of the boundary. It follows that we can also characterize chains by generalizing the geodesic equation of a manifold. Recall that on a Riemannian manifold, geodesics are defined by the Hamiltonian

$$H = \frac{1}{2} \sum g^{jk}(x) \xi_j \xi_k, \qquad (g^{jk}) > 0.$$

Similarly we can define chains by the Hamiltonian

$$H = \sum a_{jk}(x) \xi_j \xi_k + \sum b_l(x) \xi_l + V(x), \qquad (a_{jk}) \geqslant 0.$$

Note that (a_{jk}) will now merely be positive semidefinite; this means that chains need not have good behavior globally, or even locally. For example chains can be constructed that spiral inward; on the other hand, Burns-Schneider proved the global theorem that there are pairs of points that cannot be connected by chains.

6. Refined linear analysis. In §3 we saw that the Bergman kernel for the unit ball is

$$K(z,w) = \frac{c_n}{(1 - z \cdot \bar{w})^{n+1}}.$$

Note that

$$K(z,z) = \frac{c_n}{\left(1 - |z|^2\right)^{n+1}}.$$

In order to write an expansion of the Bergman kernel in the general case, we will need the

LEMMA. *If $\psi(z) \in C^\infty$, then there exists an extension $\psi(z,w)$ satisfying*
(i) $\bar{\partial}_z \psi|_{z=w} = 0$ *to infinite order,*
(ii) $\partial_w \psi|_{z=w} = 0$ *to infinite order,*
(iii) $\psi(z,z) = \psi(z)$.

Now let $D = \{\psi(z) > 0\} \subset \mathbf{C}^n$ be a strictly pseudoconvex domain.

THEOREM. *The Bergman kernel has the expansion*
(a) $K_D(z,z) = \phi(z)/\psi^{n+1}(z) + \tilde{\phi}(z)\log\psi(z)$,
where ϕ, $\tilde{\phi}$ are smooth and ϕ is nonvanishing;
(b) $K_D(z,w)$ *is obtained from* $K_D(z,z)$ *by extending* ϕ, $\tilde{\phi}$, ψ *as in the lemma above.*

Note the analogy between this expansion and the asymptotic expansion of the heat kernel on a Riemannian manifold,

$$K_t(z,z) \sim \frac{C_n}{t^{n/2}}\left\{1 + \sum_1^\infty \phi_k(z)t^k\right\}.$$

We can give a more precise description of the Bergman kernel using the Monge-Ampère equation. Again, a clue is provided by the unit ball, where the kernel takes the form

$$K(z,z) = \frac{C_n}{u^{n+1}(z)},$$

for $u(z)$ a solution of the Monge-Ampère equation (4). This suggests that we compare $K_D(z,z)$ and $C_n/u^{n+1}(z)$ for a general D. If D is locally in Moser normal form

$$\operatorname{Im} z_n > |z'|^2 + \sum_{p,q,r,s<n} A_{pq\bar{r}\bar{s}} z_p z_q \bar{z}_r \bar{z}_s + (\text{other terms}),$$

then by a calculation of Christoffers

$$K(z_t, z_t) = \frac{C_n}{t^{n+1}}\left\{1 + \gamma_n \sum_{p,q,r,s<n} |A_{pq\bar{r}\bar{s}}|^2 t^2 + O(t^3)\right\},$$

$$\frac{C_n}{u^{n+1}(z_t)} = \frac{C_n}{t^{n+1}}\left\{1 + \gamma_n' \sum_{p,q,r,s<n} |A_{pq\bar{r}\bar{s}}|^2 t^2 + O(t^3)\right\},$$

where $z_t = (0, 0, \ldots, it)$ and $\gamma'_n \neq \gamma_n$. We can write this invariantly by introducing the metric

$$ds^2 = \partial\bar{\partial}\big(|z_0|^2 u(z)\big) = \partial\bar{\partial}(U).$$

7. Reflections. Let D be a strictly pseudoconvex domain with real-analytic boundary, say $D = \{r(z) > 0\}$, $r \in C^\omega$. Let $r(z, w)$ be the power series obtained from that of $r(z)$ by replacing each $\bar{z}$ by $\bar{w}$. For instance, the defining function for an ellipsoid is

$$r(z) = \sum_k \left(\lambda_k z_k^2 + \lambda_k \bar{z}_k^2 + z_k \bar{z}_k\right),$$

so

$$r(z, w) = \sum_k \left(\lambda_k z_k^2 + \lambda_k \bar{w}_k^2 + z_k \bar{w}_k\right).$$

In general, $r(z)$ is defined up to a nonvanishing real-analytic factor; hence, so is $r(z, w)$. It follows that the set

$$\{(z, w) \mid r(z, w) = 0\}$$

is invariantly associated to ∂D, i.e. it is preserved under biholomorphic mappings of domains with real-analytic boundaries.

For domains in $\mathbf{C}^1$, this amount to the Schwarz reflection principle: we reflect $z \in D$ across the boundary to w which solves $r(z, w) = 0$. In more than one dimension, we obtain a tool much more powerful than Schwarz reflection. For, the family of codimension one varieties $V_w = \{z \in \mathbf{C}^n \mid r(z, w) = 0\}$, $w \in \mathbf{C}^n$, must be preserved by biholomorphic maps. Thus, for instance, an analytic automorphism of an ellipsoid necessarily carries a variety $\{z \mid \sum_k \lambda_k z_k^2 + \lambda_k \bar{\alpha}_k^2 + z_k \bar{\alpha}_k = 0\}$ to a variety $\{z \mid \sum_k \lambda_k z_k^2 + \lambda_k \bar{\beta}_k^2 + z_k \bar{\beta}_k = 0\}$, which strongly restricts how the automorphism can look.

As a simple application of the method of reflections, we shall prove in Chapter 8 that if $\partial D = \{r(z) = 0\}$, $\partial\tilde{D} = \{\tilde{r}(z) = 0\}$ and r, $\tilde{r}$ are polynomials on $\mathbf{C}^n$ $(n > 1)$, then any biholomorphic map Φ from D to $\tilde{D}$ is algebraic. To get an idea of the proof, we can show already that Φ is algebraic on each V_w.

In fact, for $z \in V_w$ we have also $w \in V_z$ (because $r(z)$ is real, hence $r(z, w) = \overline{r(w, z)}$). For typical $z \in V_w$, V_z will be nonsingular at w, so that we can associate to each $z \in V_w$ the hyperplane $T(z) =$ tangent space to V_z at w. All of this may be written down explicitly in terms of the polynomial $r(z, w)$, so that $z \to T(z)$ and its inverse are easily seen to be algebraic. (Here $T(z)$ is regarded as a point in the projective space of hyperplanes through w.) If Φ: $D \to \tilde{D}$ and $\Phi(w) = \tilde{w}$ then we have the natural maps

$$\begin{array}{ccc} V_w \ni z & \underset{\text{map}}{\overset{\text{Algebraic}}{\to}} & T(z) \\ \downarrow \Phi & & \downarrow D_\Phi \\ \tilde{V}_{\tilde{w}} \ni \tilde{z} & \underset{\text{map}}{\overset{\text{Algebraic}}{\to}} & \tilde{T}(z) \end{array}$$

where D_Φ carries a hyperplane H of tangent vectors at w to the image of H under the differential $\Phi'(w)$. One checks that D_Φ is a projective transformation, and, in particular, algebraic. The diagram commutes, so $\Phi|_{V_w}$ must be algebraic since the other maps are all algebraic. If we are in $\mathbf{C}^1$, then we have shown that Φ is algebraic when restricted to a single point, so we have gained nothing. But in $\mathbf{C}^n$ $(n > 1)$ we know that Φ is algebraic on a codimension one variety, and a single additional trick completes the proof that Φ is algebraic. A harder application of the method of reflection is Webster's classification of ellipsoids: in $\mathbf{C}^n$ $(n > 1)$ two ellipsoids are biholomorphically equivalent only if they are already equivalent by a (complex) linear transformation of $\mathbf{C}^n$. An ellipsoid other than the ball has no biholomorphic self-maps other than linear transformations. See Chapter 8.

All of this depends on the implicit assumption that a biholomorphic map of two C^ω strictly pseudoconvex domains continues analytically past the boundary. This was originally proved by Chern-Moser theory, but Hans Lewy gave a simple proof based on the method of reflection. The ideas were extended by Nirenberg-Webster-Yang to prove C^∞ regularity of biholomorphic maps of strictly pseudoconvex domains. More recently, a very simple proof of C^∞ regularity was discovered by Bell and Ligocka. Their ideas opened up the possibility of dropping the assumption of strict pseudoconvexity.

Finally, we refer to the book by Krantz [41] for background material in several complex variables as well as further discussion of some of the topics mentioned here.

CHAPTER 7. ANALYSIS ON THE SIEGEL DOMAIN AND ITS BOUNDARY, THE HEISENBERG GROUP

1. Solution to $\bar\partial_b$, $\Box_b$ on the Heisenberg group (Folland-Stein [25]). Recall that the Heisenberg group H^n is the boundary of the Siegel domain

$$\{\operatorname{Im} z_{n+1} > |z_1|^2 + \cdots + |z_n|^2\} \subset \mathbf{C}^{n+1},$$

with the group operation

$$(\zeta, t) \cdot (\zeta', t') = (\zeta + \zeta', t + t' + 2\operatorname{Im}\zeta \cdot \bar\zeta'),$$

where the coordinates (ζ, t) are defined by $\zeta = (z_1, \ldots, z_n) = x + iy \in \mathbf{C}^n$ and $t = \operatorname{Re} z_{n+1}$. The Lie algebra is spanned by the left-invariant vector fields

$$X_j = \frac{\partial}{\partial x_j} + 2y_j\frac{\partial}{\partial t}, \qquad Y_j = \frac{\partial}{\partial y_j} - 2x_j\frac{\partial}{\partial t}, \qquad T = \frac{\partial}{\partial t}, \qquad j = 1, \ldots, n,$$

which satisfy the commutation relations

$$[Y_j, X_k] = 4\delta_{jk}T, \qquad [X_j, X_k] = [Y_j, Y_k] = [X_j, T] = [Y_j, T] = 0.$$

The Haar measure on the group is given by Lebesgue measure $d\mathrm{vol} = dx_1\, dy_1 \cdots dx_n\, dy_n\, dt$.

Define the $(0, 1)$-forms $\bar\omega^j$ as the duals of the vector fields $\bar Z_j$, where

$$Z_j = \tfrac{1}{2}(X_j - iY_j), \qquad \bar Z_j = \tfrac{1}{2}(X_j + iY_j),$$

and choose a metric so that

$$\langle \bar{\omega}^j, \bar{\omega}^k \rangle = \delta_{jk}.$$

We can now define $\bar{\partial}_b$ acting on functions f and (0, 1)-forms $\Sigma\, \phi_j \bar{\omega}^j$:

$$\bar{\partial}_b f = \sum_{j=1}^{n} \left(\bar{Z}_j f\right) \bar{\omega}^j, \qquad \bar{\partial}_b\left(\sum_{j=1}^{n} \phi_j \bar{\omega}^j\right) = \sum_{j=1}^{n} \left(\bar{Z}_j \phi_k - \bar{Z}_k \phi_j\right) \bar{\omega}^j \wedge \bar{\omega}^k.$$

For a general Cauchy-Riemann manifold, there would be other terms involving $[\bar{Z}_j, \bar{Z}_k]$. A calculation shows that the operator $\Box_b = \bar{\partial}_b \bar{\partial}_b^* + \bar{\partial}_b^* \bar{\partial}_b$ acts on (0, 1)-forms by

$$\Box_b\left(\sum_{j=1}^{n} f_j \bar{\omega}^j\right) = \sum_{j=1}^{n} \left(\mathfrak{L}_\alpha f_j\right) \bar{\omega}^j,$$

where

$$\mathfrak{L}_\alpha = -\frac{1}{2} \sum_{j=1}^{n} \left(Z_j \bar{Z}_j + \bar{Z}_j Z_j\right) + i\alpha T \quad \text{and} \quad \alpha = n - 2.$$

In other words $\Box_b$ restricted to (0, 1)-forms is "diagonal" and is determined by an operator $\mathfrak{L}_\alpha$ that depends only on the dimension n.

In order to solve $\bar{\partial}_b$ and $\Box_b$, we will construct a fundamental solution to $\mathfrak{L}_\alpha$ for $\alpha \in \mathbf{C}$ except for a dicrete set of bad α; that is, an operator with convolution kernel K_α on H^n satisfying $\mathfrak{L}_\alpha K_\alpha = \delta_0$. In the coordinates (ζ, t), the operator $\mathfrak{L}_\alpha$ is invariant under rotations in ζ and dilations of the form $(r\zeta, r^2 t)$. These properties imply that the kernel K_α satisfies

$$K_\alpha = \frac{\Omega(u)}{\left(|\zeta|^4 + |t|^2\right)^{\text{power}}},$$

where $u = |\zeta|^2 + it$ and $\Omega(u)$ is homogeneous of degree 0 with respect to the standard dilations. In other words, with $u = re^{i\theta}$, Ω can be regarded as a function of θ alone. Because $\mathfrak{L}_\alpha K_\alpha = 0$ away from the origin, the function $\Omega(\theta)$ satisfies a differential equation; we readily find that

$$\Omega(e^{i\theta}) = e^{ik\theta},$$

where k depends upon α.

THEOREM 1 (FOLLAND-STEIN). *The kernel*

$$K_\alpha(\zeta, t) = \frac{c_\alpha}{\left(|\zeta|^2 - it\right)^{(n+\alpha)/2} \left(|\zeta|^2 + it\right)^{(n-\alpha)/2}}$$

satisfies $\mathfrak{L}_\alpha K_\alpha = \delta_0$ *and therefore can be used to invert*

$$\Box_b\left(\sum_{j=1}^{n} f_j \bar{\omega}^j\right) = \sum_{j=1}^{n} \left(\mathfrak{L}_\alpha f_j\right) \bar{\omega}^j.$$

PROOF. We will not compute the constant c_α. Observe that $\mathfrak{L}_\alpha K_\alpha$ is supported at the origin and is a distribution with the same homogeneity as that of the δ

function. This means that either $\mathcal{L}_\alpha K_\alpha$ is a multiple of the δ function or is zero. That is, $\mathcal{L}_\alpha K_\alpha = \gamma(\alpha)\delta_0$, and the only question is whether $\gamma(\alpha) = 0$. To evaluate $\gamma(\alpha)$, we test the distribution $\mathcal{L}_\alpha K_\alpha$ against a smooth function

$$\phi(\zeta, t) = \begin{Bmatrix} 1 & \text{for } |\zeta|^2 + |t|^2 \leqslant 1 \\ 0 & \text{for } |\zeta|^2 + |t|^2 \geqslant 2 \end{Bmatrix}.$$

Evidently, $\gamma(\alpha) = \langle \mathcal{L}_\alpha K_\alpha, \phi \rangle = \lim_{\varepsilon \to 0+} \langle \mathcal{L}_\alpha K_\alpha^\varepsilon, \phi \rangle$, where

$$K_\alpha^\varepsilon(\zeta, t) = \frac{c_\alpha}{\left(|\zeta|^2 - it + \varepsilon\right)^{(n+\alpha)/2}\left(|\zeta|^2 + it + \varepsilon\right)^{(n-\alpha)/2}}.$$

(Here we use $K_\alpha^\varepsilon \to K_\alpha$ as distributions.) Now $\int_{|\zeta|^2 + t^2 \geqslant 1} |\mathcal{L}_\alpha K_\alpha^\varepsilon(\zeta, t)|\, d\zeta\, dt \to 0$ as $\varepsilon \to 0+$, so

$$\gamma(\alpha) = \lim_{\varepsilon \to 0+} \langle \mathcal{L}_\alpha K_\alpha^\varepsilon, \phi \rangle = \lim_{\varepsilon \to 0+} \int_{H^n} \mathcal{L}_\alpha K_\alpha^\varepsilon(\zeta, t)\, d\zeta\, dt.$$

However, for different $\varepsilon > 0$, the kernels $\mathcal{L}_\alpha K_\alpha^\varepsilon$ are dilates of one another, and it follows that the integral on the right is independent of ε. Taking $\varepsilon = 1$, we obtain

$$\gamma(\alpha) = \int_{H^n} \mathcal{L}_\alpha K_\alpha^1(\zeta, t)\, d\zeta\, dt,$$

which can be evaluated in terms of gamma functions. The result is $\gamma(\alpha) \neq 0$ as long as $\pm\alpha \neq n, n+2, n+4, \ldots$. Except for these discrete values of α, we obtain $1/(\gamma(\alpha))K_\alpha$ as the fundamental solution for $\mathcal{L}_\alpha$. $\Box$

The idea of viewing the equations of complex analysis as the analogues of the Laplacian on the Heisenberg group is due to Stein. Theorem 1 is the simplest confirmation of this point of view.

2. $\Box$ on the Siegel domain with $\bar\partial$-Neumann boundary conditions (Phong [53]). In this section we will consider $\Box$ acting on (0, 1)-forms defined on the Siegel domain

$$D = \left\{(z, z_{n+1}) \in \mathbf{C}^{n+1} \colon \operatorname{Im} z_{n+1} > |z|^2\right\}.$$

The analysis of the operator $\Box$ is difficult because of the presence of two types of homogeneity: the operator $\Box$ is elliptic, while the nontrivial components of the boundary conditions have Heisenberg homogeneity. Because of this, the solution of $\Box$ involves composing two kernels: one is Heisenberg group invariant on the boundary and the other is a kernel of isotropic homogeneity. To understand this complication we begin by splitting $\Box$ into two pieces. To define this splitting we introduce the (1, 0)-forms

$$\omega^j = dz_j, \qquad j = 1, 2, \ldots, n,$$

$$\omega^{n+1} = -\sqrt{2} \sum_{j=1}^{n} \bar z_j\, dz_j - \frac{i}{\sqrt{2}}\, dz_{n+1}$$

and choose a metric so that these forms are orthonormal. We can now define the dual frame

$$Z_j = \frac{\partial}{\partial z_j} + 2i\bar{z}_j \frac{\partial}{\partial z_{n+1}}, \qquad j = 1, 2, \ldots, n,$$

$$Z_{n+1} = i\sqrt{2}\,\frac{\partial}{\partial z_{n+1}}$$

and the vector field

$$T = \frac{1}{\sqrt{2}}\left(\bar{Z}_{n+1} - Z_{n+1}\right).$$

Recall that the domain of $\square$ restricted to (0, 1)-forms $\eta = \sum_{j=1}^{n+1} \phi_j \bar{\omega}^j$ is

$$D(\square) = \left\{\eta\colon \phi_{n+1} = 0 \text{ on } \partial D;\ \bar{Z}_{n+1}\phi_j = 0 \text{ on } \partial D, j = 1, \ldots, n\right\},$$

and the action of $\square$ on (0, 1)-forms can be written

$$\square(\eta) = \sum_{j=1}^{n} \left(\square^+ \phi_j\right)\bar{\omega}^j + \left(\square^{\#}\phi_{n+1}\right)\bar{\omega}^{n+1},$$

where

$$\text{(1)}\qquad \begin{aligned} &\square^+ = \square_b - Z_{n+1}\bar{Z}_{n+1}, \qquad &&\square_b = -\frac{1}{2}\sum_{k=1}^{n}\left(Z_k\bar{Z}_k + \bar{Z}_k Z_k\right) - i(n-2)T,\\ &\square^{\#} = \square_b^* - Z_{n+1}\bar{Z}_{n+1}, \qquad &&\square_b^{\#} = \square_b - 2T. \end{aligned}$$

The calculations below are simpler if we use the coordinates (z, t, ζ), where

$$\zeta = \operatorname{Im} z_{n+1} - |z|^2, \qquad t = \operatorname{Re} z_{n+1},$$

for $(z, z_{n+1}) \in D$, and solve the problem on surfaces $\{\zeta = \text{constant}\}$. For these coordinates,

$$Z_j = \frac{\partial}{\partial z_j} + i\bar{z}_j \frac{\partial}{\partial t}, \qquad j = 1, 2, \ldots, n,$$

$$Z_{n+1} = i\frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial t} - i\frac{\partial}{\partial \zeta}\right), \qquad T = -i\frac{\partial}{\partial t}.$$

Since the operators $Z_1, \ldots, Z_n$, T do not contain $\partial/\partial\zeta$, they are tangential to the surfaces $\{\zeta = \text{constant}\}$.

Consider the equation

$$\text{(2.i)}\qquad \square\eta = \square\left(\sum_{j=1}^{n+1}\phi_j\bar{\omega}^j\right) = \left(\square^+ \oplus \square^{\#}\right)\left(\sum_{j=1}^{n+1}\phi_j\bar{\omega}^j\right) = \sum_{j=1}^{n+1}\psi_j\bar{\omega}^j.$$

This equation breaks into two parts

$$\text{(2.ii)}\qquad \begin{cases} \square^+ \phi_j = \psi_j & \text{in } D,\\ \bar{Z}_{n+1}\phi_j = 0 & \text{on } \partial D, j = 1, \ldots, n, \end{cases}$$

$$\text{(2.iii)} \qquad \begin{cases} \square^{\#}\phi_{n+1} = \psi_{n+1} & \text{in } D, \\ \phi_{n+1} = 0 & \text{on } \partial D. \end{cases}$$

The $\bar{\partial}$-*Neumann problem* is to find a solution of the system (2.i) or equivalently of the system (2.ii) and (2.iii). We can now state more precisely the goal of this chapter: it is to find an explicit approximate formula for the kernel associated with equation (2.i). We will see shortly that the kernel $G^{\#}$ associated with (2.iii) is easy to find. The bulk of the remainder of this chapter is devoted to finding an approximate formula for the kernel K associated with (2.ii)—for the precise statement see Theorem 2 at the beginning of §3.

Recall from §7 of Chapter 3 that when L is elliptic, the problem

$$\begin{aligned} Lu &= f \quad \text{in } D, \\ u &= g \quad \text{on } \partial D \end{aligned}$$

can be solved by transforming it into an equation for a pseudodifferential operator on ∂D; a Green's function then gives the solution u. Applying this to the boundary value problem (2.iii) gives

$$\phi_{n+1} = G^{\#}\psi_{n+1},$$

where $G^{\#}$ is the appropriate Green's function.

It is not as easy to find the kernel associated with (2.ii); i.e., the solution K of

$$\text{(3)} \qquad \begin{aligned} \square^{+} K &= \delta_p \quad \text{in } D, \\ \bar{Z}_{n+1} K &= 0 \quad \text{on } \partial D. \end{aligned}$$

We do this in several steps.

Step 1. We derive two 2nd-order equations (A, B) satisfied by K.

Step 2. We note by the symmetries of the problem that K is a function ϕ of four auxiliary variables.

Step 3. We derive a first order equation for ϕ.

Step 4. We analyze the 1st-order equation for ϕ.

Step 4 is the subject of §3.

Step 1. Since $\bar{Z}_{n+1}$ commutes with T, Z_j, $\bar{Z}_j, j = 1, \ldots, n$, equation (3) yields

$$\begin{aligned} &\bar{Z}_{n+1}(\square^{+} K) = \square^{+}\left(\bar{Z}_{n+1} K\right) = \bar{Z}_{n+1}\delta_p \quad &&\text{in } D, \\ &\left(\bar{Z}_{n+1} K\right) = 0 &&\text{on } \partial D. \end{aligned}$$

This is an elliptic boundary value problem for $\bar{Z}_{n+1}K$ with Dirichlet boundary conditions. Applying the elliptic theory of §7 of Chapter 3 again, we find

$$\bar{Z}_{n+1} K = G^{+}\left(\bar{Z}_{n+1}\delta_p\right),$$

for an appropriate Green's function G^{+}. Using the definition of $\square^{+}$ from (1) gives equation (A),

$$\text{(A)} \qquad \square_b K = \delta_p + Z_{n+1} G^{+} \bar{Z}_{n+1}\delta_p.$$

We now derive the second equation (equation B) satisfied by K. Let H be the Bergman projection operator. In Chapter 6 we saw that if $\alpha = \bar{\partial} f$, then the solution of $\bar{\partial} u = \alpha$ that is perpendicular to the space of holomorphic functions is given by either of the expressions

$$u = (\mathrm{I} - H)f = (\bar{\partial}^* \Box^{-1} \bar{\partial})f.$$

This gives us the identity

$$\mathrm{I} - H = \bar{\partial}^* \Box^{-1} \bar{\partial}.$$

We apply this identity, component by component, to δ_p. Now

$$\bar{\partial}\delta_p = \sum_{j=1}^{n} \left(\bar{Z}_j \delta_p\right)\bar{\omega}^j + \left(\bar{Z}_{n+1}\delta_p\right)\bar{\omega}^{n+1},$$

$$\Box^{-1}\bar{\partial}\delta_p = \sum_{j=1}^{n} \left(K\bar{Z}_j \delta_p\right)\bar{\omega}^j + \left(G^{\#}\bar{Z}_{n+1}\delta_p\right)\bar{\omega}^{n+1}$$

and

$$\bar{\partial}^*\left(\sum_{j=1}^{n+1} \phi_j \bar{\omega}^j\right) = -\sum_{j=1}^{n+1} Z_j \phi_j$$

so that

$$(\mathrm{I} - H)\delta_p = -\left(\sum_{j=1}^{n} Z_j K \bar{Z}_j \delta_p + Z_{n+1} G^{\#} \bar{Z}_{n+1}\delta_p\right),$$

or

$$\text{(B)} \qquad \sum_{j=1}^{n} Z_j K \bar{Z}_j \delta_p = -\delta_p - Z_{n+1} G^{\#} \bar{Z}_{n+1}\delta_p + H\delta_p,$$

which is equation (B), the second equation for the kernel K. In principle equation (B) is contained in equation (A), but in practice it provides additional information.

Step 2. Because of the symmetries of the problem it is reasonable to assume that the kernel K is invariant under Heisenberg translations and rotations in the first n coordinates and vanishes at infinity. These conditions mean that K must be of the form

$$K((z, t, \zeta), (w, s, \mu)) = \phi\left(|z - w|^2, t - s + 2\,\mathrm{Im}(z \cdot \bar{w}), \zeta, \mu\right),$$

for some function ϕ. We will treat ζ and μ as parameters and consider $\phi = \phi(v, \tau)$ as a function of

$$v = |z - w|^2, \qquad \tau = t - s - iz \cdot \bar{w} + iw \cdot \bar{z}.$$

Step 3. We begin by calculating how equations (A) and (B) restrict ϕ. Recall that if $Tf(z) = \int K(z, w) f(w)\, dw$ and X is a vector field with iX selfadjoint on the Heisenberg group, then

$$TXf(z) = \int K(z, w) X_w f(w)\, dw = -\int X_w K(z, w) f(w)\, dw.$$

We will use this identity below with $X = Z_j, \bar{Z}_j, T$. In equation (B) the term $K\bar{Z}_j$ occurs; applying the identity above with $X = \bar{Z}_j$, we find

$$K\bar{Z}_j = -\left(\frac{\partial}{\partial \bar{w}_j} - iw_j\frac{\partial}{\partial s}\right)\phi\left(|z-w|^2, t-s-iz\cdot\bar{w}+iw\cdot\bar{z}\right)$$
$$= (z_j - w_j)\phi_v + iz_j\phi_\tau - iw_j\phi_\tau = (z_j - w_j)\left(\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right)\phi.$$

Similarly, we calculate

$$Z_j\psi = \left(\frac{\partial}{\partial z_j} + i\bar{z}_j\frac{\partial}{\partial t}\right)\psi\left(|z-w|^2, t-s+2\,\mathrm{Im}\, z\cdot\bar{w}\right)$$
$$= \left(\overline{z_j - w_j}\right)\psi_v - i\bar{w}_j\psi_\tau + i\bar{z}_j\psi_\tau = \left(\overline{z_j - w_j}\right)\left\{\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right\}\psi.$$

Combining these expressions produces

$$Z_jK\bar{Z}_j = \left\{\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right\}\phi + |z_j - w_j|^2\left\{\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right\}^2\phi$$

and so

$$\sum_{j=1}^n Z_jK\bar{Z}_j = n\left\{\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right\}\phi + v\left\{\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right\}^2\phi.$$

We can now write equation (B) as

$$\text{(B}'\text{)} \qquad v\left(\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right)^2\phi + n\left(\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right)\phi = -\mathrm{I} - Z_{n+1}G^{\#}\bar{Z}_{n+1} + H.$$

Turning next to equation (A), we calculate

$$Z_jK = \left(\overline{z_j - w_j}\right)\left\{\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right\}\phi, \qquad \bar{Z}_jK = (z_j - w_j)\left\{\frac{\partial}{\partial v} - i\frac{\partial}{\partial \tau}\right\}\phi,$$

so that

$$Z_j\bar{Z}_jK = Z_j(z_j - w_j)\left\{\frac{\partial}{\partial v} - i\frac{\partial}{\partial \tau}\right\}\phi$$
$$= \left\{\frac{\partial}{\partial v} - i\frac{\partial}{\partial \tau}\right\}\phi + |z_j - w_j|^2\left\{\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right\}\left\{\frac{\partial}{\partial v} - i\frac{\partial}{\partial \tau}\right\}\phi$$

and after summing,

$$\sum_{j=1}^n Z_j\bar{Z}_jK = n\phi_v - in\phi_\tau + v\left\{\frac{\partial^2}{\partial v^2} + \frac{\partial^2}{\partial \tau^2}\right\}\phi.$$

A similar calculation shows that

$$\sum_{j=1}^n \bar{Z}_jZ_jK = n\phi_v + in\phi_\tau + v\left\{\frac{\partial^2}{\partial v^2} + \frac{\partial^2}{\partial \tau^2}\right\}\phi;$$

and therefore

$$-\frac{1}{2}\sum_{j=1}^{n}\left(Z_k\bar{Z}_k+\bar{Z}_kZ_k\right)K=-n\phi_v-v(\Delta_{v,\tau}\phi).$$

Since $-(n-2)TK=(n-2)i\phi_\tau$, we know

$$\Box_b K=-n\phi_v-v(\Delta_{v,\tau}\phi)+(n-2)i\phi_\tau;$$

and this allows us to write equation (A) as

$$\text{(A}'\text{)}\qquad -v\left(\frac{\partial^2}{\partial v^2}+\frac{\partial^2}{\partial\tau^2}\right)\phi-n\frac{\partial\phi}{\partial v}+i(n-2)\frac{\partial\phi}{\partial\tau}=\mathrm{I}+Z_{n+1}G^+\bar{Z}_{n+1}.$$

Adding equations (A′) and (B′) yields

$$2iv\frac{\partial^2\phi}{\partial v\partial\tau}-2v\frac{\partial^2}{\partial\tau^2}\phi+i(2n-2)\frac{\partial\phi}{\partial\tau}=H+Z_{n+1}(G^+-G^\#)\bar{Z}_{n+1},$$

and after integrating with respect to τ and dividing by $2i$, we get

$$\text{(4.i)}\quad v\left(\frac{\partial}{\partial v}+i\frac{\partial}{\partial\tau}\right)\phi+(n-1)\phi=\frac{1}{2i}\int\left(H+Z_{n+1}(G^+-G^\#)\bar{Z}_{n+1}\right)d\tau$$
$$=\alpha_n+\beta_n,$$

where

$$\text{(4.ii)}\qquad \alpha_n=\frac{1}{2i}\int Z_{n+1}(G^+-G^\#)\bar{Z}_{n+1}\,d\tau,$$

$$\text{(4.iii)}\qquad \beta_n=\frac{1}{2i}\int H\,d\tau.$$

This is our desired first-order equation for ϕ. Note that α_n has isotropic homogeneity and β_n has Heisenberg homogeneity.

3. Phong's theorem. We are trying to solve the boundary value problem (3):

$$\Box^+K((z,z_{n+1}),(w,w_{n+1}))=\delta_{(w,w_{n+1})},$$
$$\bar{Z}_{n+1}K((z,z_{n+1}),(w,w_{n+1}))=0\quad\text{on }\partial D.$$

Using the coordinates

$$(z,z_{n+1})\to(z,t,\zeta),\qquad t=\operatorname{Re}z_{n+1},\qquad \zeta=\operatorname{Im}z_{n+1}-|z|^2,$$
$$(w,w_{n+1})\to(w,s,\mu),\qquad s=\operatorname{Re}w_{n+1},\qquad \mu=\operatorname{Im}w_{n+1}-|w|^2,$$

we found in the last section that

$$K=\phi\left(|z-w|^2,t-s+2\operatorname{Im}(z\cdot\bar{w}),\zeta,\mu\right)\equiv\phi(v,\tau,\zeta,\mu),$$

where ϕ satisfies the first-order equation (4). In this section we will write down the first term of an expansion for K.

We will measure homogeneities by assigning the following weights:

	Weight 1	Weight 2
isotropic homogeneities:	τ,ζ,μ	v
Heisenberg homogeneities:	v,ζ,μ,τ	—

Now define

$$\Phi_- = 2v + |\zeta - \mu|^2 + \tau^2 \approx \operatorname{dist}^2((z, z_{n+1}), (w, w_{n+1})),$$
$$\Phi_+ = 2v + |\zeta + \mu|^2 + \tau^2, \qquad \Psi = v + \zeta + \mu - i\tau.$$

By the table, Φ_+, Φ_- are of isotropic weight 2 and Ψ is of Heisenberg weight 1.

THEOREM 2 (PHONG). *The kernel K has the expansion*

$$K = K^0 + (\textit{terms with weaker singularities}),$$

where

$$K^0 = c_n \Phi_-^{-n} + \frac{\theta_n + \lambda_n}{v^{n-1}}$$

and the terms θ_n, λ_n satisfy (with P and Q polynomials)

(i) $\theta_n = P(v, \tau, \zeta, \mu)/\Phi_+^{a_n}$ *has isotropic homogeneity* $\geqslant 2$,

(ii) $\lambda_n = Q(v, \tau, \zeta, \mu)/\Psi^{b_n}$ *has Heisenberg homogeneity* $\geqslant -2$,

(iii) $\theta_n + \lambda_n = 0$ *to order* $(n-1)$ *when* $v = 0$.

Moreover there is an algorithm for finding θ_n and λ_n for each dimension.

REMARK. As we have seen, the Neumann kernel $K(z, w)$ solves an elliptic PDE (well behaved under Euclidean dilations) under a boundary condition with a natural homogeneity under Heisenberg dilations. Phong's theorem explains the Heisenberg and Euclidean aspects of the kernel $K(z, w)$. In fact, K can be written as a sum $K = K_{\text{Euclidean}}(z, w) + K_{\text{Heisenberg}}(z, w)$ with each piece homogeneous under the appropriate dilations. The catch is that the two pieces $K_{\text{Euclidean}}(z, w)$ and $K_{\text{Heisenberg}}(z, w)$ have a spurious singularity which cancels when the two pieces are combined into the full kernel K. In order to write K without spurious singularities, one can state Phong's theorem in the form $K \sim \sum_l K^l_{\text{Euclidean}} \cdot K^l_{\text{Heisenberg}}$.

PROOF. We use the same notation and coordinates as in the last section. The equation for ϕ is

$$\left[v\left(\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right) + (n-1)\right]\phi = \alpha_n + \beta_n, \tag{5}$$

where

$$\alpha_n = \frac{1}{2i}\int Z_{n+1}(G^+ - G^\#)\overline{Z}_{n+1}\, d\tau, \qquad \beta_n = \frac{1}{2i}\int H\, d\tau.$$

Recall that for the Siegel domain D the Bergman kernel is

$$H = c_n\left(\frac{z_{n+1} - \overline{w}_{n+1}}{2i} - z \cdot \overline{w}\right)^{-(n+2)};$$

therefore

$$\beta_n = c(\Psi)^{-(n+1)}.$$

To handle the other term, recall that

$$Z_{n+1} = i\frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial t} - i\frac{\partial}{\partial \zeta}\right);$$

after computing G^+, $G^\#$ by the methods of Chapter 3, §7, we obtain

$$\alpha_n = \alpha_{n,1} + \alpha_{n,2} + (\text{terms of weaker homogeneity}),$$

where

$$\alpha_{n,1} = c(\Phi_-)^{-(n+1)}[(n-1)\Phi_- - 2nv],$$
$$\alpha_{n,2} = c(\overline{\Psi}_0)^2(\Phi_+)^{-(n+1)}, \qquad \Psi_0 = \zeta + \mu - i\tau.$$

Term	Isotropic Homogeneity	Heisenberg Homogeneity
α_n	yes	—
$\alpha_{n,1}$; $\alpha_{n,2}$	$-2n$	—
β_n	—	yes
Ψ_0	1	1

It is convenient to rewrite equation (5) in terms of $\tilde{\phi}$, where $\phi = v^{1-n}\tilde{\phi}$. We assume that $\tilde{\phi}$ vanishes to order $n-1$ at $v = 0$. Because

$$v\left(\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right)(v^{1-n}\tilde{\phi}) = v^{2-n}\left(\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right)\tilde{\phi} + (1-n)v^{1-n}\tilde{\phi},$$

we can write equation (5) as

$$\begin{aligned}\left[v\left(\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right) + (n-1)\right]\phi &= \left[v\left(\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right) + (n-1)\right](v^{1-n}\tilde{\phi}) \\ &= v^{2-n}\left(\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right)\tilde{\phi} = \alpha_n + \beta_n,\end{aligned}$$

or

$$\left(\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right)\tilde{\phi} = v^{n-2}(\alpha_n + \beta_n).$$

Neglecting terms of weaker homogeneity, this is

$$\text{(6.i)} \qquad \left(\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right)\tilde{\phi} = v^{n-2}(\alpha_{n,1} + \alpha_{n,2} + \beta_n).$$

We now solve (6) by constructing solutions η_n, θ_n, λ_n for the three equations

$$\text{(6.ii)} \qquad \left(\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right)\eta_n = v^{n-2}\alpha_{n,1} + (\text{weaker singularities}),$$

$$\text{(6.iii)} \qquad \left(\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right)\theta_n = v^{n-2}\alpha_{n,2} + (\text{weaker singularities}),$$

$$\text{(6.iv)} \qquad \left(\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right)\lambda_n = v^{n-2}\beta_n.$$

We begin with (6.ii). We claim that $c_n v^{n-1}\Phi_-^{-n}$ is a solution. This is verified by calculating

$$\left(\frac{\partial}{\partial v}+i\frac{\partial}{\partial \tau}\right)\left(c_n v^{n-1}\Phi_-^{-n}\right)$$
$$=(n-1)c_n v^{n-2}\Phi_-^{-n}-nc_n v^{n-1}\Phi_-^{-(n-1)}\cdot\left\{\left(\frac{\partial}{\partial v}+i\frac{\partial}{\partial \tau}\right)\Phi_-\right\}$$

and using the fact that

$$\left(\frac{\partial}{\partial v}+i\frac{\partial}{\partial \tau}\right)\Phi_-=\frac{\partial}{\partial v}\Phi_-+(\text{weaker singularities})$$
$$=2+(\text{weaker singularities}).$$

We turn next to the equation (6.iv):

$$\left(\frac{\partial}{\partial v}+i\frac{\partial}{\partial \tau}\right)\lambda_n=v^{n-2}\beta_n.$$

We will construct a solution λ_n by induction. If $n=2$, the right-hand side is Ψ^{-3} and we can take $\lambda_2=c\Psi^{-2}$. Since

$$\left(\frac{\partial}{\partial v}+i\frac{\partial}{\partial \tau}\right)\beta_n=c\beta_{n+1},$$

we can set

$$\lambda_{n+1}=c\left[(n-1)\lambda_n-v^{n-1}\beta_n\right],$$

so that

$$\left(\frac{\partial}{\partial v}+i\frac{\partial}{\partial \tau}\right)\lambda_{n+1}=c\left[(n-1)v^{n-2}\beta_n-(n-1)v^{n-2}\beta_n-v^{n-1}c\beta_{n+1}\right]$$
$$=cv^{n-1}\beta_{n+1},$$

as required.

This leaves equation (6.iii):

$$\left(\frac{\partial}{\partial v}+i\frac{\partial}{\partial \tau}\right)\theta_n=v^{n-2}\alpha_{n,2}+(\text{weaker singularities})$$

or

$$\frac{\partial}{\partial v}\theta_n=v^{n-2}\alpha_{n,2}+(\text{weaker singularities}).\tag{7}$$

Again we will construct a solution θ_n using induction on the dimension n. Note that the theorem will be proved if

(8)
(i) $(\theta_n+\lambda_n)$ vanishes to order $(n-1)$ at $v=0$.
(ii) θ_n is a rational function $P(v,\tau,\zeta,\mu)/\Phi_+^{\text{power}}$, with all terms having isotropic homogeneity $\geqslant -2$.

Since $(\partial/\partial v)\theta_2=\alpha_{2,2}$, we set

$$\theta_2=c\left(\overline{\Psi}_0\right)^2(\Phi_+)^{-2}$$

and note that at $v = 0$, $\Phi_+ = \Psi_0\overline{\Psi}_0$ and $\Psi = \overline{\Psi}_0$, so that

$$\theta_2 + \lambda_2 = 0,$$

as required. Define

$$\theta_n' = c\big[(n-1)\theta_n - v^{n-1}\alpha_{n,2}\big].$$

It is easy to check that equation (7) and property (8.ii) are satisfied; we will modify it so that property (8.i) holds. Note, by the induction step,

$$\theta_{n+1}' + \lambda_{n+1} = c\big[(n-1)\{\theta_n + \lambda_n\} - v^{n-1}\{\alpha_{n,2} + \beta_n\}\big]$$

vanishes to order $n-1$; if we correct θ_{n+1}' with something of higher homogeneity then equation (7) and property (8.ii) will not be affected. Put

$$b_{n+1} = \left[\left(\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right)^{n-1}(\theta_{n+1}' + \lambda_{n+1})\right]\Big|_{v=0},$$

which by (7) is

$$\begin{aligned} b_{n+1} &= \left(\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right)^{n-2}\left(v^{n-1}\alpha_{n+1,2} + (\text{weaker singularities}) + v^{n-1}\beta_{n+1}\right)\big|_{v=0} \\ &= \left(\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right)^{n-2} \\ &\quad (\text{weaker singularities})\,|_{v=0}. \end{aligned}$$

Since α_{n+1} has isotropic weight $-2(n+1)$, b_{n+1} has isotropic weight

$$\text{weight} > 2(n-1) - 2(n+1) - 2(n+2);$$

in other words, b_{n+1} has weight $\geqslant -2n+1$. We now define the solution θ_{n+1} by

$$\theta_{n+1} = \theta_{n+1}' + cv^{n-1}b_{n+1}^*,$$

where b_{n+1}^* is obtained from b_{n+1} by appropriately adjusting the denominator to be a power of Φ_+ (this only gives weaker singularities). Equation (7) and property (8.ii) still hold for θ_{n+1}, since $v^{n-1}b_{n+1}^*$ has isotropic weight

$$\text{weight} \geqslant 2(n-1) - 2n + 1 = -1.$$

Finally,

$$\left(\frac{\partial}{\partial v} + i\frac{\partial}{\partial \tau}\right)^{n-1}\left(\tilde{c}v^{n-1}b_{n+1}^* + \theta_{n+1}' + \lambda_{n+1}\right)\Big|_{v=0} = 0,$$

by the definition of b_{n+2}. This verifies property (8.i) and completes the proof of the theorem. $\square$

4. Approximate solution to $\square_b$ on strictly pseudoconvex domains. In this section we explain how to pass from the exact solution of $\square_b$ on the Heisenberg group to a good approximate solution for a general strictly pseudoconvex domain D with boundary $M = \{r(z) = 0\} \subseteq \mathbf{C}^{n+1}$. Our remarks summarize crudely the constructions in Folland-Stein [25]. The process is analogous to solving variable-coefficient elliptic equations by freezing coefficients. There are three main steps.

(i) Finding how closely the boundary M may be approximated by the Heisenberg group.
(ii) Transplanting the fundamental solution of $\square_b$ from the Heisenberg group to M.
(iii) Extracting useful information from the transplanted solution.

Now, the sharpest answer to (i) is to place M in Moser's normal form about a point $p \in M$. Actually, what we need for $\square_b$ is much simpler: After an analytic change of coordinates in a neighborhood of p, we may suppose $p = 0$ and M osculates the Heisenberg group $H^n = \partial(\text{Siegel domain}) \subseteq \mathbf{C}^{n+1}$ to third order at 0.

Since Chapter 9 discusses Moser's normal form, we omit here the construction of the change of coordinates. (What is involved is a calculation of Taylor series to order 3.)

Thus, there is a map $\Theta_p: M \to H^n$ which carries p to the origin and satisfies

(a) Θ_p is a local diffeomorphism.

(b) Θ_p is analytic (i.e., annhiliated by the $\bar\partial_b$-operator on M) to third order at p, when H^n is viewed as sitting inside $\mathbf{C}^{n+1}$.

Moreover, we may make Θ_p depend smoothly on p. We write $\Theta(x, y)$ for $\Theta_y(x)$. Thus $\Theta: M \times M \to H^n$. An important property of Θ is

(c) for x close to y in M, $\Theta(x, y)$ is very close to $[\Theta(y, x)]^{-1}$ in H^n.

EXAMPLE. If already $M = H^n$, then we can take $\Theta(x, y) = xy^{-1}$.

Next we carry out step (ii) by using Θ to transplant the fundamental solution of §1 from H^n to M. We start by picking a good metric on M. The natural compatibility condition is that the metric acting on tangent vectors of type $(1,0)$ be proportional to the Levi form, i.e.,

$$\|Z\|^2 = (\text{scalar})\cdot \sum_{jk} \frac{\partial^2 r}{\partial z_j \partial \bar z_k} \xi_j \bar\xi_k \text{ when } Z = \sum_{k=1}^{n+1} \xi_k \frac{\partial}{\partial z_k} \text{ is tangent to } M = \{r = 0\}.$$

Such a metric is called a *Levi metric.* Since the purpose of solving $\square_b$ is to solve $\bar\partial_b$ and thus construct analytic functions, we can restrict attention to Levi metrics without serious loss of generality.

Next we pick $(0,1)$-forms $\bar\omega^1, \ldots, \bar\omega^n$ locally on M so that the $\bar\omega^k$ form an orthonormal basis for the $(0,1)$-forms at each point of M. Here "orthonormal" refers to the Levi metric. The general $(0,1)$-form is $u = \sum_k u_k \bar\omega^k$, and the $\square_b$-equation for $(0,1)$-forms may be written as

$$\square_b\Big(\sum_k u_k \bar\omega^k\Big) = \sum_k f_k \bar\omega^k. \tag{9}$$

Recall that the solution for $\square_b$ on the Heisenberg group is $u_k(x) = \int_{H^n} K_\alpha(xy^{-1}) f_k(y)\, dy$ for $\alpha = n - 2$. Now we can guess an approximate solution to (9), namely

$$u_k^0(x) = \int_M K_\alpha(\Theta(x, y)) f_k(y)\, d\mu(y) \tag{10}$$

for a suitable smooth measure μ on M. Step (ii) is complete.

Finally, we want to make use of our approximate solution. The obvious starting point is to apply $\square_b$ to (10); one finds that $\square_b u^0 = f + \varepsilon f$, where $\varepsilon f = \Sigma_k(\varepsilon_k f)\overline{\omega}^k$ and $\varepsilon_k f(x)$ is a sum of expressions of the form

$$\int_M G(x, y)K'(\Theta(x, y))f_l(y)\, d\mu(y) \tag{11}$$

with $G \in C^\infty(M \times M)$ and K' homogeneous on the Heisenberg group with a singularity weaker than δ_0, the delta mass at the origin. Now, εf may be regarded as a small perturbation of the identity, because $f = \int_M \delta_0(\Theta(x, y))f(y)\, dy$ and δ_0 is homogeneous on H^n. Thus, the integral kernel in (11) has a weaker singularity than the identity.

There are two natural strategies to exploit this. We can either use successive approximations starting with (10) to build ever more accurate approximate solutions of $\square_b u = f$, or we can define the analogues of Sobolev spaces for $\square_b$ and prove that the integral operator (11) carries H^s to $H^{s'}$ ($s' > s$). Either approach yields C^∞ regularity of solutions of $\square_b$ and much more. The result of calculations with successive approximations is as follows.

THEOREM 3. *$\square_b(\Sigma_k u_k\overline{\omega}^k) = (\Sigma_k f_k\overline{\omega}^k)$ may be solved modulo C^∞-smoothing errors by $u_j(x) = \Sigma_k \int_M K_{jk}(x, y)f_k(y)\, d\mu(y)$ with $K_{jk}(x, y) = \delta_{jk}K_\alpha(\Theta(x, y)) +$ an asymptotic series of weaker singularities.*

The natural Sobolev spaces for $\square_b$ on the Heisenberg group are

$$H^s_{\text{Heisenberg}} = \left\{ f \in L^2(H^n) \;\middle|\; \left(\prod_k X_k^{\mu_k}\right)\left(\prod_k Y_k^{\nu_k}\right)T^\alpha f \in L^2 \right.$$
$$\left. \text{whenever } \sum_k \mu_k + \sum_k \nu_k + 2\alpha \leqslant s \right\}$$

for s a positive integer. For s not a positive integer the definition is harder and will not be given here. Note that these Sobolev spaces are natural for the non-Euclidean dilations on the Heisenberg group.

On a strictly pseudoconvex boundary M, we define H^s_M for s a positive integer to consist of all $f \in L^2(M)$ such that $Z_1 Z_2 \cdots Z_m \overline{W}_1 \overline{W}_2 \cdots \overline{W}_l T_1 \cdots T_q f \in L^2$ whenever the Z's and W's are of type (1, 0), the T's are arbitrarily smooth vector fields, and $m + l + 2q \leqslant s$. Again there is a natural extension to s not a positive integer.

THEOREM 4. *The mapping $f \to \int_M K_\alpha(\Theta(x, y))f(y)\, d\mu(y)$ carries H^s_M to H^{s+2}_M, while $f \to \varepsilon f$ carries H^s_M to H^{s+1}_M.*

COROLLARY. *$\square_b u = f \in H^s_M$ implies $u \in H^{s+2}_M$.*

See Folland-Stein [25] for the proofs.

We should note that Phong [53] gave an approximate solution of the Neumann equation $\square u = \omega$ for $\bar\partial$ on general strictly pseudoconvex domains D. The transplantation from H^n to D follows the same philosophy as for $\Box_b$, but this time is harder technically because of the presence of the boundary and the two interacting homogeneities on D. As in the Heisenberg case, the integral kernel for Phong's approximate solution of $\Box u = \omega$ is a sum of products of purely "Euclidean" and purely "Heisenberg" kernels. The approximate solution u^0 satisfies the boundary conditions exactly, and $\Box u^0 = \omega + \varepsilon\omega$ where ε has a weak singularity.

So far we have described how to use the Heisenberg group for strictly pseudoconvex domains. However, we left out a very important ingredient in the application of nilpotent Lie groups to a broader class of problem: the idea of "lifting" (Rothschild-Stein **[54]**), which we now briefly describe.

The key step in solving $\Box_b u = f$ on a strictly pseudoconvex boundary was to approximate $\Box_b$ by a translation-invariant operator on the Heisenberg group. For more general differential equations, this idea apparently cannot be applied. Consider the Grusin operator $L = (\partial/\partial x)^2 + (x\partial/\partial y)^2$ on $\mathbf{R}^2$; this is elliptic at $x \neq 0$ and degenerate at $x = 0$. On the other hand, a translation-invariant operator on a Lie group looks the same at all points and so cannot provide a good approximation to L. The way out of this dilemma is to introduce a new variable z and study the "lifted" operator $\tilde{L} = (\partial/\partial x)^2 + (x\partial/\partial y + \partial/\partial z)^2$ on functions on $\mathbf{R}^3$. The new operator $\tilde{L}$ behaves the same at all points; in fact, after a change of coordinates, one can realize $\tilde{L}$ as a translation-invariant operator on the Heisenberg group. Now, $\tilde{L}$ and L are equal when they act on functions independent of z. So any regularity theorem for $\tilde{L}$ automatically yields a regularity theorem for L. Also, one can build a fundamental solution for L by starting with the (essentially known) fundamental solution for $\tilde{L}$ and integrating out the extra variables z.

A similar construction can be carried out for any Hörmander operator **[32]** $L = \sum_j X_j^2 + X_0$, where the X_j's are noncommuting vector fields. At any point p we can introduce extra variables $z_1, z_2, \ldots, z_N$ and lift L to an operator $\tilde{L}$ which is well approximated about p by a translation-invariant operator $\tilde{L}_0$ on a nilpotent Lie group. From this follow sharp regularity theorems for L, and a parametrix for L may be written as the integral of a fundamental solution for $\tilde{L}_0$ over the z_k. A complete discussion of these matters appears in **[54]**.

Before the work of Folland-Stein and Phong, the C^∞-regularity for $\bar{\partial}$, $\bar{\partial}_b$, $\Box$, $\Box_b$ on strictly pseudoconvex domains was established by Kohn **[39]** using energy estimates. More recently, by a powerful extension of his earlier techniques, Kohn **[40]** and his students Catlin **[9]** and D'Angelo **[13]** (see also Diederich-Pflug **[14]**) achieved a deep understanding of C^∞-regularity of $\bar{\partial}$-Neumann on weakly pseudoconvex domains. Roughly speaking, there is C^∞-regularity for the $\bar{\partial}$-Neumann problem for (0, 1)-forms on D if and only if no complex-analytic curve sits inside the boundary of D. It would be a major achievement to write down the basic fundamental solutions for the Bergman or Szegö kernel for any nontrivial class of weakly pseudoconvex domains. Only a few simple examples are known.

We should mention also the important work of Trèves **[62]** and Tartakoff **[60]** on real-analyticity of solutions of $\bar\partial$. To understand this requires the real-analytic analogue of Fourier integral operators, which we have not treated here.

Finally, we refer the reader to Chapter 12, where the idea of transplantation recurs in connection with the Bergman and Szegö kernels.

CHAPTER 8. ELLIPSOIDS AND THE METHOD OF REFLECTIONS

1. Introduction. Let D, D' be two bounded strictly pseudoconvex domains in $\mathbf{C}^n$ with real-analytic boundaries, say

$$D = \{z \in \mathbf{C}^n \mid r(z) > 0\}, \qquad D' = \{z' \in \mathbf{C}^n \mid r'(z') > 0\} \tag{1}$$

where r, r' are real-valued convergent power series in z, $\bar z$. We shall study here whether D and D' are biholomorphic. Recall this means $\Phi\colon D \to D'$ for some complex analytic one-to-one, onto map Φ. The known methods of classifying domains under biholomorphic maps start with the following result.

THEOREM 1. *If D, D' are strictly pseudoconvex with smooth boundaries, then any biholomorphic map $\Phi\colon D \to D'$ is smooth up to the boundary. If the boundaries ∂D, $\partial D'$ are real-analytic, then Φ continues analytically to a biholomorphic map of a neighborhood of $\bar D$ onto a neighborhood of $\bar D'$.*

A simple proof of Theorem 1 will be given in §4. Now in the real-analytic case we can regard Φ as defined in a neighborhood of D, and therefore

$$\left\{\begin{matrix} r'(z') = \mu(z) r(z) \\ z' = \Phi(z) \end{matrix}\right\}, \tag{2}$$

where $\mu(z)$ is a convergent power series in z, $\bar z$, and $\mu(z) \neq 0$ for z near D.

The idea of the method of reflections is to pass from $r(z)$ to the analytic function $r(z, \bar w)$ of $z, \bar w \in \mathbf{C}^n$ such that $r(z, \bar z) = r(z)$. The power series for $r(z, \bar w)$ is obtained from that of $r(z)$ just by substituting $\bar w$ in place of $\bar z$. In particular, $r(z, \bar w)$ is uniquely determined, by $r(z)$. From (2) we obtain

$$\left\{\begin{matrix} r'(z', \bar w') = \mu(z, \bar w) r(z, \bar w) \\ z' = \Phi(z),\ w' = \Phi(w) \end{matrix}\right\}, \tag{3}$$

where $\mu(z, \bar w) \neq 0$ for z, w near the same boundary point $p \in \partial D$. In particular $r'(z', \bar w') = 0$ if and only if $r(z, \bar w) = 0$ for $z' = \Phi(z)$, $w' = \Phi(w)$, z and w near $p \in \partial D$. To $w \in \mathbf{C}^n$ near p we associate the analytic variety $Q_w = \{z \mid r(z, \bar w) = 0\}$. The family $\{Q_w \mid w \in \mathbf{C}^n\}$ of codimension-one varieties in $\mathbf{C}^n$ is thus invariantly associated to the domain D. That is, if $\Phi\colon D \to D'$ is biholomorphic and $\{Q_w\}$, $\{Q'_{w'}\}$ are the families of varieties associated to D, D', then $\Phi\colon Q_w \to Q'_{\Phi(w)}$. In $\mathbf{C}^n$ $(n > r)$, these subvarieties are a powerful tool to classify domains. It was S. Webster who realized how much could be done by this technique, which is called the method of reflections.

In §2 we prove

THEOREM 2 (WEBSTER). *Assume the functions $r(z)$, $r'(z')$ defining the domains $D, D' \subseteq \mathbf{C}^n$ in (1) are polynomials. If $n \geqslant 2$ and Φ: $D \to D'$ is biholomorphic, then Φ is an algebraic map, i.e. its components are algebraic functions.*

Given specific D, D' it is hard to decide whether a biholomorphic map exists. In §3 we consider the classification of ellipsoids, i.e. domains in $\mathbf{C}^n$ ($n \geqslant 2$) of the form

$$\sum_{j,k=1}^{n} \left(a_{jk}x_jx_k + b_{jk}y_jy_k + c_{jk}x_jy_k\right) < d, \tag{4}$$

where $z_j = x_j + \sqrt{-1}\,y_j$ and the quadratic form is positive. We prove

THEOREM 3 (WEBSTER). *If two ellipsoids $D, D' \subseteq \mathbf{C}^n$ ($n \geqslant 2$) are biholomorphic, then already they are equivalent by a linear transformation of $\mathbf{C}^n$. If D, D' are not biholomorphic to the unit ball, then any biholomorphic map Φ: $D \to D'$ is linear.*

Theorems 2 and 3 will be proved by the method of reflections.

EXAMPLE. Consider the ellipsoid

$$\left\{r(z) = |z|^2 + \sum_{j=1}^{n} A_j\left(z_j^2 + \bar{z}_j^2\right) - 1 < 0\right\}.$$

Then

$$r(z,\bar{w}) = z \cdot \bar{w} + \sum_{j=1}^{n} A_j\left(z_j^2 + \bar{w}_j^2\right) - 1$$

and

$$w \mapsto Q_w = \left\{z \in \mathbf{C}^n \,\middle|\, z \cdot \bar{w} + \sum_{j=1}^{n} A_j z_j^2 = 1 - \sum_{j=1}^{n} A_j \bar{w}_j^2\right\}.$$

An analytic automorphism of this ellipsoid must carry each Q_w to some $Q_{w'}$. This places strong restrictions on the automorphism.

We conclude this section with a quick review of algebraic and rational maps. Recall that a (multiple-valued) analytic function $w = F(z_1 \cdots z_n)$ defined in the complement of a variety $V \subseteq \mathbf{C}^n$ is called *algebraic* if the polynomial equation $P(w, z_1,\dots,z_n) = 0$ holds, whenever $w = F(z_1,\dots,z_n)$, $z_1,\dots,z_n \in \mathbf{C}^n \setminus V$. Thus w can be obtained from $(z_1,\dots,z_n)$ by solving a polynomial equation. A map Φ: $(z_1,\dots,z_n) \to (w_1,\dots,w_n)$ is called algebraic if each w_j is an algebraic function of $(z_1,\dots,z_n)$. The composite of two algebraic maps is again algebraic; if Φ is algebraic and invertible, then Φ^{-1} is algebraic. We omit the proofs. A map Φ: $(z_1,\dots,z_n) \to (w_1,\dots,w_n)$, single valued and well defined outside a subvariety $V \subseteq \mathbf{C}^n$ is called *rational* if each w_j is a rational function of $(z_1,\dots,z_n)$. Φ is called *birational* if Φ and Φ^{-1} are both rational. The next proposition is needed in §3; we prove it here since it clarifies the relation between algebraic and rational maps.

PROPOSITION 1. *Let* $F\colon \mathbf{C}^n \setminus V \to \mathbf{C}^1$ *be a single-valued algebraic function. Then* F *is rational.*

PROOF. When $w = F(z)$, $z \in \mathbf{C}^n \setminus V$; we have

$$P(w, z) = \sum_{k=0}^{m} P_k(z) w^k \doteq 0,$$

by definition. (Note that $m \geq 1$, $P_m \not\equiv 0$.) This gives us a bound

$$|w| \leq C(1 + |z|)^M / P_m(z).$$

So $G(z) \equiv P_m(z)F(z)$ is a single-valued analytic function of polynomial growth on $\mathbf{C}^n \setminus V$. By the removable singularity theorem, G extends to an analytic function of polynomial growth on $\mathbf{C}^n$. So G must be a polynomial, and $F(z) = G(z)/P_m(z)$ is a rational function. $\square$

2. The proof of Webster's theorem on algebraic mappings. In this section we prove Theorem 2. We return to the notation of §1. Thus $\Phi\colon D \to D'$ is biholomorphic and extends to a neighborhood of D; we have the families $\{Q_w\}$, $\{Q'_{w'}\}$ associated to D, D', and we know that $\Phi\colon Q_w \to Q'_{w'}$ for $w' = \Phi(w)$, w near ∂D.

The first step in proving Theorem 2 is to show that the restriction of Φ to Q_w is algebraic. To see this, we begin with an observation: Since $r(z, \bar{z})$ is real, it follows by analytic continuation that $r(z, \bar{w}) = r(w, \bar{z})$. In particular $y \in Q_w$ means $r(y, \bar{w}) = 0$, hence $r(w, \bar{y}) = 0$ so that also $w \in Q_y$. Now to $y \in Q_w$ we associate $T_w(Q_y)$, the tangent space of Q_y at w. $T_w(Q_y)$ is a hyperplane in $\mathbf{C}^n$, given in coordinates $(p_1, \ldots, p_n)$ as $\{(\zeta_1, \ldots, \zeta_n) \in \mathbf{C}^n \mid \Sigma_1^n p_k \bar{\zeta}_k = 0\}$. Since $(p_1, \ldots, p_n)$ and $(\lambda p_1, \ldots, \lambda p_n)$ define the same hyperplane ($\lambda \neq 0$), we may regard $T_w(Q_y)$ as a point in complex projective space $\mathbf{C}P^{n-1}$. So we have a natural map $\sigma_w\colon Q_w \to \mathbf{C}P^{n-1}$.

PROPOSITION 2. *The map* σ_w *is a local diffeomorphism of* Q_w *and an algebraic map.*

PROOF. Immediately from the definition we get

$$\sigma_w(y) = (p_1, \ldots, p_n) \quad \text{with } p_k = \partial r(y, \bar{w})/\partial \bar{w}_k,$$

so that $\sigma_w\colon y \to (p_1, \ldots, p_n)$ is a polynomial map. To see that σ_w is a diffeomorphism locally near ∂D, we check the differential $\sigma'_w\colon T_y(Q_w) \to T(\mathbf{C}P^{n-1})$ at $y = w = z^0 \in \partial D$. We can pick coordinates in $\mathbf{C}^n$ so that

$$\frac{\partial r}{\partial z_k}\left(z^0, \overline{z^0}\right) = \begin{Bmatrix} 1 & \text{if } k = n \\ 0 & \text{if } k \neq n \end{Bmatrix}.$$

Then

$$T_{z^0}(Q_{z^0}) = \{(\zeta_1, \ldots, \zeta_n) \mid \zeta_n = 0\},$$

and

$$\sigma_{z^0}(z^0) = (p_1, \ldots, p_n) = (0, 0, \ldots, 0, 1) \in \mathbf{C}P^{n-1},$$

so that p_k/p_n $(k = 1,\ldots,n-1)$ serve as local coordinates on $\mathbf{C}P^{n-1}$. One computes in these coordinates that σ_w' is the matrix $(\partial^2 r(z^0,\overline{z^0})/\partial z_j \partial \overline{w}_k)_{i \leq j,k \leq n-1}$, which is nonsingular by the strict pseudoconvexity of D. Thus σ_w is a local diffeomorphism. □

Now we can show that $\Phi|_{Q_w}$ is algebraic. In fact we have a commutative diagram:

$$\begin{array}{ccc} Q_w & \xrightarrow{\Phi} & Q_{w'} \\ \downarrow \sigma_w & & \downarrow \sigma_{w'} \\ \mathbf{C}P^{n-1} & \xrightarrow[\Phi'(w)]{} & \mathbf{C}P^{n-1} \end{array}$$

The map on the bottom takes hyperplanes in $T_w(\mathbf{C}^n)$ to hyperplanes in $T_{w'}(\mathbf{C}^n)$ by composing with the differential $\Phi'(w)$. In homogeneous coordinates $(p_1,\ldots,p_n)$, this map is linear; in local coordinates p_k/p_n, it is a linear fractional transformation. So the map $\Phi'(w)$ is certainly algebraic. Proposition 2 shows that σ_w, $\sigma_{w'}$ are also algebraic. Consequently $\Phi|_{Q_w} = \sigma_{w'}^{-1}\Phi'(w)\sigma_w$ is algebraic, as claimed.

Now we shall prove that Φ is algebraic. We begin with $D \subseteq \mathbf{C}^2$. Fix $z^0 \in \partial D$ and two points p^1, p^2 on Q_{z^0} in general position (to be defined soon). For z near z^0, we have

$$Q_z \cap Q_{p^1} = \{w^1\}, \qquad Q_z \cap Q_{p^2} = \{w^2\}.$$

Consider

(5) $$D \to Q_{p^1} \times Q_{p^2}, \qquad z \mapsto (w^1, w^2),$$

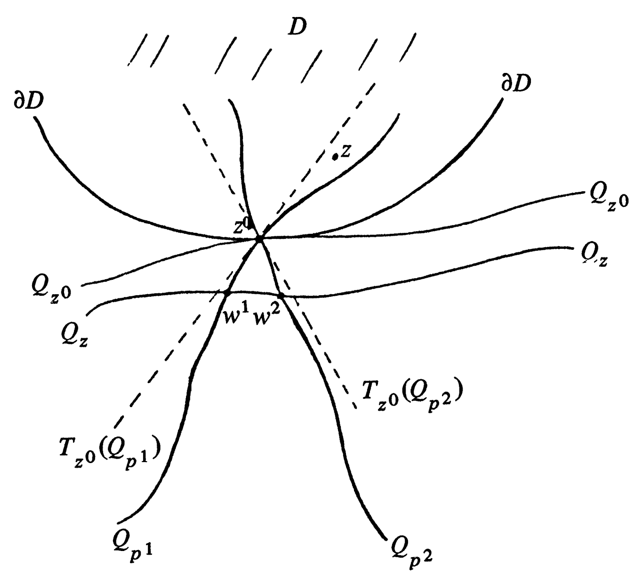

FIGURE 1

which is defined for z near z^0. The points p^1, p^2 are said to be in *general position* if (5) is a local diffeomorphism. Indeed for z sufficiently close to z^0, the map (5) is well approximated by $D \to T_{z^0}(Q_{p^1}) \times T_{z^0}(Q_{p^2})$, $z \mapsto (\bar{w}_1, \bar{w}_2)$, where $\bar{w}_i$ is defined by $Q_z \cap T_{z^0}(Q_{p^i}) = \{\bar{w}_i\}$. Proposition 2 shows that we may pick p^1, p^2 to make the tangent spaces $T_{z^0}(Q_{p^i})$ transverse, and a computation of the derivative of (5) confirms that p^1, p^2 are in general position. See Figure 1. Note that (5) is an algebraic map, since w^i is found by solving the simulataneous polynomial equations

$$r(w^i, \bar{z}) = 0, \qquad r(w^i, \bar{p}^i) = 0.$$

Putting $p^{i'} = \Phi(p^i)$ and repeating the construction yields an algebraic map $D' \to Q_{p^{1'}} \times Q_{p^{2'}}$, which is a local diffeomorphism. Since $\Phi\,|_{Q_{p^i}}$ is known to be algebraic, we have the commutative diagram

$$\begin{array}{ccc} D & \xrightarrow{\quad\Phi\quad} & D' \\ \downarrow & & \downarrow \\ Q_{p^1} \times Q_{p^2} & \xrightarrow[\Phi|_{Q_{p^1}} \times \Phi|_{Q_{p^2}}]{} & Q_{p^{1'}} \times Q_{p^{2'}} \end{array}$$

with all maps except Φ known to be algebraic. Hence Φ is algebraic as required. This completes the proof when $n = 2$.

For $D \subseteq \mathbf{C}^n$, fix $z^0 \in \partial D$ and choose $p^1,\ldots,p^n \in Q_{z^0}$ in general position (to be defined shortly). For each $i = 1,\ldots,n$, the variety $\mathcal{Q}_i = \bigcap_{j \neq i} Q_{p^j}$ is one dimensional, so we can define w^i by $Q_z \cap \mathcal{Q}_i = \{w^i\}$ for $z \in D$ near z^0. As before, the map $D \to \mathcal{Q}_1 \times,\ldots,\times \mathcal{Q}_n$, $z \mapsto (w^1,\ldots,w^n)$ is defined for z near z^0, and the points p^i are said to be in general position if it is a local diffeomorphism. This map is algebraic, and the rest of the argument is the same as in the two-dimensional case. This completes the proof of Theorem 2. $\square$

3. Ellipsoids and biholomorphic maps between them. Recall the boundary of an ellipsoid in $\mathbf{C}^n$ has the form

$$\sum_{j,k=1}^{n} \left(a_{jk}x_jx_k + b_{jk}y_jy_k + c_{jk}x_jy_k\right) = 1, \tag{6}$$

where $z_j = x_j + \sqrt{-1}\,y_j \in \mathbf{C}$. We can rewrite this using the conjugate variables z_j, $\bar{z}_j$ as $\langle z, \bar{z}\rangle + Q(z, z) + \overline{Q}(\bar{z}, \bar{z}) = 1$, where $\langle\,\cdot\,\rangle$ is a Hermitian form and $Q(\cdot\,,\,\cdot)$ is a quadratic form. The positivity of the quadratic form (6) implies that $\langle\,,\rangle$ is positive-definite. (Apply (6) to $e^{i\theta}z$ and average over θ.) Therefore, we can diagonalize it by a complex linear transformation, and our ellipsoid becomes

$$\sum_{j=1}^{n} |z_j|^2 + Q(z, z) + \overline{Q}(\bar{z}, \bar{z}) = 1.$$

Now consider the restriction of the real quantity $Q(z, z) + \overline{Q}(\bar{z}, \bar{z})$ to the unit sphere in $\mathbf{C}^n$. After a unitary transformation, we may assume this quantity is maximized at $z = (1, 0, \ldots, 0) \in$ Unit Sphere. This unitary transformation puts $Q(z, z) + \overline{Q}(\bar{z}, \bar{z})$ in the special form $A_1(z_1^2 + \bar{z}_1^2) + Q'(z', z') + \overline{Q}'(\bar{z}', \bar{z}')$, where $(z_1, \ldots, z_n) = (z_1, z')$. We can apply the same maximization argument to $Q' + \overline{Q}'$ in fewer variables, and proceed inductively. This proves

PROPOSITION 3. *The boundary of an ellipsoid $D \subseteq \mathbf{C}^n$ can be written*

$$\sum_{j=1}^{n} |z_j|^2 + \sum_{j=1}^{n} A_j\left(z_j^2 + \bar{z}_j^2\right) - 1 = 0$$

after a complex linear transformation.

PROOF OF THEOREM 3. Suppose that Φ: $D \to D'$ is a biholomorphic map between ellipsoids in $\mathbf{C}^n$. We may assume D, D' are given as in Proposition 3 with coefficients $\{A_i\}$, $\{A_i'\}$ respectively. The proof of Webster's theorem proceeds in several steps.

Step 1. Φ is algebraic.

Step 2. Φ is birational.

Step 3. The coefficients $\{A_i\}$ and $\{A_i'\}$ agree.

Step 4. Φ is linear unless all the A_j vanish.

We already know that Φ is algebraic: this follows from Theorem 2. We begin with

Step 2. Since Φ is algebraic, we can think of it as a globally-defined, possibly multiple-valued map Φ: $\mathbf{C}^n \setminus V \to \mathbf{C}^n$, where V is a subvariety of $\mathbf{C}^n$. By a careful analytic continuation, we shall show that Φ has a single-valued branch on $\mathbf{C}^n \setminus V$. Proposition 1 will then imply that Φ is rational. Applying the same argument to Φ^{-1} shows that Φ is birational, as desired. So it is enough to define Φ as a single-valued function globally on $\mathbf{C}^n \setminus V$.

Let $\Phi(z) = (w_1, \ldots, w_n)$ satisfy polynomial equations

$$P^j(z, w_j) = \sum_{k=0}^{m_j} P_k^j(z)(w_j)^k = 0, \qquad z \in \mathbf{C}^n \setminus V. \tag{7}$$

After possibly enlarging V (throw in $z \in \mathbf{C}^n$ where $w_j \mapsto P^j(z, w_j)$ has discriminant zero) we may assume that locally there is only one solution w of the equations $P^j(z, w_j) = 0$.

To show that Φ is globally single valued, we consider the analytic continuation of Φ along a path $\gamma \subseteq \mathbf{C}^n \setminus V$ connecting z_{initial} and z_{final}. That is, we continue each coordinate function w_j along the path according to the equation $P^j(z, w_j) = 0$; denote this by $z'(t) = \Phi(\gamma(t))$. If we can prove that $z'_{\text{final}} = \Phi(z_{\text{final}})$ is independent of the path γ, then we have shown that $\Phi(z)$ is single valued as desired.

It is sufficient to prove this for γ a generator of $\pi_1(\mathbf{C}^n \setminus V)$. To describe a set of generators, recall that a complex line l in $\mathbf{C}^n$, in general position, intersects the variety V in a finite number of points B_i: $l \cap V = \{B_1, \ldots, B_n\}$. It is not difficult to show that $\pi_1(\mathbf{C}^n \setminus V)$ is generated by those loops $t \to \gamma(t) \in \mathbf{C}^n \setminus V$ that originate at a fixed basepoint z_{initial}, travel to a neighborhood of B_i, encircle B_i once inside l, and then return to z_{initial} along the original path to $z_{\text{final}} = z_{\text{initial}}$.

(See Figure 2.) We can assume that $z_{\text{initial}} \in \partial D$, where D is our ellipsoid. Let $\tilde{V} = \{y \mid Q_y \subseteq V\}$; this is a proper subvariety of $\mathbf{C}^n$. Note that we can arrange γ so that

$$\gamma(t) \notin \tilde{V}. \tag{8}$$

We will need (8) later.

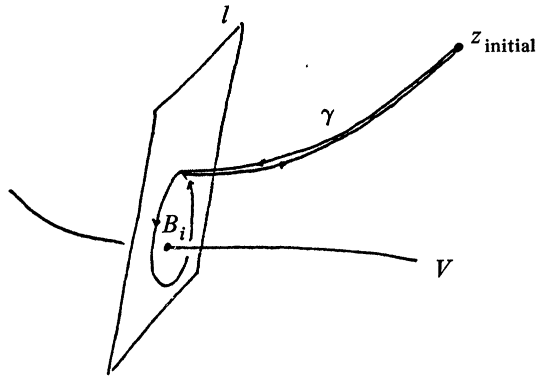

FIGURE 2

LEMMA 1 *Let w be a point of $Q_{z_{\text{initial}}}$ near z_{initial}. After possibly reparametrizing γ, we can find a closed loop $\gamma^\# = \{t \to \gamma^\#(t)\} \subseteq \mathbf{C}^n \setminus V$ starting at w and satisfying*

$$\text{(9)} \qquad \begin{aligned} &\text{(i)} \quad r\left(\gamma(t), \overline{\gamma^\#(t)}\right) \equiv 0, \\ &\text{(ii)} \quad \gamma^\# \textit{ is homotopically trivial in } \mathbf{C}^n \setminus V. \end{aligned}$$

We shall prove this shortly. For now, we show that Lemma 1 implies $z'_{\text{initial}} = z'_{\text{final}}$, where $z'_{\text{initial}} = \Phi(z_{\text{initial}})$. Indeed, let $w'(t) = \Phi(\gamma^\#(t))$ be the analytic continuation of Φ along $\gamma^\#$; this is a closed loop by virtue of (9ii). Now (9i) yields $w' = w'_{\text{final}} \in Q'_{z'_{\text{final}}}$. Thus, if w' belongs to a neighborhood of z'_{initial} in $Q'_{z'_{\text{initial}}}$, then w' belongs to $Q'_{z'_{\text{final}}}$. Since any small piece of $Q'_{z'}$ already determines z', we can conclude that $z'_{\text{initial}} = z'_{\text{final}}$. Thus Φ is single valued, hence birational. All that remains is to prove Lemma 1.

PROOF OF LEMMA 1. We can assume that γ makes a very small loop around the point B_1. Thus $\gamma(t)$ travels from z_{initial} to l during the time interval $0 \leqslant t \leqslant t_0$, then $\gamma(t)$ loops around B_1 during $t_0 \leqslant t \leqslant t_1$, and finally $\gamma(t)$ returns from l to z_{initial} during $t_1 \leqslant t \leqslant 1$. We first define $\gamma^\#(t)$ for $0 \leqslant t \leqslant t_0$. To do so, it is enough to assume $\gamma^\#(t)$ has been constructed for $0 \leqslant t \leqslant \tau$, and then shows how to extend $\gamma^\#(t)$ into $0 \leqslant t \leqslant \min\{\tau + h, t_0\}$ with h independent of τ. There are two cases.

Easy case. If $\gamma^{\#}(\tau)$ is far from V, then we just continue $\gamma^{\#}$ in any reasonable fashion to satisfy (9i), and $\gamma^{\#}$ will continue to avoid V.

Delicate case. If $\gamma^{\#}(\tau)$ is near V, then by reparametrizing γ we may assume that $\gamma(t)$ remains constant for a short time, starting at $t = \tau$. During that short time, we let $\gamma^{\#}(t)$ move inside $Q_{\gamma(\tau)}$, starting at $\gamma^{\#}(\tau)$ and ending at some point far from V. This is possible by virtue of (8). Now we can proceed as in the easy case. So $\gamma^{\#}(t)$ can be defined for $0 \leqslant t \leqslant t_0$. Moreover, by repeating the argument of the Delicate case above, we can suppose that $\gamma^{\#}(t_0)$ is far from V.

Next we define $\gamma^{\#}(t)$ for $t_0 \leqslant t \leqslant t_1$. Since the loop in $\gamma(t)$ is small, we have $t_1 - t_0 \ll 1$, so we can continue $\gamma^{\#}(t)$ into $0 \leqslant t \leqslant t_1$ so that (9i) holds, and $\gamma^{\#}(t)$ still misses V. Although $\gamma(t_0) = \gamma(t_1)$, it may happen that $\gamma^{\#}(t_1) \neq \gamma^{\#}(t_0)$. To correct this, we reparametrize γ so that $\gamma(t)$ is constant for a short time starting at $t = t_1$; during this short time we can let $\gamma^{\#}(t)$ move back to $\gamma^{\#}(t_0)$, staying within $Q_{\gamma(t_0)}$. Thus, we have defined $\gamma^{\#}(t)$, $0 \leqslant t \leqslant t_1$. Note that $\gamma^{\#}(t_0) = \gamma^{\#}(t_1)$ now, and the loop $\gamma^{\#}(t)$, $t_0 \leqslant t \leqslant t_1$ is contained in a small ball in $\mathbf{C}^n \setminus V$. Finally, we have to define $\gamma^{\#}(t)$ for $t_1 \leqslant t \leqslant 1$. This is easy—since $\gamma(t)$ retraces its initial path during $t_1 \leqslant t \leqslant 1$, we can simply let $\gamma^{\#}(t)$ do the same. Now we have defined $t \to \gamma^{\#}(t) \in \mathbf{C}^n \setminus V$ for $0 \leqslant t \leqslant 1$. We know that (9i) holds, and (9ii) is obvious, since the part of $\gamma^{\#}$ from $t = t_0$ to $t = t_1$ can be shrunk to a point. $\square$

Step 3. In this step we prove $\{A_i\} = \{A_i'\}$. Recall that

$$D = \{z \in \mathbf{C}^n \mid r(z, \bar{z}) = 0\}, \qquad D' = \{z' \in \mathbf{C}^n \mid r'(z', \bar{z}') = 0\},$$

and if $z' = \Phi(z)$, $w' = \Phi(w)$, then

$$r'(z', \overline{w}') = \mu(z, \overline{w}) r(z, \overline{w}). \tag{10}$$

So far, we have been using Q_w, defined by the vanishing of $r(z, \overline{w})$. Now we are going to make use of the varieties defined by the vanishing of a linear combination $\Sigma_j \lambda_j r(z, \overline{w}_j)$. The key observation that makes this possible is the following.

LEMMA 2. *If Φ is birational, then $\mu(z, \overline{w}) = \beta(z)\overline{\beta(w)}$ for a rational function $\beta(\cdot)$.*

PROOF. Since $\mu(z, \overline{w}) = r'(\Phi(z), \overline{\Phi(w)})/r(z, w)$ with Φ rational, we know that $\mu(z, \overline{w})$ is rational. Similarly if we define $\mu'(z', \overline{w}')$ by $r(z, \overline{w}) = \mu'(z', \overline{w}')r'(z', \overline{w}')$, then $\mu'(z', \overline{w}')$ is rational and satisfies $\mu(z, \overline{w})\mu'(z', \overline{w}') = 1$. Since Φ is rational, $z' = F(z)/g(z)$, where $F(z) = (f_1(z), \ldots, f_n(z))$ and $f_i(z)$, $g(z)$ are polynomials. Therefore

$$r(z, \overline{w})\mu(z, \overline{w}) = r'\left(\frac{F(z)}{g(z)}, \frac{\overline{F(w)}}{\overline{g(w)}}\right) = \frac{P_0(z, \overline{w})}{Q(z)\overline{Q(w)}} \tag{11}$$

for some polynomials P_0, Q.

We claim $r(z, \overline{w})$ divides $P_0(z, \overline{w})$. For, we know that $\mu(z, \overline{w})$ has no singularities for z, w near $p_0 \in \partial D$. Therefore if $r(z, \overline{w}) = 0$ and z, w are near p_0, it follows that $P_0(z, \overline{w}) = 0$. In other words, P_0 vanishes in an open subset of $\{r = 0\}$. However, $\{r = 0\}$ is connected, and so by analytic continuation $r(z, \overline{w}) = 0$

implies $P_0(z, \bar{w}) = 0$. The Hilbert Nullstellensatz [43, p. 256] shows that r divides P_0^k for some k; since r is irreducible, r divides P_0 as claimed. (One can get away without using the Nullstellensatz—we leave this to the reader.)

Now since r divides P_0, we can rewrite (11) in the form

$$\mu(z, \bar{w}) = \frac{P(z, \bar{w})}{Q(z)\,\overline{Q(w)}}. \tag{12}$$

Similarly,

$$\mu'(z', \bar{w}') = \frac{P'(z', \bar{w}')}{Q'(z')\,\overline{Q'(w')}}. \tag{13}$$

Since $\mu(z, w)\mu'(z', w') = 1$, we get from (12), (13) that

$$1 = \frac{P(z, \bar{w})P'(z', \bar{w}')}{Q(z)\,\overline{Q(w)}\,Q'(z')\,\overline{Q'(w')}} = \frac{P(z, \bar{w})\tilde{P}(z, \bar{w})}{Q(z)\,\overline{Q(w)}\,Q'(z')\,\overline{Q'(w')}\,S(z)\,\overline{S(w)}},$$

for some polynomials $\tilde{P}$, S. Therefore $P(z, \bar{w})\tilde{P}(z, \bar{w}) = R(z)\overline{R(w)}$ for some rational function R. Clearing denominators gives

$$P(z, \bar{w})\tilde{\tilde{P}}(z, \bar{w}) = \tilde{S}(z)\,\overline{\tilde{S}(w)}, \quad \text{for polynomials } \tilde{\tilde{P}}, \tilde{S}.$$

Since the polynomials form a unqiue factorization domain, it follows that $P(z, \bar{w}) = A(z)\overline{B(w)}$. From (12) we obtain $\overline{P(z, \bar{w})} = P(w, \bar{z})$, which implies $A = B$. Substituting back into (12), we now have

$$\mu(z, \bar{w}) = \left(\frac{A(z)}{Q(z)}\right) \cdot \left(\frac{\overline{A(w)}}{\overline{Q(w)}}\right) = \beta(z)\,\overline{\beta(w)}. \quad \square$$

Now we can pass to linear combinations of the $r(\cdot, w_j)$.

LEMMA 3. *Φ carries the hypersurface $\{z \mid \Sigma_j \lambda_j r(z, \bar{w}_j) = 0\}$ to a hypersurface $\{z' \mid \Sigma_j \lambda'_j r(z', \bar{w}'_j) = 0\}$.*

PROOF. Consider the spaces

$$H = \text{linear span } \{r(\cdot, \bar{w}) \text{ as } w \text{ varies}\},$$
$$H' = \text{linear span } \{r'(\cdot, \bar{w}') \text{ as } w' \text{ varies}\}$$

These spaces consist of second-degree polynomials and are thus finite dimensional. We can pick generators of H, $r(\cdot, \bar{w}_1), \ldots, r(\cdot, \bar{w}_m)$ so that w_j is neither a zero nor a pole of $\beta(w)$ in Lemma 2. Set $w'_j = \Phi(w_j)$. By enlarging our list, we may assume that the $r'(\cdot, \bar{w}'_j)$ span H'. Now by Lemma 2,

$$\sum_{j=1}^{m} \lambda'_j r'(z', \bar{w}'_j) = 0 \quad \text{if and only if} \quad \sum_{j=1}^{m} \lambda'_j \beta(z)\,\overline{\beta(w_j)}\,r(z, \bar{w}_j) = 0,$$

i.e.

$$\beta(z) \cdot \sum_{j=1}^{m} \lambda_j r(z, \bar{w}_j) = 0, \quad \text{where } \lambda_j = \lambda'_j \overline{\beta(w_j)}. \tag{14}$$

Assume for the moment that $\beta(z) = 0$. This gives $\mu(z, \overline{w}_j) = \beta(z)\beta(w_j) = 0$ for $j = 1,\dots,m$, since $\beta(w_j)$ never has a pole; therefore

$$r'(z', \overline{w}_j') = \mu(z, \overline{w}_j)r(z, \overline{w}_j) = 0, \qquad j = 1,\dots,m.$$

But that is a contradiction, since the $r'(\cdot, \overline{w}_j')$ span H'. This means $\beta(z)$ is never zero when $z' = \Phi(z)$ is well defined, i.e. for $z \in \mathbf{C}^n \setminus V$. Therefore (14) implies

$$\sum_{j=1}^{m} \lambda'_j r'(z', \overline{w}'_j) = 0 \quad \text{if and only if } \sum_{j=1}^{m} \lambda_j r(z, \overline{w}_j) = 0.$$

which is the claim of the Lemma. □

Denote by T the isomorphism between H and H' used in the proof of Lemma 3, i.e.

$$(15) \qquad T\colon H' \to H, \qquad \sum_{j=1}^{m} \lambda'_j r'(z', \overline{w}'_j) \to \sum_{j=1}^{m} \left(\lambda'_j \overline{\beta(w_j)}\right) r(z, \overline{w}_j).$$

Sitting inside H' and H are the linear polynomials. We shall prove that T takes linear polynomials to linear polynomials. Lemma 3 then shows that Φ carries hyperplanes to hyperplanes. This is the next main step in the proof of Theorem 3. In order to give a useful characterization of the linear polynomials, we will introduce inner products so that T becomes unitary. This requires the map i: $\mathbf{C}^n \to H$ which sends w to $r(\cdot, \overline{w})$.

LEMMA 4. *There is a unique inner product $\langle \cdot, \cdot \rangle$ on H satisfying*

(i) $\langle\, , \rangle$ is nondegenerate,

(ii) $\mathrm{r}(z, \overline{w}) = \langle \mathrm{i}z, \mathrm{i}w \rangle$.

PROOF. Let $\{\psi_\alpha\}_{\alpha \leq m}$ be a basis of H. Then we can find $\overline{\chi_\alpha(w)}$ so that

$$r(z, \overline{w}) = \sum_{\alpha=1}^{m} \overline{\chi_\alpha(w)}\, \psi_\alpha(z).$$

Since $r(z, \overline{w}) = \overline{r(w, \overline{z})}$, this gives

$$r(w, \overline{z}) = \sum_{\alpha=1}^{m} \chi_\alpha(w)\, \overline{\psi_\alpha(z)},$$

or after changing notation

$$(17) \qquad r(z, \overline{w}) = \sum_{\alpha=1}^{m} \chi_\alpha(z)\, \overline{\psi_\alpha(w)}.$$

This means $H \subseteq$ linear span$\{\chi_\alpha(z), \alpha = 1,\dots,m\}$; since $\dim H = m$, $\{\chi_\alpha(z)\}_{\alpha \leq m}$ is a basis of H. The change of basis equations $\chi_\alpha = \sum_\beta g_{\alpha\beta}\psi_\beta$, where $(g_{\alpha\beta})$ is an invertible matrix, give us the formula

$$r(z, \overline{w}) = \sum_{\alpha,\beta=1}^{m} g_{\alpha\beta}\psi_\beta(z)\, \overline{\psi_\alpha(w)}.$$

Thus (16) holds if we define $\langle\ ,\ \rangle$ by

$$\left\langle \sum_{\alpha} \lambda_{\alpha} \chi_{\alpha}, \sum_{\beta} \tilde{\lambda}_{\beta} \chi_{\beta} \right\rangle = \sum_{\alpha,\beta} g_{\alpha\beta} \bar{\lambda}_{\alpha} \tilde{\lambda}_{\beta}. \tag{18}$$

The uniqueness of the inner product is immediate, since the image of $\imath$ spans H. $\square$

Lemma 4 also defines a natural inner product on H'. Let

$$h' = \sum_j \lambda'_j r'(\cdot, \bar{w}'_j) = \sum_j \lambda'_j \imath(w'_j), \qquad \tilde{h}' = \sum_k \tilde{\lambda}'_k r'(\cdot, \bar{w}'_k) = \sum_k \tilde{\lambda}'_k \imath(w'_k),$$

so that

$$\langle h', \tilde{h}' \rangle = \sum_{jk} \bar{\lambda}'_j \tilde{\lambda}'_k r'(w'_j, \bar{w}'_k).$$

Since

$$Th' = \sum_j \left(\lambda'_j \overline{\beta(w_j)}\right) r(\cdot, w_j) = \sum_j \left(\lambda'_j \overline{\beta(w_j)}\right) \imath(w_j),$$

$$T\tilde{h}' = \sum_k \left(\tilde{\lambda}'_k \overline{\beta(w_k)}\right) r(\cdot, \bar{w}_k) = \sum_k \left(\tilde{\lambda}'_k \overline{\beta(w_k)}\right) \imath(w_k),$$

we have

$$\begin{aligned}\langle Th', T\tilde{h}' \rangle &= \sum_{jk} \bar{\lambda}'_j \beta(w_j) \tilde{\lambda}'_k \overline{\beta(w_k)}\, r(w_j, \bar{w}_k) \\ &= \sum_{jk} \bar{\lambda}'_j \tilde{\lambda}'_k r'(w'_j, \bar{w}'_k) = \langle h', \tilde{h}' \rangle,\end{aligned}$$

showing that T is unitary with respect to $\langle\ ,\ \rangle$.

Next note that T carries $r(\cdot, \bar{w}) \in H$ to $\beta(w) r'(\cdot, w') \in H'$, whenever $w' = \Phi(w)$ is well defined. So if we set

$$\mathfrak{Q} = \text{closure of } \left\{\lambda r(\cdot, \bar{w}) \mid \lambda \in \mathbf{C}, w \in \mathbf{C}^n\right\} \subseteq H,$$

$$\mathfrak{Q}' = \text{closure of } \left\{\lambda' r(\cdot, w') \mid \lambda' \in \mathbf{C}, w' \in \mathbf{C}^n\right\} \subseteq H',$$

it follows that $T(\mathfrak{Q}) \subseteq \mathfrak{Q}'$. Thus, T preserves both an inner product and a distinguished subset $\mathfrak{Q}$.

Let us now specialize to the case of the ellipsoid of Proposition 3. We find that the general element $h \in H$ is given by $h = \chi_0 \psi_0 + \chi_* \psi_* + \sum_1^n \chi_j \psi_j$, where the χ's are arbitrary numbers, and $\psi_0 = 1$, $\psi_* = \sum_1^n A_j z_j^2$, $\psi_j = z_j$. The ψ's form a basis for H except for the ball (i.e. all $A_i = 0$) in which case ψ_* may evidently be deleted. Since we know that dim H is a biholomorphic invariant for ellipsoids, it follows that $\{|z|^2 + \sum_j 2A_j \mathrm{Re}(z_j^2) < 1\}$ cannot be biholomorphic to the ball unless all $A_j = 0$. From now on assume some $A_j \neq 0$.

Next we write down the inner product $\langle\ ,\ \rangle$ and the subset $\mathfrak{Q} \subseteq H$. Routine computations yield

$$\|h\|^2 = \sum_i^n |\chi_j|^2 - |\chi_0|^2 + 2\,\mathrm{Re}\left(\chi_0 \bar{\chi}_*\right),$$

$$h \in \mathfrak{Q} \quad \text{if and only if} \quad \sum_i^n A_j \chi_j^2 - \chi_*^2 - \chi_* \chi_0 = 0.$$

Now $T\colon H \to H'$ preserves $\|h\|^2$ and $\mathfrak{Q} = \{h \in H \mid (h, h) = 0\}$, where $(h, h) = \Sigma_1^n A_j \chi_j^2 - \chi_*^2 - \chi_* \chi_0$. A linear transformation preserving the zero set $\mathfrak{Q}$ of (h, h) must fix (h, h) up to a nonvanishing scalar factor β. In other words, T preserves $\langle\ ,\ \rangle$ exactly and $(\ ,\ )$ up to a factor β.

Since T preserves two quadratic forms, it is natural to look at the eigenvectors and eigenvalues of $(\ ,\ )$ with respect to $\langle\ ,\ \rangle$. That is, we seek $\omega_j \in H$, $\lambda_j \in \mathbf{C}$ such that

$$(19) \qquad (\xi, \omega_j) = \lambda_j \langle \xi, \omega_j \rangle.$$

This is slightly different from an ordinary eigenvalue problem, since $(\ ,\ )$ is bilinear and $\langle\ ,\ \rangle$ sesquilinear. Multiplying ω_j by $\rho e^{i\theta}$ has the effect of multiplying λ_j by $e^{2i\theta}$, so it is the $|\lambda_j|$ that have invariant meaning. Since T preserves the two inner product (up to factor β), it follows that the eigenvalues $\{\lambda_j\}$, $\{\lambda'_j\}$ corresponding to the two ellipsoids coincide up to a factor of $|\beta|$.

It is simple enough to compute the λ's explicitly. Substituting our formulas for $(\ ,\ )$ and $\langle\ ,\ \rangle$ into (19), we obtain the equations

$$(20) \qquad \lambda \bar{\chi}_j = A_j \chi_j, \qquad (1 \leqslant j \leqslant n),$$

$$(21) \qquad \lambda(-\bar{\chi}_0 + \bar{\chi}_*) = -\tfrac{1}{2}\chi_*,$$

$$(22) \qquad \lambda \bar{\chi}_0 = -\chi_* - \tfrac{1}{2}\chi_0.$$

If any $\chi_j \neq 0$, then (20) yields $|\lambda_j| = A_j$; and for each $|\lambda_j| = A_j$ there is an obvious solution w_j of (20)–(23) with a single nonzero component in w_j.

If all $\chi_j = 0$ $(1 \leqslant j \leqslant n)$, then we are left with the task of solving (21), (22). After elementary manipulation one finds two additional λ's, namely the roots $\lambda_\pm$ of $4\lambda^2 - 8\lambda + 1 = 0$. Thus our list of all $|\lambda_j|$'s is $\{A_1, A_2, \ldots, A_n, \lambda_+, \lambda_-\}$. This list therefore differs from $\{A'_1, A'_2, \ldots, A'_n, \lambda_+, \lambda_-\}$ only by a factor of $|\beta|$. Since the quadratic form defining an ellipsoid is positive, we have $0 \leqslant A_j < 1$, while $\lambda_- < 1 < \lambda_+$. Hence $\lambda_+ = \max\{A_1, A_2, \ldots, A_n, \lambda_+, \lambda_-\} = \max\{A'_1, \ldots, A'_n, \lambda_+, \lambda_-\}$, so $|\beta| = 1$ and $\{A_1, \ldots, A_n\} = \{A'_1, \ldots, A'_n\}$. The two ellipsoids D, D' are therefore the same, and Step 3 is complete. $\square$

We omit Step 4 (linearity of Φ). It is clear that the information from Step 3 places powerful constraints on the map T, therefore on Φ. For instance if $A_1, A_2, \ldots, A_n, \lambda_+, \lambda_-$ are distinct, then T must be diagonal.

Our discussion of Steps 1 and 2 follows the paper of Webster [**64**]. Step 3 follows an unpublished proof by Webster. The discussion of Steps 3 and 4 in [**64**] is actually shorter than here, but the ideas here obviously apply to much more general domains than the ellipsoids, so we thought them worth presenting.

4. Smoothness of biholomorphic maps. In this section we prove that a biholomorphic map of strictly pseudoconvex domains is smooth up to the boundary. The original proof in [**21**] was long and hard, but E. Ligocka and S. Bell [**2**] found a very simple proof using only the smoothness of solutions of the $\bar{\partial}$-Neumann problem. (See also Nirenberg-Webster-Yang **[51]** for a proof based on reflections.) Nowadays one can deal also with classes of weakly pseudoconvex domains. We are grateful to Bell for providing the following particularly simple argument in the strictly pseudoconvex case.

Proof of Theorem 1. Suppose $f: D_1 \to D_2$ is a biholomorphic mapping between smooth bounded strictly pseudoconvex domains contained in $\mathbf{C}^n$. Let P_1 and P_2 denote the Bergman orthogonal projections associated to D_1 and D_2, respectively, and let $u = \det[f']$. A simple Hilbert space argument using the fact that $|u|^2$ is equal to the *real* Jacobian determinant of f viewed as a mapping on $\mathbf{R}^{2n}$ reveals that the Bergman projections transform under f according to the formula $u((P_2\phi) \circ f) = P_1(u(\phi \circ f))$.

A key element in the proof of the theorem is the fact that if h is a holomorphic function on D_2 that is in $C^\infty(\bar{D}_2)$, then it is possible to construct a function ϕ in $C^\infty(\bar{D}_2)$ that vanishes to infinite order on ∂D_2 such that $P_2\phi = h$. Indeed, the C^∞ version of the Cauchy-Kowalewski theorem asserts that the Cauchy problem

$$\begin{aligned} \Delta\psi &= h && \text{on } D_2, \\ \psi = \nabla\psi &= 0 && \text{on } \partial D_2 \end{aligned}$$

can be solved modulo functions which vanish to infinite order on ∂D_2, i.e., there is a function ψ in $C^\infty(\bar{D}_2)$ such that $\psi = \nabla\psi = 0$ on ∂D_2 and such that $h - \Delta\psi$ vanishes to infinite order on ∂D_2. Now $\Delta\psi$ is orthogonal to holomorphic functions on D_2 via integration by parts. Hence, $\phi = h - \Delta\psi$ is a function in $C^\infty(\bar{D}_2)$ vanishing to infinite order on ∂D_2 such that $P_2\phi = P_2 h - P_2\Delta\psi = h$.

One more auxiliary fact is required before the theorem can be proved; namely, that if $\phi \in C^\infty(\bar{D}_2)$ vanishes to infinite order on ∂D_2, then $u(\phi \circ f)$ is in $C^\infty(\bar{D}_1)$. Since the components of f are bounded holomorphic functions on D_1, the classical Cauchy estimates yield that there are constants C_α independent of z such that $|\partial^\alpha f_k/\partial z^\alpha(z)| \leq C_\alpha \operatorname{dist}(z, \partial D_1)^{-|\alpha|}$ for all z in D_1. Hence, to see that all the derivatives of $u(\phi \circ f)$ are bounded functions on D_1, it suffices to prove that

$$\operatorname{dist}(f(z), \partial D_2) \leq C \operatorname{dist}(z, \partial D_1).$$

The proof of this classical estimate is due to Henkin and is quite simple. Let r be a defining function for D_1 that is strictly plurisubharmonic in a neighborhood of $\bar{D}_1$. Then $r \circ f^{-1}$ is a negative subharmonic function on D_2. Since $r \circ f^{-1}$ attains its maximum value, zero, on ∂D_2, the classical Hopf lemma implies that $(r \circ f^{-1})(z) \leq -C \operatorname{dist}(z, \partial D_2)$. The desired inequality now follows because $-r(w) \leq C' \operatorname{dist}(w, \partial D_1)$.

The proof of the theorem can now be completed. Let h be a holomorphic function on D_2 that is in $C^\infty(\bar{D}_2)$. Let ϕ be a function in $C^\infty(\bar{D}_2)$ vanishing to infinite order on ∂D_2 such that $P_2\phi = h$. The transformation formula for the Bergman projections yields that $u(h \circ f) = u((P_2\phi) \circ f) = P_1(u(\phi \circ f))$. Kohn's formula, $P_1 = I - \bar{\partial}^* N \bar{\partial}$, which relates the Bergman projection P_1 to the $\bar{\partial}$-Neumann operator N, together with Kohn's subelliptic estimates for N, reveal that P_1 preserves the space $C^\infty(\bar{D}_1)$. Hence, since $u(\phi \circ f)$ is in $C^\infty(\bar{D}_1)$, the relation $u(h \circ f) = P_1(u(\phi \circ f))$ implies that $u(h \circ f)$ is in $C^\infty(\bar{D}_1)$. Let $h = 1$ to obtain that u is in $C^\infty(\bar{D}_1)$. Let $h = z_k$, the kth coordinate function of $\mathbf{C}^n$, to obtain that uf_k is in $C^\infty(\bar{D}_1)$. Hence f extends smoothly to ∂D_1 where u does not vanish. But u cannot vanish on ∂D_1. To see this, observe that the same arguments can be applied to the inverse mapping $F = f^{-1}$. Hence, $U = \det[F']$ is in $C^\infty(\bar{D}_2)$. Now, the relation

$$U(z) = 1/u(F(z))$$

implies that u is bounded away from zero. This completes the proof of the theorem. □

CHAPTER 9. MOSER'S NORMAL FORM

Suppose we are given a strictly ψ-convex domain in $\mathbf{C}^{n+1} = \{(z_1, \ldots, z_n, w)\}$, where $w = u + iv$. We know that the boundary has a defining function satisfying the Levi condition. In the case we consider first, this defining function is given by a formal power series, which we can assume is centered at the origin. The goal is to find coordinates so that the boundary near the origin is in Moser's normal form

$$(*) \qquad v = |z|^2 + \sum_{|\alpha|,|\beta| \geqslant 2} \sum_{l=0}^{\infty} A^l_{\alpha\bar{\beta}} u^l z^\alpha \bar{z}^\beta,$$

where the coefficients $A^l_{\alpha\bar{\beta}} = A^l_{j_1,\ldots,j_r\bar{k}_1,\ldots,\bar{k}_s}$ satisfy

(a) $\sum_{p=1}^n A^l_{pj_2\bar{p}\bar{k}_2} = 0$, for all $l, j_2, \bar{k}_2$.

(b) $\sum_{p,q=1}^n A^l_{pqj_3\bar{p}\bar{q}} = 0$, for all l, j_3.

(c) $\sum_{p,q,r=1}^n A^l_{pqr\bar{p}\bar{q}\bar{r}} = 0$, for all l,

and the indices are tensor indices varying from 1 to n. Requirement (a) says that $\operatorname{Trace} A_{22} = 0$; (b) that $\operatorname{Trace}^3 A_{23} = \operatorname{Trace}^2 A_{32} = 0$ and (c) that $\operatorname{Trace}^3 A_{33} = 0$. Of course the coordinate change will also be given by formal power series. In the last section of the chapter we will consider the question of when these series actually converge.

1. Moser's theorem for formal power series. We are interested not only in bringing a boundary to normal form but also in the maps that preserve this form. Define a *biholomorphism of normal forms* to be a biholomorphic map that takes the surface $(*)$ in normal form to any other hypersurface in normal form and preserves the origin.

THEOREM 1 (MOSER). (1) *Let ∂D be a strictly pseudoconvex real analytic boundary and let p be a point in ∂D. There is a biholomorphic map defined in a neighborhood of p which places ∂D in Moser's normal form and takes p to the origin.*

(2) *Suppose ∂D is already in normal form, and $p = 0$. Then the class of all biholomorphic maps which fix 0 and carry ∂D to a boundary in normal form is*

parametrized by $\mathcal{H}$, the group of linear fractional transformations of the hyperquadric which fix the origin.

PROOF. After translating and rotating, we can assume that the power series is centered at 0 and that the tangent space at 0 is $v = 0$. This means that for a real-valued F our surface is given by

$$
\begin{aligned}
&v = F(z, \bar{z}, u), \\
(1) \qquad &F(0,0,0) = 0, \qquad \frac{\partial F}{\partial z_k} = \frac{\partial F}{\partial \bar{z}_k} = \frac{\partial F}{\partial u} = 0 \quad \text{at } (0,0,0).
\end{aligned}
$$

The *weights* of the terms in the formal power series are determined by the table

term	weight
$z, \bar{z}$	1
u, v	2

Equation (1) means that F has no terms of weight 1 and 3 terms of weight 2,

$$\langle z, \bar{z} \rangle, \qquad Q(z), \qquad \bar{Q}(\bar{z}),$$

where $\langle\, , \rangle$ is a nondegenerate quadratic form and Q is a quadratic form. The form $\langle\, , \rangle$ is nondegenerate because the boundary was assumed to be strictly ψ-convex. Equation (1) becomes

$$v = \langle z, \bar{z} \rangle + Q(z) + \bar{Q}(\bar{z}) + (\text{terms of weight} \geqslant 3).$$

After the change of variables

$$(z, w) \to (z^*, w^*), \qquad z^* = z, \qquad w^* = w - 2iQ(z),$$

the variable v^* becomes

$$v^* = \operatorname{Im} w^* = \operatorname{Im} w - 2 \operatorname{Re} Q(z) = v - Q(z) - \bar{Q}(\bar{z})$$

so that in the new coordinates

$$(2) \qquad v = \langle z, \bar{z} \rangle + (\text{terms of weight} \geqslant 3).$$

In order to study the surface (2), we introduce the group $\mathcal{G}_1$ consisting of its biholomorphic transformations; in other words $\mathcal{G}_1$ consists of all transformations given by formal power series that take surfaces of the form (2) to surfaces of the form (2).

LEMMA 1. *Any $\Phi \in \mathcal{G}_1$ can be uniquely factored as $\Phi = \Psi \circ \phi$, where ϕ is a fractional linear transformation in $\mathcal{H}$ and Ψ agrees with the identity in the initial part of its Taylor expansion; in other words, at the origin*

$$\frac{\partial w^*}{\partial z_k} = 0, \qquad \frac{\partial w^*}{\partial w} = 1, \qquad \frac{\partial^2 w^*}{\partial z_j \partial z_k} = 0, \qquad \frac{\partial z_k^*}{\partial z_j} = \delta_{jk}, \qquad \operatorname{Re}\left(\frac{\partial^2 w^*}{\partial w^2}\right) = 0$$

where $\Psi(z, w) = (z^*, w^*)$.

PROOF OF LEMMA 1. Recall that the group

$$\mathcal{H} = \text{fractional linear transformations of } H \text{ which fix the origin}$$

consists of compositions of

(i) Heisenberg dilations,

(ii) unitary rotations in z,

(iii) translations at ∞ parametrized by $(v_1,\dots,v_n,\zeta) \in \mathbf{C}^n \times \mathbf{R}$.

The translations (iii) are given by

$$(z_k,w) \mapsto (z_k^*,w^*) = \frac{(z_k+\bar{v}_k w/2i,w)}{1+\Sigma_k v_k z_k + w\left(\zeta-(i/4)\Sigma_k|v_k|^2\right)},$$

where (z_k,w) is an abbreviation for $(z_1,\dots,z_k,\dots,z_n,w)$.

We are given $\Phi \in \mathcal{G}_1$ mapping

$$\Phi\colon (z_k,w) \to (z_k^*,w^*)$$

which satisfies

$$z_k^*=0, \qquad w^*=0, \qquad \partial w^*/\partial w=\lambda>0, \qquad \partial w^*/\partial z_k=0$$

at the origin. Since λ is positive and real, we can find a dilation such that

$$\Phi = (\text{better } \Phi)\circ(\text{Heisenberg dilation}),$$

where (better Φ): $(z_k,w)\mapsto(z_k^*,w^*)$ satisfies $\partial w^*/\partial w=1$. This means

$$\begin{aligned} &w^* = w+\textstyle\sum a_{jk}z_jz_k + (\text{terms of weight} \geqslant 3),\\ (3)\qquad &z_k^* = \sum_j c_{kj}z_j+\beta_k w + (\text{terms quadratic in } z) + (\text{terms of weight} \geqslant 3). \end{aligned}$$

From the definition of v^* and equation (2), we get

$$v^* = \operatorname{Im} w^*, \qquad v^* = |z^*|^2 + (\text{terms of weight} \geqslant 3);$$

comparing the terms of weight 2 in these equations to those in equation (3) shows $(\partial^2 w^*/\partial z_j \partial z_k)\,|_0 = 0$, so that $a_{jk}=0$. Similarly, (c_{kj}) must be a unitary matrix; therefore, for a suitable rotation

$$(\text{better } \Phi) = (\text{still better } \Phi)\circ(\text{unitary rotation}),$$

where (still better Φ): $(z_k,w)\mapsto(z_k^*,w^*)$ and satisfies $\partial z_j^*/\partial z_k=\delta_{jk}$. Equation (3) now becomes

$$\begin{aligned} &w^* = w + (\text{terms of weight} \geqslant 3),\\ &z_k^* = z_k+\beta_k w + (\text{terms quadratic in } z) + (\text{terms of weight} \geqslant 3). \end{aligned}$$

The first term of the Taylor series of the translations (iii), with $\zeta=0$, is

$$(z_k,w)\mapsto(z_k+\bar{v}_k w/2i,w).$$

Choosing the parameters $(v_1,\dots,v_n,0)$ of the translation so that $\bar{v}_k/2i=\beta_k$ allows us to factor (still better ϕ) as

$$(\text{still better } \Phi) = (\text{much better } \Phi)\circ\left(\text{translation } (v_k,0) \text{ at } \infty\right),$$

where (much better Φ): $(z_k,w)\mapsto(z_k^*,w^*)$ satisfies

$$\begin{aligned} &w^* = w + (\text{terms of weight} \geqslant 3),\\ &v^* = z_k + (\text{terms quadratic in } z) + (\text{terms of weight} \geqslant 3). \end{aligned}$$

Finally we need to consider translations parametrized by $(0, \zeta)$; these take the form

$$(z_k, w) \mapsto \left(\frac{z_k}{1+\zeta w}, \frac{w}{1+\zeta w} \right).$$

For appropriate ζ, the second component can be expanded as $w/(1+\zeta w) = w - \zeta w^2 + \cdots$, which shows that we can find a factorization

$$(\text{much better } \Phi) = \Psi \circ (\text{translation } (0, \zeta) \text{ at } \infty)$$

so that the map $\Psi\colon (z_k, w) \mapsto (z_k^*, w^*)$ satisfies $\operatorname{Re}(\partial^2 w^*/\partial w^2) = 0$. Reviewing the choices made shows that we have constructed a unique Ψ satisfying the conditions of the lemma. □

Let $\mathcal{G}_0$ be the subgroup of $\mathcal{G}_1$ consisting of all holomorphic changes of coordinates about 0

$$\Psi\colon (z, w) \mapsto (z^*, w^*)$$

satisfying

(i) through first order, $\Psi = \mathrm{I}$,
(ii) $(\partial^2 w^*/\partial z_j \partial z_k)\,|_0 = 0$,
(iii) $\operatorname{Re}(\partial^2 w^*/\partial w^2)\,|_0 = 0$.

The lemma shows that $\mathcal{G}_1 = \mathcal{G}_0 \circ \mathcal{H}$, uniquely.

In the next section we will prove

THEOREM 2 (MOSER). *If ∂D is given by*

$$v = |z|^2 + (\textit{terms of weight} \geqslant 3),$$

then there is precisely one transformation in $\mathcal{G}_0$ which carries ∂D to a hypersurface in normal form.

COROLLARY. *The only biholomorphic transformations (given by formal power series) taking the hyperquadric H to itself are the fractional linear transformations.*

Clearly now Theorem 1 is reduced to the more precise Theorem 2.

2. Proof of Moser's theorem (Theorem 2). Recall that the boundary ∂D is given by

$$v = |z|^2 + F(z, \bar{z}, u), \tag{4}$$

where F is a formal power series of weight $\geqslant 3$, $(z, w) \in \mathbf{C}^{n+1}$ and $w = u + iv$. We will find a tranformation

$$z_k^* = z_k + f_k(z, w), \qquad w^* = w + g(z, w)$$

and a formal power series F^* of weight $\geqslant 3$ satisfying

(i) at the origin,

$$\frac{\partial f_k}{\partial z_l} = \frac{\partial f_k}{\partial w} = \frac{\partial g}{\partial z_l} = \frac{\partial g}{\partial w} = \frac{\partial^2 g}{\partial z_k \partial z_l} = \operatorname{Re}\left(\frac{\partial^2 g}{\partial w^2} \right) = 0;$$

(ii)

$$(5)\quad |z|^2 + 2\,\mathrm{Re}\langle f, z\rangle + \langle f, f\rangle + F^*\left(z + f, \bar{z} + \bar{f}, u + \tfrac{1}{2}(g + \bar{g})\right)$$
$$= |z|^2 + F(z, \bar{z}, u) + \frac{1}{2i}(g - \bar{g}),$$

where f, g are evaluated at (z, w) and $w = u + i|z|^2 + iF(z, \bar{z}, u)$;

(iii) the formal power series F^* satisfies the trace conditions (a), (b), (c). Condition (i) implies that the transformation $(z_k, w) \mapsto (z_k^*, w^*)$ is an element of $\mathcal{G}_0$; conditions (ii) and (iii) imply that (4) is equivalent to

$$v^* = |z|^2 + 2\,\mathrm{Re}\langle f, z\rangle + \langle f, f\rangle + F^*\left(z + f, \bar{z} + \bar{f}, u + \tfrac{1}{2}(g + \bar{g})\right)$$

or

$$v^* = |z^*|^2 + F^*(z^*, \bar{z}^*, u^*),$$

where F^* satisfies the Moser trace conditions (a), (b), (c). We will prove Theorem 2 by showing that given any F, there exists precisely one F^*, and functions f_k, g, such that conditions (i), (ii) and (iii) hold. Let

$$\mathrm{I} = \left\{\begin{array}{lll} \text{terms in (5) of weight:} & \nu & \text{in } F, F^* \\ & \nu & \text{in } g \\ & \nu - 1 & \text{in } f \end{array}\right\},$$

$$\mathrm{II} = \left\{\begin{array}{lll} \text{terms in (5) of weight:} & < \nu & \text{in } F, F^* \\ & < \nu & \text{in } g \\ & < \nu - 1 & \text{in } f \end{array}\right\}.$$

Lemma 2. *The terms of weight ν in* (5) *depend linearly on elements of* I, *nonlinearly on elements of* II, *and not at all on any other terms in f_k and g.*

Proof of Lemma 2. We begin with three observations about the Taylor series expansions of g and f_k. Consider the Taylor series

$$g(z, w)\Big|_{w = u + i|z|^2 + iF(z,\bar{z},u)} = \sum_\alpha \frac{1}{\alpha!}\left(\partial_w^\alpha g\,\big|_{u+i|z|^2}\right)\left(iF(z,\bar{z},u)\right)^\alpha$$
$$= \sum_\alpha \frac{1}{\alpha!}\left(\text{wt. } a_\alpha + [\text{wt.}(a_\alpha) + 1] + \cdots\right)(\text{wt.}\,3 + \text{wt.}\,4 + \cdots)^\alpha$$

and note

(i) a term of weight ν comes from the product of a term of weight ν_0 in $(\partial_w^\alpha g\,|_{u+i|z|^2})$ with terms of weight $\nu_1, \ldots, \nu_{|\alpha|}$ in $(iF(z,\bar{z},u))^\alpha$ satisfying $\nu_0 + \nu_1 + \cdots + \nu_{|\alpha|} = \nu$.

(ii) The first term of the series arises from a term of weight $(\nu_0 + 2|\alpha|)$ in g; furthermore, $\nu_1 \geqslant 3, \ldots, \nu_{|\alpha|} \geqslant 3$.

(iii) If $|\alpha| > 0$, then $(\nu_0 + 2|\alpha|), \nu_1, \ldots, \nu_{|\alpha|}$ are all $< \nu$. Observations (i) and (ii) are obvious; to verify (iii), consider

$$\nu = \nu_0 + \nu_1 + \cdots + \nu_{|\alpha|} \geqslant \nu_0 + 3|\alpha| > \nu_0 + 2|\alpha|.$$

Exactly the same observations apply to the Taylor series expansions of f_k. Applying these observations to the Taylor series

$$F^*\left(z+f, \bar{z}+\bar{f}, u+\tfrac{1}{2}(g+\bar{g})\right)$$
$$= \sum_{\alpha,\beta,k} \frac{1}{\alpha!\beta!k!} \partial_z^\alpha \partial_{\bar{z}}^\beta \partial_u^k F^* \big|_{(z,\bar{z},u)} \left(f^\alpha \bar{f}^\beta \left(\frac{g+\bar{g}}{2} \right)^k \right),$$

shows that

$$F^*\left(z+f, \bar{z}+\bar{f}, u+\tfrac{1}{2}(g+\bar{g})\right) = \{\text{terms of wt. } \nu \text{ in } F^*(z,\bar{z},u)\}$$
$$+ \left\{ \text{expression in terms of wt.} \begin{array}{ll} < \nu & \text{in } F^* \\ < \nu - 1 & \text{in } f \\ < \nu & \text{in } g \end{array} \right\}.$$

Similarly, if we let F_ν, F_ν^*, g_ν, f_ν denote terms of weight ν in the functions F, F^*, g, f, respectively, then (5) gives

$$F_\nu + \operatorname{Im} g_\nu \big|_{w=u+i|z|^2} + \{\text{expression in } F_\mu, g_\mu \text{ for } \mu < \nu\}$$
$$= 2\operatorname{Re}\langle f_{\nu-1}, z\rangle \big|_{w=u+i|z|^2} + \left\{ \text{expression in} \begin{array}{l} f_\mu \text{ for } \mu < \nu - 1 \\ \\ F_\mu \text{ for } \mu < \nu \end{array} \right\}$$
$$+ F_\nu^*(z,\bar{z},u) + \{\text{expression in } F_\mu^*, g_\mu, F_\mu \text{ for } \mu < \nu\}.$$

and this proves the lemma. □

Given all the F_ν, we can prove the theorem by using induction on the weight ν to solve for F_ν^*, $f_{\nu-1}$, g_ν. Assume that we have already determined the functions for $0, 1, \ldots, \nu-1$; then by (5)

$$F_\nu + \operatorname{Im} g_\nu \big|_{w=u+i|z|^2} = 2\operatorname{Re}\langle f_{\nu-1}, z\rangle \big|_{w=u+i|z|^2} + F_\nu^*$$
$$+ \{\text{already determined function}\}.$$

We will solve this equation by using the linear operator

$$L^\nu \colon (f_{\nu-1}, g_\nu) \mapsto \left(\operatorname{Im} g_\nu - 2\operatorname{Re}\langle f_{\nu-1}, z\rangle\right) \big|_{w=u+i|z|^2}.$$

Let $\mathfrak{N}_\nu$ denote a complement to the range of L^ν. Since $F_\nu - \{\text{already determined function}\}$ is given and since one can show that L^ν is injective, the equation

$$F_\nu - \{\text{already determined function}\} = F_\nu^* - L^\nu(f_{\nu-1}, g_\nu)$$

uniquely determines $F_\nu^* \in \mathfrak{N}_\nu$ and $f_{\nu-1}$, g_ν. Taking the direct sum over ν gives a linear operator acting on formal power series

$$L \colon (f, g) \to \left(\operatorname{Im} g - 2\operatorname{Re}\langle f, z\rangle\right) \big|_{w=u+i|z|^2},$$

with the property that, if $\mathfrak{N} = \oplus_\nu \mathfrak{N}_\nu$, then $\mathfrak{N}$ is a complement to the range of L. If we follow the prescription above, we get unique formal power series f, g and F^* with $F^* \in \mathfrak{N}$ and satisfying equation (5). This leaves conditions (i) and (iii): condition (i) was incorporated in the weight analysis at the beginning of the proof, and an examination of the linear algebra shows that $\mathfrak{N}$ may be taken to be precisely the set of F^* which satisfy trace condition (iii). This finishes our sketch of the proof of Moser's theorem. The reader should note that we did not check the two crucial points of linear algebra: L^v is injective, and a complement for its range is the space of F^* satisfying the trace conditions. These verifications are the hard parts of Moser's discussion of formal power series.

3. The convergence of the Moser normal form. We will prove in this section that for a surface given by a convergent power series the formal argument of the last section actually gives a convergent change of coordinates into Moser normal form. Note that a surface in normal form always contains the line

$$L = \{z = 0, w \text{ real}\}.$$

The idea of the proof is to expand the power series not around a point but around a curve; and then to find a transformation that maps this curve into the line L.

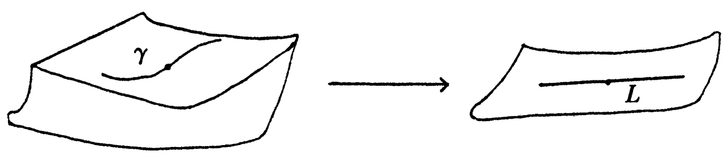

FIGURE 3

Let $t \mapsto \gamma(t)$ denote a real analytic curve in the given surface and assume that $\gamma'(0)$ is transverse to the "holomorphic" part of the tangent space $T(\partial D)$. (That is, $\gamma'(0)$ is tangent to ∂D but does not remain so after multiplication by $\sqrt{-1}$.) Let $\{e_\alpha\}$ be vectors of type $(1,0)$ varying analytically along γ and spanning the holomorphic part of the tangent space $T(\partial D)$ at each point of γ.

Step 1. As a first step, we place the boundary ∂D into a partial normal form

$$v = \langle z, z\rangle_u + F(z, \bar{z}, u),$$

where $\langle\,,\rangle_u$ is a Hermitian form depending upon u and F is a convergent power series, all of whose terms contain at least two z_j's and at least two $\bar{z}_j$'s. Recall that $(z, w) \in \mathbf{C}^n \times \mathbf{C}$ and $w = u + iv$. We use the term "partial normal form" to indicate the fact that although the boundary satisfies many of the conditions defining a normal form, no requirements are made yet about the trace conditions. We also demand that the change of coordinates be biholomorphic, that the curve $\gamma = \{t \mapsto \gamma(t)\}$ be sent to the line $L = \{t \mapsto (z = 0, w = t)\}$, and that the frame $\{e_\alpha\}$ be sent to the frame $\{\partial/\partial z_\alpha\}$. Given γ, $\{e_\alpha\}$, we claim that there is precisely one coordinate change meeting these demands.

Step 2. The purpose of this step is to replace the Hermitian form $\langle\,,\rangle_u$ with one that is independent of u. In fact, we claim that given a curve $\gamma(t)$ and a frame $\{e_\alpha\}$ satisfying

$$\langle e_\alpha, e_\beta\rangle_{\text{Levi form}} = \mu(t)\delta_{\alpha\beta},$$

we can find a unique transformation described by a convergent process taking the boundary into partial normal form

$$v = z \cdot \bar{z} + F(z, \bar{z}, u),$$

the curve γ into the line L, and the frame $\{e_\alpha\}$ into the frame $\{\zeta\partial/\partial z_\alpha\}$, for a suitable scalar factor ζ.

Step 3. In this step we show that by suitably choosing the parametrized curve γ and the frame $\{e_\alpha\}$ we can satisfy the trace conditions. In fact the vanishing of $\mathrm{Trace}^2\, F_{32}$ is a second order system of ODE's for the unparametrized curve γ. If we pick a γ that is a solution of this system, then the requirement that $\mathrm{Trace}\, F_{22}$ vanish becomes the condition that a framing of this curve satisfy a certain system of ODE's. Finally, suppose that we chose such a frame. The remaining requirement is that $\mathrm{Trace}^2\, F_{33}$ vanish: this becomes a single 3rd-order ODE involving the parametrizations, which can also be satisfied. This completes the sketch of the proof of the convergence of the Moser normal form; we will now fill in a few details.

Proof of Step 1. Say $\gamma(0)$ is the origin and the holomorphic part of $T_{\partial D}|_0$ is $\{(z, w) \in \mathbf{C}^n \times \mathbf{C} \mid w = 0\}$. Assume that the curve γ is given by $t \mapsto (p(t), q(t)) \in \mathbf{C}^n \times \mathbf{C}$, where p and q are analytic functions and satisfy $(p(0), q(0)) = 0$. The transversality condition becomes $dq/dt \neq 0$. Suppose that we extend p, q, to complex values of t and use a local change of variables φ such that

$$(z^*, w^*) \mapsto (z, w) = (z^* + p(w^*), q(w^*)).$$

Then φ is biholomorphic and maps the line L to the curve γ via

$$\{t \mapsto (z^* = 0, w^* = t)\} \underset{\varphi}{\rightarrow} \{t \mapsto (z = p(t), q(t) = w)\};$$

in other words, the transformation φ^{-1} straightens out the curve γ, allowing us to assume from the start that γ is given by

$$t \mapsto (0, \mathrm{Re}\, w = t, \mathrm{Im}\, w = 0).$$

So far, since the straight line γ lies on ∂D, our surface is given by a power series G satisfying

$$v = G(z, \bar{z}, u), \qquad G(0, 0, u) \equiv 0.$$

We claim that we can find a transformation

$$w^* = w + g(z, w), \qquad g(0, w) \equiv 0, \qquad z^* = z,$$

given by a convergent power series g, that transforms the boundary into

$$v^* = G^*(z^*, \bar{z}^*, u^*),$$

where G^* contains no pure powers $z^{*\alpha}$ or $\bar{z}^{*\alpha}$; in other words, $G^*(z^*, 0, u^*) \equiv 0$ and $G^*(0, \bar{z}^*, u^*) \equiv 0$. Note that the condition $g(0, w) \equiv 0$ means that the line γ is preserved. To verify this claim note that

$$v^* = v + \frac{1}{2i}(g(z, w) - \bar{g}(z, w)) = G^*\Big(z, \bar{z}, u + \frac{1}{2}(g(z, w) + \bar{g}(z, w))\Big),$$

whenever $w = u + iv = u + iG(z, \bar{z}, u)$, and this implies

$$G(z, \bar{z}, u) + \frac{1}{2i}(g(z, w) - \bar{g}(z, w)) = G^*\Big(z, \bar{z}, u + \frac{1}{2}(g(z, w) + \bar{g}(z, w))\Big).$$

Treating z and $\bar{z}$ as independent and putting $\bar{z} = 0$ gives

$$G(z, 0, u) + \frac{1}{2i}g(z, u + iG(z, 0, u)) = G^*\Big(z, 0, u + \frac{1}{2}g(z, u + iG(z, 0, u))\Big),$$

since $g(0, w) \equiv 0 \Rightarrow g = \sum_{|\alpha| \geq 1} c_{\alpha j} z^\alpha w^j \Rightarrow g(\bar{z}, \bar{w})\,|_{\bar{z}=0} = 0$. Consider the equation

$$G(z, 0, u) + \frac{1}{2i}g(z, u + iG(z, 0, u)) \equiv 0. \tag{6}$$

This implies $G^*(z, 0, u + \frac{1}{2}\, g(z, u + iG(z, 0, u))) \equiv 0$, which after the real analytic transformation

$$(z, u) \mapsto (z, u + \tfrac{1}{2}g(z, u + iG(z, 0, u))),$$

becomes $G^*(z^*, 0, u^*) \equiv 0$, which is what we want to prove; in other words, if we can satisfy (6), we are done. Note that

$$(z, u) \mapsto (z, u + iG(z, 0, u))$$

is also a real analytic transformation; this reduces (6) to

$$\tilde{G}(\tilde{z}, \tilde{u}) + \frac{1}{2i}g(\tilde{z}, \tilde{u}) \equiv 0,$$

which is solved by putting $g = -2i\tilde{G}$. Retracing our steps shows that we have determined a power series $G^*(z^*, \bar{z}^*, u^*)$ which contains no pure powers z^{*^α} or $\bar{z}^{*^\alpha}$.

This means that we can assume that

$$v = \langle z, z \rangle_u + F(z, \bar{z}, u), \qquad F(z, 0, u) \equiv 0, \tag{7}$$

where F contains no terms in z, $\bar{z}$ of type (1, 1). Our second claim is that we can find a biholomorphic change of coordinates

$$z^* = z + f(z, w), \qquad w^* = w,$$

with $f(0, w) \equiv 0$ (to preserve the line γ) and $(\partial f/\partial z_\alpha)(0, w) = 0$, so that (7) is transformed into a surface of the same form, but this time with an F^* involving only $z^\alpha \bar{z}^\beta$, where $|\,\alpha\,|$, $|\,\beta\,| \geq 2$. To verify this claim consider

$$\begin{aligned} v &= \langle z^*, z^* \rangle_u + F^*(z^*, \bar{z}^*, u) \\ &= \langle z, z \rangle_u + 2\operatorname{Re}\langle f, z \rangle_u + \langle f, f \rangle_u + F^*(z^*, \bar{z}^*, u). \end{aligned}$$

Comparing this to (7) gives

$$F^*(z^*, \bar{z}^*, u) = F - 2\operatorname{Re}\langle f, z \rangle_u - \langle f, f \rangle_u.$$

Note that the last term on the right contains the appropriate powers as long as $f(0, w) = (\partial f/\partial z_\alpha)(0, w) = 0$. Given F, we will choose f so as to kill the remaining terms of type $(m, 1)$, for $m > 1$; that is so that $(\partial F^*/\partial \bar{z}_\beta)(z, 0, u) \equiv 0$. This will happen if $(\partial/\partial \bar{z}_\beta)F^*(z^*, 0, u) \equiv 0$, i.e. $\sum_\alpha g_{\alpha\beta}(u) f_\alpha(z, w) + \partial F/\partial \bar{z}_\beta = 0$, and if $(g_{\alpha\beta}(u)) > 0$ (which holds since ∂D is strictly pseudoconvex). So we can find the required f by inverting the matrix $(g_{\alpha\beta})$, and this proves the claim.

We now consider the condition on the frame. To review—the boundary is given by (7), with all terms in F containing at least two z's and two $\bar{z}$'s, and γ is the curve $\{z = 0, w = t\}$. Introduce the coordinate transformation

$$z^* = c(w)z, \qquad c(w) \text{ a matrix,}$$
$$w^* = w;$$

this acts on the frame by

$$e_\alpha \mapsto c(t)e_\alpha\,\big|_{\gamma(t)}.$$

It is easy to find matrices $c(w)$ such that

$$c(t)e_\alpha\,\big|_{\gamma(t)} = \frac{\partial}{\partial z_\alpha},$$

and it is not much harder to show that these matrices are unique. This concludes Step 1.

Proof of Step 2. Writing (7) as

$$v = \theta(u)|z|^2 + F,$$

suggests the coordinate transformation

$$z^* = c(w)z, \qquad c(w) \text{ a number,}$$
$$w^* = w,$$

giving, for appropriate $c(w)$, $v = |z|^2 + F$. The desired result now follows immediately.

Proof of Step 3. Given a boundary ∂D, a parametrized curve $\gamma\colon t \to (p(t), q(t)) \in \partial D$ and a frame $(e_\alpha(t))$ defined along γ, we have constructed a biholomorphic map $\Phi\colon\ \partial D \to \partial\tilde{D}$, where $\partial\tilde{D} = \{v = |z|^2 + \tilde{F}(z, \bar{z}, u)\}$ and $\tilde{F} = \sum_{|\alpha|,|\beta| \geq 2} \tilde{F}_{\alpha\bar{\beta}}(u)z^\alpha\bar{z}^\beta$. By following the construction carefully, we can in principle read off the Taylor series of $\tilde{F}$ from those of ∂D, γ, (e_α). In particular, the trace conditions $\operatorname{Trace}\tilde{F}_{22}(u) = 0$, $\operatorname{Trace}^2\tilde{F}_{32}(u) = 0$, $\operatorname{Trace}^3\tilde{F}_{33}(u) = 0$ are really conditions on the Taylor series of ∂D, $\gamma(t)$, $e_\alpha(t)$ about the points on the curve γ. Regarding ∂D as fixed, and $\gamma(t)$, $e_\alpha(t)$ as arbitrary, we find after much work that

(a) $\operatorname{Trace}^2\tilde{F}_{32}(u) = 0$ means that the unparametrized curve γ satisfies a second-order ordinary differential equation. The standard existence theorem for ODE's therefore shows that we can pick the unparametrized curve to satisfy the trace condition on $\tilde{F}_{32}$. In particular, the initial direction of γ (transverse to the complex part of $T(\partial D)$) may be prescribed arbitrarily, and the frame $e_\alpha(t)$ and parametrization of γ have no effect on the trace condition for $\tilde{F}_{32}$. γ is called a "chain".

(b) Assume the unparametrized curve has been picked to satisfy $\operatorname{Trace}^2\tilde{F}_{32}(u) = 0$. Then $\operatorname{Trace}\tilde{F}_{22}(u) = 0$ if and only if the frame $e_\alpha(t)$ satisfies a first-order ordinary differential equation along γ. Again applying the standard existence theorem for ODE's we conclude that the frame $e_\alpha(t)$ may be picked to satisfy the trace condition on $\tilde{F}_{22}$. We are free to specify the frame at $t = 0$; the propagation of the frame along the curve is then uniquely determined.

(c) Finally, suppose the unparametrized curve γ and the frame (e_α) have been picked as above. Then the last trace condition $\operatorname{Trace}^3 \tilde{F}_{33}(u) = 0$ becomes a third-order ordinary differential equation for the parametrization of γ. In fact, the equation for the parameter τ involves a Schwartzian derivative, so that τ is uniquely determined up to linear fractional transformations.

Thus, we may transform ∂D to Moser's normal form by convergent power series, by picking first the unparametrized curve γ, then the frame e_α, and finally the parametrization of γ. The choices can all be made properly by solving ordinary differential equations. We may pick arbitrarily an initial direction for γ, an initial frame e_α, and a linear fractional transformation of the line. The number of free parameters agrees exactly with the dimension of the isotropy group $\mathcal{H}$, so all the formal power series transformations to normal form arise from suitable γ, e_α, τ and thus converge. This completes Moser's convergence proof. It is remarkable that one can get away without using anything more than existence of solutions of ODE's.

Note that the convergence proof made repeated use of the assumption that the boundary ∂D was real-analytic. In fact, various series, e.g. $(p(t), q(t))$, are defined initially for real t but are then used for complex t. One can easily give a reasonable definition of Moser's normal form for C^∞ boundaries, but it is also easy to write down smooth strictly pseudoconvex domains which cannot be placed in normal form. Chains may still be defined on C^∞ boundaries as solution curves for a family of second-order differential equations; the equations are obtained by osculating the C^∞ boundary to high order with a real-analytic boundary.

The theorems of this chapter are due to Moser [**48**] and Chern and Moser [**11**].

CHAPTER 10. CHERN'S THEOREM

The first half of this chapter contains a description of the relevant bundles, groups and forms, beginning with the flat case and then considering the general case. The second half of the chapter consists of a statement of the theorem and a brief outline of its proof.

1. The flat case. In this chapter we will change notation slightly and denote the boundary of the Siegel domain as

$$Q = \left\{(\omega, z) \in \mathbf{C} \times \mathbf{C}^n : \operatorname{Im} \omega = \sum_1^n |z_\alpha|^2\right\}.$$

Using the quadratic form

$$\langle v, v' \rangle = \frac{-i}{2}\left(v_{n+1}\bar{v}'_0 - v_0 \bar{v}'_{n+1}\right) - \sum_1^n v_\alpha \bar{v}'_\alpha$$

in $\mathbf{C}^{n+2}$ to define

$$E^* = \left\{v \in \mathbf{C}^{n+2} : \langle v, v \rangle = 0\right\},$$

allows us to construct a line bundle

$$E^* \to Q, \qquad v \mapsto \left(\frac{v_{n+1}}{v_0}, \frac{v}{v_0} \right) = (\omega, z).$$

There are several groups naturally associated with this bundle:

(i) The indefinite unitary group $U(n+1,1)$ consisting of matrices over $\mathbf{C}^{n+2}$ preserving the quadratic form $\langle\ ,\ \rangle$.

(ii) The group of Heisenberg translations $N \subset U(n+1,1)$. For $(\omega', z') \in Q$, the translation $T_{(\omega',z')} \in N$ is defined by

$$T_{(\omega',z')}: (\omega, z) \mapsto (\omega + \omega' + 2iz \cdot z', z + z')$$

and is identified with the matrix

$$\left[\begin{array}{c|ccc|c} 1 & 0 & \cdots & 0 & 0 \\ \hline z_1' & & & & 0 \\ \vdots & & I & & \vdots \\ z_n' & & & & 0 \\ \hline \omega' & 2i\bar{z}_1' & \cdots & 2iz_n' & 1 \end{array}\right] \in U(n+1,1).$$

(iii) The Heisenberg dilations $\mathbf{R}^\times \subseteq U(n+1,1)$. For $t \in \mathbf{R}^\times$, the Heisenberg dilation D_t acts on Q by

$$D_t: (\omega, z) \mapsto (t^2\omega, tz)$$

and is identified with the matrix

$$\left(\begin{array}{c|c|c} t^{-1} & 0 & 0 \\ \hline 0 & I & 0 \\ \hline 0 & 0 & t \end{array}\right) \in U(n+1,1).$$

(iv) The Heisenberg rotations $U(n) \subseteq U(n+1,1)$. The matrix $(U_{\alpha\beta}) \in U(n)$ is associated with the rotation $R_{U_{\alpha\beta}}$ which acts on Q by

$$R_{U_{\alpha\beta}}: (\omega, z) \mapsto \left(\omega, \sum_\beta U_{\alpha\beta} z_\beta\right)$$

and is identified with the matrix

$$\left(\begin{array}{c|c|c} 1 & 0 & 0 \\ \hline 0 & U_{\alpha\beta} & 0 \\ \hline 0 & 0 & 1 \end{array}\right).$$

(v) The isotropy groups H^+ and H. These are defined by

$$\begin{aligned} H^+ &= \{\text{group of LFT of } Q \text{ fixing the origin}\} \\ &= \{T \in U(n+1,1) \colon Te^0 = \lambda e^0, \text{ for some } \lambda \in \mathbf{C}, \text{ modulo } S^1\}, \end{aligned}$$

$$\begin{aligned} H &= \{\text{group of LFT } \phi \text{ of } Q \text{ fixing } 0 \text{ with } |\det \phi'(0)| = 1\} \\ &= \{T \in U(n+1,1) \colon Te^0 = e^0\}, \end{aligned}$$

where LFT = linear fractional transformation, S^1 is thought of as the unit circle in $\mathbf{C}$, and $(e^0, \ldots, e^{n+1})$ is the standard basis of $\mathbf{C}^{n+2}$. If we think of $H \subseteq U(n+1,1)$ as matrices, then H consists of matrices of the form

$$\left(\begin{array}{c|c|c} 1 & v_1 \cdots v_n & s + i\sum |v_\alpha|^2 \\ \hline 0 & & \frac{i}{2}\sum\limits_\alpha u_{1\alpha}\bar{v}_\alpha \\ \vdots & u_{\alpha\beta} & \vdots \\ 0 & & \frac{i}{2}\sum\limits_\alpha u_{n\alpha}\bar{v}_\alpha \\ \hline 0 & 0 \cdots 0 & 1 \end{array}\right),$$

where $(u_{\alpha\beta})$ is unitary, $v_\alpha \in \mathbf{C}$ and $s \in \mathbf{R}$. The group H^+ is generated by H and the Heisenberg dilations and consists of matrices of the obvious form. Recall that the Lie algebras of $U(n+1,1)$, N, H^+, H are denoted by $u(n+1,2)$, n, h^+, h and were described in a previous chapter, where it was shown that

(1) $$u(n+1,1) \cong n + h^+.$$

The next step is to construct a frame bundle

$$\begin{array}{c} Y^* \\ \pi \downarrow \\ E^* \end{array},$$

where Y^* consists of all frames $(z_0, \ldots, z_{n+1})$ such that

$$\langle z_j, z_k \rangle = \langle e^j, e^k \rangle$$

and the projection is

$$(z_0, \ldots, z_{n+1}) \mapsto z_0.$$

Here $z_j \in \mathbf{C}^{n+2}$ and z_0 satisfies $\|z_0\|^2 = \|e^0\|^2 = 0$. If we associate the frame $(z_0, \ldots, z_{n+1})$ with the matrix which maps

$$(e^0, \ldots, e^{n+1}) \mapsto (z_0, \ldots, z_{n+1}),$$

then the space Y^* can be identified with $U(n+1,1)$.

The splitting (1) defines a natural connection in the frame bundle. At a point $p \in Y^*$, the tangent space splits

$$T_pU(n+1,1) \cong u(n+1,1) \cong n + h^+,$$

and we can use

$$h^+ \cong T_p(\pi^{-1}(\text{point of } E^*)) = \text{"verticle part of } T_pY^*\text{"}$$

to define a connection; in other words, the connection is defined by putting n as the horizontal part of T_pY^*.

The final step is to mod out by the S^1 action consisting of multiplication by a complex number with modulus 1. This gives bundles

$$\begin{array}{c} Y \\ \downarrow \\ E \\ \downarrow \\ Q \end{array}$$

where Y is a principal H-bundle over E; that is, the fibres are isomorphic to H. Roughly speaking, in the general case Chern tries to find bundles over C-R manifolds and connections that resemble those in the flat case.

2. The general case. Let M be a $(2n+1)$-dimensional real manifold which is strictly ψ-convex and has an integrable C-R structure $J^2 = -\mathrm{I}$. This defines forms of type $(1,0)$ and $(0,1)$ on M:

$$\left.\begin{aligned} \omega(JX) &= i\omega(X) \quad \text{type} \quad (1,0) \\ &= -i\omega(X) \quad \text{type} \quad (0,1) \end{aligned}\right\},$$

where $X \in HM \equiv \{\text{complex part of } TM\}$.

Using the C-R structure, we can construct an $\mathbf{R}^+$ bundle over M

$$\begin{array}{c} E \\ \pi^1 \downarrow \\ M \end{array}$$

by specifying the fibre E_p, for $p \in M$:

$$E_p = \{\text{real covectors } \omega \in T_p^*M \colon \omega \perp H_pM\}.$$

A $(1,0)$- or $(0,1)$-form on E is defined to be the pullback of a $(1,0)$- or $(0,1)$-form on M.

In order to define the analogue of the bundle of frames, first we need to define tautological forms. The tautological 1-form $\hat{\omega}$ on E is defined by

$$\hat{\omega}(X) = \left\langle \omega, \pi_*^1 X \right\rangle,$$

for $X \in T_{(q,\omega)}E$ and $q \in M$. This makes sense since $\omega \in T_q^*M$ and $\pi_*^1 X \in T_qM$.

LEMMA 1. *The tautological* 1-*form* $\hat{\omega}$ *on* E *satisfies*

$$\text{(S1)} \qquad d\hat{\omega} = i\sum_{\alpha} \omega^\alpha \wedge \overline{\omega}^\alpha + \hat{\omega} \wedge \phi,$$

where ω^α *are* $(1,0)$-*forms on* E *and* ϕ *is a real* 1-*form on* E.

The lemma is an immediate consequence of the strict ψ-convexity of M. After these preliminaries, we can define a bundle

$$\text{(2)} \qquad \begin{array}{c} Y \\ \pi^2 \downarrow \\ E \end{array}$$

by defining the fibre to be

$$Y_q = \{ (1,0)\text{-covectors } \omega^\alpha \text{ and } 1\text{-covectors } \phi \text{ at } q \in E\colon (\mathrm{S}1) \text{ holds}\}.$$

With the obvious projection this becomes a bundle.

LEMMA 2. *The bundle $Y \to E$ is a principal H-bundle.*

PROOF. We want to define an action of H on the fibres. Given a point $(q, \omega^{\alpha *}, \phi^*) \in Y_q$ and

$$\left(\begin{array}{c|c|c} 1 & v_1 \cdots v_n & s + i\sum |v^\alpha|^2 \\ \hline 0 & & \frac{i}{2}\sum u_{1\alpha}\bar{v}_\alpha \\ \vdots & u_{\alpha\beta} & \vdots \\ 0 & & \frac{i}{2}\sum u_{n\alpha}\bar{v}_\alpha \\ \hline 0 & 0 \cdots 0 & 1 \end{array}\right) \in H,$$

define

$$T(\omega^{\alpha *}, \phi^*) = (\omega^\alpha, \phi), \tag{3}$$

where

$$\omega^\alpha = v_\alpha \hat{\omega} + \sum_\beta u_{\alpha\beta}\omega^{\beta *},$$

$$\phi = s\hat{\omega} + i\sum_\alpha \bar{v}_\alpha \omega^{\alpha *} - i\sum_\alpha v_\alpha \overline{\omega^{\alpha *}} + \phi^*.$$

It is easy to verify that $(q, \omega^\alpha, \phi) \in Y_q$ and that the bundle is principal. □

Our next task is to define the connection; this too involves tautological forms. Let $\xi = (q, \omega^\alpha, \phi) \in Y$ and $X \in T_\xi Y$. The tautological 1-forms $\overset{+}{\omega}$, $\overset{+}{\omega}{}^\alpha$ and $\overset{+}{\phi}$ on Y are defined by

$$\overset{+}{\omega} = (\pi^2)^*\hat{\omega}, \qquad \overset{+}{\omega}{}^\alpha(X) = \left\langle \omega^\alpha, \pi^2_* X\right\rangle, \qquad \overset{+}{\phi}(X) = \left\langle \phi, \pi^2_* X\right\rangle,$$

where $\hat{\omega}$ is the tautological 1-form on E. This is well defined since $\omega^\alpha \in T_q^* E$ and $\pi^2_* X \in T_q E$. Note that (S1) holds on Y:

$$d\overset{+}{\omega} = i\sum \overset{+}{\omega}{}^\alpha \wedge \overline{\overset{+}{\omega}{}^\alpha} + \overset{+}{\omega} \wedge \overset{+}{\phi}. \tag{S1}$$

Recall that a connection for the principal H-bundle (2) is an h-valued 1-form on Y that is invariant under the H-action (3) in the fibre; in other words, a connection is a matrix (π_{pq}) of 1-forms such that $(\pi_{pq}(X)) \in h$, for $X \in TY$. The "horizontal subspace" of the connection is the kernel of this 1-form.

In the flat case $M = Q$ described in the first section, we used the splitting (1) of the tangent space to define a connection in the principal H-bundle $Y \to E$. This determines the connection 1-form (π_{pq}). Not only does the first structural equation (S1) hold in the flat case, but also the following equations:

$$(\mathrm{S2})_{\mathrm{flat}} \qquad d\overset{+}{\omega}{}^{\alpha} = \sum_{\beta} \overset{+}{\omega}{}^{\beta} \wedge \phi_{\beta\bar{\alpha}} + \overset{+}{\omega} \wedge \bar{\phi}_{\alpha},$$

$$(\mathrm{S3})_{\mathrm{flat}} \qquad d\overset{+}{\phi} = i\sum \overset{+}{\omega}{}^{\beta} \wedge \phi_{\beta} + i\sum \bar{\phi}_{\beta} \wedge \overline{\overset{+}{\omega}{}^{\beta}} + \overset{+}{\omega} \wedge \psi,$$

where the forms $\overset{+}{\omega}$, $\overset{+}{\omega}{}^{\alpha}$, ϕ_{α}, $\phi_{\alpha\beta}$ and ψ are defined by the connection according to

$$(4) \qquad \overset{+}{\omega} = \frac{1}{2}\pi_{n+1,0}, \qquad \overset{+}{\omega}{}^{\alpha} = \pi_{\alpha 0}, \qquad \overset{+}{\phi} = -\pi_{00} + \pi_{n+1,n+1},$$
$$\phi\alpha = 2\bar{\pi}_{\alpha,n+1}, \qquad \psi = -4\pi_{0,n+1}, \qquad \phi_{\alpha\bar{\beta}} = \pi_{\beta\alpha} - \delta_{\alpha\beta}\pi_{00}.$$

That is, knowning the connection is equivalent to knowing all the forms in (S1), (S2) and (S3). Finally, in the flat case we can check that the following equations are also valid:

$$(\mathrm{S4})_{\mathrm{flat}} \qquad d\phi_{\beta\bar{\alpha}} = \sum_{\gamma} \phi_{\beta\bar{\gamma}} \wedge \phi_{\gamma\bar{\alpha}} + i\overline{\overset{+}{\omega}{}^{\beta}} \wedge \bar{\phi}_{\alpha}$$
$$-i\phi_{\beta} \wedge \overset{+}{\omega}{}^{\alpha} - i\delta_{\beta\alpha}\left(\sum_{\sigma} \phi_{\sigma} \wedge \overset{+}{\omega}{}^{\sigma}\right) - \frac{1}{2}\delta_{\beta\alpha}\psi \wedge \overset{+}{\omega},$$

$$(\mathrm{S5})_{\mathrm{flat}} \qquad d\bar{\phi}_{\alpha} = \overset{+}{\phi} \wedge \bar{\phi}_{\alpha} + \sum_{\beta} \bar{\phi}_{\beta} \wedge \phi_{\beta\bar{\alpha}} - \frac{1}{2}\psi \wedge \overset{+}{\omega},$$

$$(\mathrm{S6})_{\mathrm{flat}} \qquad d\psi = \overset{+}{\phi} \wedge \psi + 2i\sum_{\beta} \bar{\phi}_{\beta} \wedge \phi_{\beta},$$

$$(\mathrm{T1}) \qquad \phi_{\alpha\bar{\beta}} + \bar{\phi}_{\beta\bar{\alpha}} - \delta_{\alpha\bar{\beta}}\,\overset{+}{\phi} = 0 \quad \text{and} \quad \psi \text{ is real.}$$

In the general case in which M is no longer assumed to be Q, rather than define the connection immediately, we will first write down the appropriate equations that the tautological forms satisfy. Equations (S1), (S2) and (S3) are the same as (S1), $(\mathrm{S2})_{\mathrm{flat}}$ and $(\mathrm{S3})_{\mathrm{flat}}$. Equation (S4) is obtained by adding

$$+\sum_{\rho\sigma} S_{\beta\rho\bar{\alpha}\bar{\sigma}}\, \overset{+}{\omega}{}^{\rho} \wedge \overline{\overset{+}{\omega}{}^{\sigma}} + \sum_{\rho} V_{\beta\bar{\alpha}\rho}\, \overset{+}{\omega}{}^{\rho} \wedge \overset{+}{\omega} - \sum_{\rho} \bar{V}_{\beta\bar{\alpha}\rho}\, \overset{+}{\omega}{}^{\rho} \wedge \overset{+}{\omega}$$

to the right-hand side of $(\mathrm{S4})_{\mathrm{flat}}$; similarly, (S5) arises by adding

$$-\sum_{\beta\gamma} V_{\beta\bar{\alpha}\gamma}\, \overset{+}{\omega}{}^{\beta} \wedge \overset{+}{\omega}{}^{\gamma} + \sum_{\beta\sigma} \bar{V}_{\alpha\bar{\beta}\sigma}\, \overset{+}{\omega}{}^{\beta} \wedge \overline{\overset{+}{\omega}{}^{\alpha}} + \sum_{\beta} P_{\beta\bar{\alpha}}\, \overset{+}{\omega}{}^{\beta} \wedge \overset{+}{\omega} + \sum_{\beta} Q_{\bar{\beta}\bar{\alpha}}\, \overline{\overset{+}{\omega}{}^{\beta}} \wedge \overset{+}{\omega}$$

to the right-hand side of $(\mathrm{S5})_{\mathrm{flat}}$. Finally the correction term that yields (S6) from $(\mathrm{S6})_{\mathrm{flat}}$ is another complicated expression that involves not only the tensors S, V,

P and Q but also R. These tensors are required to satisfy the symmetry conditions

$$S_{\beta\gamma\bar{\alpha}\bar{\sigma}} = S_{\gamma\beta\bar{\sigma}\bar{\alpha}} = S_{\gamma\beta\bar{\alpha}\bar{\sigma}} = \bar{S}_{\alpha\sigma\bar{\gamma}\bar{\beta}},$$

(T2) $$\sum_\alpha S_{\alpha\gamma\bar{\alpha}\bar{\sigma}} = 0, \qquad \sum_\rho V_{\beta\bar{\rho}\rho} = 0, \qquad \operatorname{Re}\Big(\sum_\alpha P_{\alpha\bar{\alpha}}\Big) = 0.$$

THEOREM 1 (CHERN). (i) *In the situation above we can find unique* 1*-forms* $\phi_{\beta\bar{\alpha}}$, ϕ_α *and* ψ *on* Y *and functions*

$$S_{\beta\gamma\bar{\alpha}\bar{\sigma}}, \qquad V_{\alpha\bar{\beta}\sigma}, \qquad P_{\beta\bar{\alpha}}, \qquad Q_{\bar{\beta}\bar{\alpha}}, \qquad R_\sigma$$

which satisfy (S1)-(S6) *and* (T1), (T2).

(ii) *There exists a unique connection* (π_{pq}) *on* Y *so that* (4) *holds.*

(iii) *All the forms and functions above can be read off at a given point of* Y *just by knowning finitely many terms in the Taylor expansion of the C-R-structure at the given point of* M.

3. The ideas in the proof of Chern's theorem. We begin this section with some general comments about Cartan's method of equivalence and end with a brief description of how this method is used to prove Chern's theorem. As background, we show how the method of equivalence can be used in Riemannian geometry.

Given a map

$$(M, \text{geometric structure}) \to (M', \text{same type of geometric structure}),$$

Cartan's method provides a means of determining whether locally there exists a diffeomorphism $f\colon M \to M'$ that preserves this structure.

EXAMPLE. A G-structure on a manifold M consists of a coframe θ_j defined up to the equivalence

(5) $$\theta_j^* = \sum_k \mu_{jk}\theta_k,$$

where

$$(\mu_{jk}) \in G \subseteq \mathrm{GL}(n, \mathbf{R}).$$

Assume now that $G = \mathrm{I}$. We say that f preserves the G-structure if

$$f^*(\theta_j') = \theta_j, \qquad j = 1, \ldots, n.$$

If we define c_{jkl} and c'_{jkl} by

$$d\theta_j = \sum_{k,l} c_{jkl}\theta_k \wedge \theta_l, \qquad d\theta_j' = \sum_{k,l} c'_{jkl}\theta_k' \wedge \theta_l',$$

then $c_{jkl} = c'_{jkl}$, whenever f preserves the G-structure. This is because $f^*(d\theta_j') = d\theta_j$. There are two extreme cases of interest.

(i) c_{jkl} contain a coordinate system,

(ii) c_{jkl} are all constant.

Given a general G, we would like to reduce to the case $G = \mathrm{I}$.

Step 1. The group G can be reduced by using the (μ_{jk}) in (5) to determine certain of the c_{jkl}. If this procedure does not reduce G to I, then we go on to Step 2.

Step 2. In this step we "prolong" the G-structure on M to a G'-structure on $M \times G$. The G'-structure is defined by a coframe $\{\omega_j, \omega_{jk}\}$, where

$$\omega_j = \sum_k \mu_{jk}\theta_k, \qquad \omega_{jk} = d\mu_{jk} + (\text{other terms}).$$

We now apply Step 1 with the new G'-structure. Note that this procedure does not necessarily stop, since G' may be larger than G.

As an illustration of this method, consider the Riemannian case in which the coframes determine a metric

$$ds^2 = \theta_1^2 + \cdots + \theta_n^2,$$

which is defined up to the equivalence

$$\theta_j^* = \mu_{jk}\theta_k, \qquad \text{where } \mu_{jk}\mu_{ik} = \delta_{ij}.$$

Here, and throughout the remainder of the section, we will use the summation convention. The coframe determines the structure constants c_{jkl} by

$$d\theta_j = c_{jkl}\theta_k \wedge \theta_l,$$

and we calculate that

$$d\theta_j^* = d\mu_{jk} \wedge \mu_{ik}\theta_i^* + \mu_{jk}c_{klm}\theta_l \wedge \theta_m.$$

Applying Step 1 to this equation gets us into trouble, because of the appearance of the $d\mu_{jk}$. Instead we prolong according to Step 2 by introducing the new coframe $\{\omega_j, \omega_{jk}\}$ defined by

$$\omega_j = \mu_{jk}\theta_k, \qquad d\omega_j = \omega_{ji} \wedge \omega_i + T_{jkl}\omega_k \wedge \omega_l,$$

where

$$\omega_{ij} + \omega_{ji} = 0, \qquad T_{jkl} + T_{jlk} = 0.$$

T_{jkl} is the torsion tensor. Using the equivalence (5) we get

$$d\omega_j = \omega_{ji} \wedge \omega_i + T_{jkl}\omega_k \wedge \omega_l = \omega_{ji}^* \wedge \omega_i + T_{jkl}^*\omega_k \wedge \omega_l \tag{6}$$

or

$$\left(\omega_{ji} - \omega_{ji}^*\right) \wedge \omega_i + \left(T_{jkl} - T_{jkl}^*\right)\omega_k \wedge \omega_l = 0,$$
$$\left[\omega_{ji} - \omega_{ji}^* + \left(T_{jki} - T_{jki}^*\right)\omega_k\right] \wedge \omega_i = 0,$$

which by linear algebra yields

$$\omega_{ji} - \omega_{ji}^* = a_{jik}\omega_k, \quad \text{where } a_{jik} + a_{ijk} = 0,$$

by the skew symmetry of the ω_{ij}. (The linear algebra used here is called the *Cartan Lemma*.) We can now apply Step 1 using the two equations

$$\omega_{ji} = \omega_{ji}^* + a_{jik}\omega_k,$$
$$\left(\tfrac{1}{2}\left(a_{jik} - a_{jki}\right) + T_{jki} - T_{jki}^*\right)\omega_k \wedge \omega_l = 0$$

for G'; this requires that we determine the a_{ijk} such that $T_{ijk}^* = 0$ in the $nn(n+1)/2 + nn(n-1)/2 = n^3$ equations

$$a_{jik} + a_{ijk} = 0, \qquad a_{jik} - a_{jki} = -2T_{jki}.$$

This system can be solved using the trick of "rotating indices", in other words, by suitably permuting the indices. We find that

$$d\omega_{ij} = \omega_{ik} \wedge \omega_{kj} + R_{ijkl}\omega_k \wedge \omega_l, \tag{7}$$

where R_{ijkl} is the curvature tensor. Equations (6) and (7) are the Cartan structural equations, which we have derived through his method of equivalence.

We now turn to the proof of Chern's theorem. Let M^{2n+1} be a C-R manifold and let $\mathcal{I}$ be a closed complex differential ideal generated by 1-forms. In the case of the hypersurface Q considered in the first section, $\mathcal{I}$ is the ideal generated by the restrictions of $(1,0)$-forms to Q. Let $\mathcal{I}_1 = \{$1-forms in $\mathcal{I}\}$, so that

$$\mathcal{I}_1 + \bar{\mathcal{I}}_1 = \Lambda^1, \qquad \mathcal{I}_1 \cap \bar{\mathcal{I}}_1 = (\theta),$$

where Λ^1 denotes the 1-forms on M and θ is the annihilator of the maximal complex tangent space HM. Choose forms θ^α so that $\{\theta, \theta^\alpha\}$ is a basis for $\mathcal{I}$ and consider the equivalence

$$\theta^* = u\theta, \qquad \theta^{\alpha^*} = v^\alpha\theta + u^\alpha_\beta\theta^\beta, \qquad \theta^{\bar{\alpha}^*} = v^{\bar{\alpha}}\theta + u^{\bar{\alpha}}_{\bar{\beta}}\theta^{\bar{\beta}},$$

where the notation of the first section has been changed slightly so that $u_{\alpha\beta}$ is now denoted u^α_β. Because of the integrability condition $d\mathcal{I} \subset \mathcal{I}$, we can write

$$d\theta = ih_{\alpha\bar{\beta}}\theta^\alpha \wedge \theta^{\bar{\beta}} \mod \theta.$$

Once again we must prolong; but this time instead of prolonging by looking at $M \times G$, we will prolong "along θ". We can define a basis $\{\omega, \omega^\alpha, \omega^{\bar{\alpha}}, \phi\}$ for a coframe on $E = M \times \mathbf{R}$ by using the equations

$$\begin{aligned} \omega &= u\theta, \\ d\omega &= iuh_{\alpha\bar{\beta}}\theta^\alpha \wedge \theta^{\bar{\beta}} + \omega \wedge (-du/u + \text{something}) \\ &= ig_{\alpha\bar{\beta}}\omega^\alpha \wedge \omega^{\bar{\beta}} + \omega \wedge \phi, \\ \omega^\alpha &= \text{multiples of } \theta^\alpha. \end{aligned} \tag{8}$$

The basis is defined up to the equivalence

$$\begin{aligned} &\omega^* = \omega, \qquad \omega^{\alpha^*} = v^\alpha\omega + u^\alpha_\beta\omega^\beta, \qquad \omega^{\bar{\alpha}^*} = v^{\bar{\alpha}}\omega + u^{\bar{\alpha}}\omega + u^{\bar{\alpha}}_{\bar{\beta}}\omega^{\beta}, \\ &\phi^* = s\omega + ig_{\rho\bar{\sigma}}\left(u^\rho_\beta v^{\bar{\sigma}}\omega^\beta - u^{\bar{\sigma}}_{\bar{\beta}} v^\rho\omega^\beta\right) + \phi. \end{aligned}$$

In order to apply Step 1 to this prolongation we need an expression for $d\phi$ and $d\phi^\alpha\beta$. To begin this calculation, differentiate (8) to get

$$0 = i\left(dg_{\alpha\bar{\beta}} - g_{\gamma\bar{\beta}}\phi^\gamma_\alpha + g_{\alpha\bar{\gamma}}\phi^{\bar{\gamma}}_{\bar{\beta}} - g_{\alpha\bar{\beta}}\phi\right)\omega^\alpha \wedge \omega^{\bar{\beta}} + \omega \wedge (d\phi + \text{something}).$$

A bit of linear algebra shows us that the ϕ^γ_α can be chosen so that

$$dg_{\alpha\bar{\beta}} - g_{\gamma\bar{\beta}}\phi^\gamma_\alpha + g_{\alpha\bar{\gamma}}\phi^{\bar{\gamma}}_{\bar{\beta}} - g_{\alpha\bar{\beta}}\phi = 0;$$

which means $\omega \wedge (d\phi + \text{something}) = 0$ or $d\phi = -(\text{something})$. In fact, we find

$$d\phi = i\omega_{\bar{\beta}} \wedge \phi^{\bar{\beta}} + i\phi_{\bar{\beta}} \wedge \omega^{\bar{\beta}} + \omega \wedge \psi,$$

where $\omega_{\bar{\beta}} = g_{\alpha\bar{\beta}}\omega^\alpha$, etc. In order to find a similar expression for $d\phi^\alpha_\beta$, note that (8) and the integrability condition give

$$d\omega^\alpha = \omega^\beta \wedge \phi^\alpha_\beta + \omega \wedge \phi^\alpha.$$

Differentiating this and working very hard yields the desired expression for $d\phi^\alpha_\beta$. Using Step 1 now, we find, after many uses of Cartan's Lemma, that G has been reduced to the identity. This concludes the sketch of the proof of Chern's theorem.

The theorems in this chapter are due independently to Chern-Moser [**11**] and Tanaka [**58**, **59**] who built upon the early work of Cartan in two variables [**8**].

4. Applications. We explain briefly how to use Theorem 1 to define chains, and also to write down finitely many differential equations on ψ which hold if and only if $D = \{\psi(z) > 0\}$ is locally biholomorphic to the ball. To find the differential equations, we use a result of Chern in [**11**] that the C-R manifold M is flat (i.e. locally biholomorphic to a sphere) if and only if the "curvatures" $(S_{\beta\gamma\bar{\alpha}\bar{\sigma}})$, $(V_{\alpha\bar{\beta}\sigma})$, $(P_{\beta\bar{\alpha}})$, $(Q_{\bar{\beta}\bar{\alpha}})$, (R_σ) in Theorem 1 all vanish. This is analogous to the fact that a Riemannian metric is locally isometric to R^n if and only if the curvature R_{ijkl} vanishes. Since $\psi \to [(S_{\beta\gamma\bar{\alpha}\bar{\sigma}})$ etc. associated to the boundary $M = \{\psi(z) = 0\}]$ is a nonlinear differential operator by Theorem 1, part (iii), it follows that the boundaries $\{\psi = 0\}$ locally biholomorphic to the sphere have been characterized by finitely many differential equations on ψ.

Regarding the chains, we start by recalling that the connection in Theorem 1 canonically identifies the tangent space T_pY at each point with the Lie algebra $\mathrm{u}(n+1, 1)$. Therefore, any element $\xi \in \mathrm{u}(n+1,1)$ gives rise to a vector field on Y. Integrating the vector field, we get family of parametrized curves in Y. Finally, projecting from Y down to M, we associate to each $\xi \in \mathrm{u}(n+1,1)$ a distinguished family of parametrized curves on M. For the particular element

$$\xi_0 = \begin{pmatrix} 0 & & & \\ \vdots & & 0 & \\ 0 & & & \\ 1 & 0 & \cdots & 0 \end{pmatrix} \in \mathrm{u}(n+1,1)$$

(infinitesimal Heisenberg translation by $(0, 0, \ldots, 0, \omega)$) the family of curves obtained by this procedure agrees with Moser's chains. Since the main step here is to integrate a vector field, it is clear that chains are given by solutions of ordinary differential equations. Curiously, the parametrization of chains given above in terms of Y does not agree with Moser's parametrization. Theorem 1 also allows us to define a notion of parallel transport of frames, which agrees with Moser's definition in Chapter 9. In two dimensions, these ideas go back to E. Cartan [**8**], who also studied the distinguished curves corresponding to other ξ in the Lie algebra.

The results on chains stated here are proved by computing the connection in the special case of a boundary in normal form. Details are formidable.

CHAPTER 11. THE COMPLEX MONGE-AMPÈRE EQUATION

The first section discusses the formal properties of a complex Monge-Ampère equation, while the second section sketches a proof of the Cheng-Yau theorem. This theorem establishes existence, uniqueness, C^∞ interior regularity, and some regularity up to the boundary for solutions of the basic equation. The final section describes chains and local invariants of pseudoconvex boundaries as an application of the formal study.

1. The Monge-Ampère equation. The problem is to find on a given strictly pseudoconvex domain $D \subseteq \mathbf{C}^N$ a complete Kähler-Einstein metric

$$ds^2 = \sum_{jk} g_{j\bar{k}}(z)\, dz_j\, \overline{dz_k}\,.$$

A Kähler metric ds^2 is called Kähler-Einstein if $\mathrm{Ric}_{j\bar{k}} = -kg_{j\bar{k}}$, where k is a constant which we take here to be positive. Recall that the Ricci tensor for a Kähler metric is given by

$$\mathrm{Ric}_{j\bar{k}} = c_n \frac{\partial^2}{\partial z_j \partial \bar{z}_k} \log \det(g_{\mu\bar{\nu}}).$$

So we are looking for solutions $g_{j\bar{k}}$ of

$$g_{j\bar{k}} = (\mathrm{const}) \frac{\partial^2}{\partial z_j \partial \bar{z}_k} \log \det(g_{\mu\bar{\nu}}). \tag{1}$$

Setting $\zeta = \det(g_{\mu\bar{\nu}})$, we see that solving (1) is equivalent to solving the complex Monge-Ampère equation

$$\text{(M-A1)} \qquad \begin{aligned} \det\left(\frac{\partial^2}{\partial z_j \partial \bar{z}_k} \log \zeta\right) &= (\mathrm{const})\zeta \quad && \text{in } D, \\ \zeta &\to \infty && \text{at } \partial D, \end{aligned}$$

since if ζ solves (M-A1) then $g_{j\bar{k}} = (\mathrm{const})(\partial^2/\partial z_j \partial \bar{z}_k) \log \zeta$ is a solution of the Kähler-Einstein equation (1).

EXAMPLE. For the unit ball $D = \{|z| < 1\} \subseteq \mathbf{C}^n$,

$$\zeta = c_n / \left(1 - |z|^2\right)^{n+1} \tag{2}$$

is a solution of (M-A1).

This suggests that for a more general domain D, we try a solution of the form

$$\zeta = c_n / \left(u(z)\right)^{n+1}, \tag{3}$$

where $u(z) = 0$ at ∂D to exactly first order. Rewriting the Monge-Ampère equation (M-A1) in terms of $u(z)$ gives

$$\text{(M-A2)} \qquad \begin{aligned} J(u) &= 1 \quad && \text{in } D, \\ u &= 0 && \text{at } \partial D, \end{aligned}$$

where

$$J(u)=(-1)^n\det\begin{pmatrix} u & \dfrac{\partial u}{\partial \bar z_1} & \cdots & \dfrac{\partial u}{\partial \bar z_n} \\ \dfrac{\partial u}{\partial z_1} & & & \\ \vdots & & \dfrac{\partial^2 u}{\partial z_j \partial \bar z_k} & \\ \dfrac{\partial u}{\partial z_n} & & & \end{pmatrix}.$$

In constructing the formal solution to this equation, we shall see that "much" of the Taylor series of $u(z)$ at ∂D is specified infinitesimally. This is in contrast to the Dirichlet problem for the Laplace equation

$$(4)\qquad \begin{aligned} \Delta u &= 0 \quad \text{on } \mathbf{R}^{n+1}_+, \\ u &= f \quad \text{on } \mathbf{R}^n, \end{aligned}$$

where $\mathbf{R}^{n+1}_+ = \{(x_1,\ldots,x_n,t)\in \mathbf{R}^{n+1} \mid t \geqslant 0\}$.

In (4), $\partial u/\partial t$ evaluated at the boundary is not determined locally by f, although it is determined by f globally. Once it is determined, all the higher derivatives $(\partial/\partial t)^k u\,|_{\mathbf{R}^n}$ on the boundary are fixed, since for example, $(\partial^2/\partial t^2)u\,|_{\mathbf{R}^n} = -\Delta_x f$. In this formal sense the Monge-Ampère equation (M-A2) is better than the Laplace equation.

Now we shall look for a "formal" solution of (M-A2), that is a function $u(z)$ vanishing at ∂D, which satisfies

$$(5)_s \qquad J(u) = 1 + O\big([\mathrm{dist}(z,\partial D)]^s\big), \quad \text{in } D,$$

with s as large as possible. To do so, we exploit two simple identities for the J-operator, which are proved by applying row and column operations to the definition. The identities are

$$(6.\mathrm{i}) \qquad J(\phi\psi) = \phi^{n+1}J(\psi) + \psi(\text{Remainder}),$$

$$(6.\mathrm{ii}) \qquad J(\psi+\phi\psi^s) = J(\psi)\cdot\big[1 + c_{n,s}\phi\psi^{s-1}\big] + \psi^s\cdot(\text{Remainder}) \qquad (s\geqslant 2)$$

where the remainders are smooth functions on $\bar D$. Now let ψ be a smooth defining function for D; this means $D=\{\psi>0\}$ and $\psi'\neq 0$ on ∂D. The strict pseudo-convexity of D shows that $J(\psi)>0$ on ∂D. Hence we may take $\phi = [J(\psi)]^{-(1/n+1)}$ in (6.i) to obtain a smooth function $u_1 = [J(\psi)]^{-(1/n+1)}\psi$ vanishing on ∂D, so that $J(u_1) = 1 + O(\psi)$. So we have solved (5) with $s=1$.

Next suppose we have found a smooth function u_{s-1} which vanishes at ∂D and satisfies $J(u_{s-1}) = 1 + O(\psi^{s-1})$ $(s\geqslant 2)$. Thus, $J(u_{s-1}) = 1 + \eta(u_{s-1})^{s-1}$ for an $\eta\in C^\infty(\bar D)$. Setting $u_s = u_{s-1} + \phi(u_{s-1})^s$ and using (6.ii) gives

$$J(u_s) = \big[1+\eta(u_{s-1})^{s-1}\big]\cdot\big[1 + c_{n,s}\phi(u_{s-1})^{s-1}\big] + O(\psi^s).$$

So if we put

$$\phi = \frac{-\eta}{c_{n,s}} = \frac{1 - J(u_{s-1})}{c_{n,s}(u_{s-1})^{s-1}},$$

then we get

$$(7) \qquad J(u_s) = 1 + O(\psi^s), \qquad u_s = u_{s-1} \cdot \left[1 + \frac{1 - J(u_{s-1})}{c_{n,s}}\right].$$

As long as $c_{n,s} \neq 0$, this provides an inductive procedure to pass from solutions of $(5)_{s-1}$ to solutions of $(5)_s$. It happens that $c_{n,s} = s(n + 2 - s)$, which vanishes exactly for $s = (n + 2)$, so our inductive procedure stops there, yielding a smooth function u which vanishes at ∂D and satisfies $J(u) = 1 + O(\psi^{n+1})$. Note that u arises by applying a nonlinear differential operator to any defining function ψ; in fact $u = P_{n+1}P_nP_{n-1} \cdots P_2P_1\psi$ where $P_1(\psi) = [J(\psi)]^{-(1/n+1)}\psi$, and P_s: $u_{s-1} \to u_s$ by (7). An easy induction on s using (6.i), (6.ii) shows that the condition $J(u) = 1 + O(\psi^s)$ specifies u up to an error $O(\psi^{s+1})$. So our formal solution u is determined modulo $O(\psi^{n+2})$. For more details see [**22**].

The fact that $c_{n,s} = 0$ for $s = n + 2$ shows that the actual solution to (M-A2) is not smooth up to the boundary for general strictly pseudoconvex D, since we will already get a contradiction by Taylor expanding to order $(n + 2)$ about a boundary point. Rather, we expect logarithmic terms, so that the full solution to (M-A 2) should have the following asymptotic expansion near the boundary:

$$(8) \qquad u \sim \phi\psi + \sum_{p,q \geq 1} \phi_{pq}\psi^p(\psi^{n+1}\log\psi)^q, \quad \text{with } \phi, \phi_{pq} \in C^\infty(\bar{D}).$$

Part of the information contained in ϕ, ϕ_{pq} is determined locally by ∂D, the rest is determined globally. It would be very interesting to understand this precisely; a beautiful result in this direction has recently been obtained by R. Graham [**29**].

From (8) we obtain formally $u \in C^{n+2-\varepsilon}(\bar{D})$. The Cheng-Yau theorem gives us almost this much regularity for the solution of (M-A2).

2. The Cheng-Yau theorem. The goal of this section is to sketch the proof of the following

THEOREM 1 (CHENG-YAU) [**10**]. *The solution u of* (M-A2) *exists and is unique. Moreover u is C^∞ in the interior of D and belongs to $C^{n+(3/2)-\varepsilon}(\bar{D})$.*

Recall that $u \in C^{n+(3/2)-\varepsilon}(\bar{D})$ if $(\partial/\partial x)^\alpha u \in \text{Lip}(\frac{1}{2} - \varepsilon)$ for $|\alpha| \leq n + 1$.

To prove the theorem, we start with the approximate solution ψ of (M-A2) constructed in §1. Thus ψ is smooth up to the boundary and $J(\psi) = e^{-F}$ where F vanishes to order $(n + 1)$ at ∂D. We shall look for solutions of (M-A2) of the form $u = \psi e^v$; we hope that v will vanish to high order at ∂D so that $u \in C^{n+(3/2)-\varepsilon}(\bar{D})$. We shall study v in terms of the approximate Poincaré metric $ds^2 = \sum_{j,k} g_{j\bar{k}}\, dz_j\, d\bar{z}_k$, $g_{j\bar{k}} = (\partial^2/\partial z_j \partial\bar{z}_k)\log 1/\psi$. In terms of the unknown v and the metric $g_{j\bar{k}}$, equation (M-A2) may be rewritten in the form

$$\text{(M-A3)}\qquad \begin{aligned} &\det(g_{j\bar{k}} + v_{j\bar{k}}) = e^{F} e^{(n+1)v} \det(g_{j\bar{k}}) && \text{in } D,\\ &(g_{j\bar{k}} + v_{j\bar{k}}) > 0 && \text{in } D,\\ &v \to 0 && \text{at } \partial D, \end{aligned}$$

where subscripts on v denote differentation. Again we note that F is smooth and tends rapidly to zero at ∂D, and we hope that v will do the same.

To solve (M-A3) we use the "continuity method"; that is, we introduce a parameter $0 \leqslant s \leqslant 1$ and look for solutions v^s of

$$\text{(9)}\qquad \det\left(g_{j\bar{k}} + v^s_{j\bar{k}}\right) = e^{sF} e^{(n+1)v^s} \det(g_{j\bar{k}}).$$

For $s = 0$ the equation has a trivial solution $v^s = 0$, while for $s = 1$ we recover (M-A3). We attempt to vary v^s continuously as s moves from 0 to 1 and still satisfy (9). Differentiating (9) in the parameter s leads to a simple linearized equation:

$$\text{(10)}\qquad \{-(n+1)I + \Delta'\}\frac{\partial V}{\partial s} = (\text{given functions defined in terms of } v),$$

where Δ' is the Laplacian in the metric $g'_{j\bar{k}} = g_{j\bar{k}} + v^s_{j\bar{k}}$. Since $c > 0$, the maximum principle lets us invert $\{-(n+1)I + \Delta'\}$, so that (10) is a kind of ODE for $s \to v^s$. Hence we can continue to solve (9) for small time s.

What we need in order to continue the solution all the way to $s = 1$ are good a priori bounds on v^s. For, if v^s is defined for $s \leqslant s_0$ and v^{s_0} satisfies good bounds independent of ε_0, then (10) may be used to extend v^s to $s \leqslant s_0 + \varepsilon$ with ε independent of s_0. Thus we can get to $s = 1$.

Now, proving good bounds on solutions of (9) is equivalent to proving good bounds for solutions of (M-A3).

To summarize: We assume v is a bounded, smooth solution of (M-A3), and we try to prove good a priori estimates for v and its derivatives. If we can do this, then the existence and regularity will follow by the continuity method.

We now explain how to estimate solutions of (M-A3) in the interior. The estimates are based on tricks rooted in a long literature on the real and complex Monge-Ampère equations. An important part is played by the

MAXIMUM PRINCIPLE. Let u be a smooth, bounded, real-valued function in D. Then there exists a sequence $x_1, x_2, \ldots$ in D such that

(i) $u(x_i) \to \sup_D u$,

(ii) $\|u'(x_i)\| \to 0$,

(iii) $\limsup_{i\to\infty}$ (highest eigenvalue of $u''(x_i)$) $\leqslant 0$.

Here the derivatives u', u'' are covariant derivatives with respect to $g_{j\bar{k}}$, and norms and eigenvalues are also taken with respect to $g_{j\bar{k}}$. The maximum principle can be proved easily by subtracting a barrier function from u.

We begin now with a 0th order estimate for solutions of (M-A3). Apply the maximum principle to v. Part (iii) above yields

$$0 \leqslant (g_{j\bar{k}} + u_{j\bar{k}})(x_i) \leqslant (1+\varepsilon)(g_{j\bar{k}})(x_i) \quad \text{for large } i,$$

so (M-A3) and (i) imply

$$(1+\varepsilon)^n \geqslant \limsup_{i\to\infty} e^{F(x_i)}e^{(n+1)v(x_i)} \geqslant e^{-\inf_D F}e^{(n+1)\sup_D v}.$$

Thus $\sup_D v \leqslant C_1$, where C_1 depends only on F. An analogous application of the maximum principle to $-v$ shows that $\sup_D(-v) \leqslant C_2$, where C_2 depends only on F. Hence we get an apriori bound

$$|v(z)| \leqslant C \quad \text{in } D.$$

Next we estimate the second derivatives $v_{j\bar{k}}$ by applying the maximum principle to $e^{-cv}(n+\Delta v)$, where Δ, Δ' denote the Laplacians in the $g_{j\bar{k}}$, $g'_{j\bar{k}} = g_{j\bar{k}} + v_{j\bar{k}}$ metrics respectively. To do this, we shall study $\Delta'[e^{-cv}(n+\Delta v)]$. In principle, this involves derivatives of v up to fourth order. However, the expression

$$\Delta\left[\det(g_{j\bar{k}} + v_{j\bar{k}})\right]$$

contains the same fourth-order terms yet equals $\Delta[e^{(n+1)v}e^F]$ by (M-A3). Thus, we may eliminate the fourth-order terms from $\Delta'[e^{-cv}(n+\Delta v)]$. After picking coordinates to diagonalize $g_{j\bar{k}}$ and $g'_{j\bar{k}}$, we can compute that

$$\begin{aligned} \Delta'[e^{-cv}(n+\Delta v)] \geqslant{}& e^{-cv}\Big\{(n+1)\Delta v + \Delta F - n^2 \inf_{i\neq l} R_{i\bar{i}l\bar{l}}\Big\} \\ &- c'e^{-cv}(n+\Delta v) \\ &+ \Big(c + \inf_{i\neq l} R_{i\bar{i}l\bar{l}}\Big) e^{-cv}e^{c'F}(n+\Delta v)^{n/(n-1)}, \end{aligned} \tag{11}$$

where $R_{i\bar{j}k\bar{l}}$ is the curvature associated to $g_{j\bar{k}}$. Although the derivation of (11) is messy, it is elementary. In fact, once $\Delta[e^{-cv}(n+\Delta v)]$ is expressed in terms of derivatives of v up to third order, as explained above, we may simply regard the derivatives v, v_j, $v_{j\bar{k}}$, v_{jk} as independent variables, and then (11) becomes a messy elementary inequality for numbers. This involves algebra, but no calculus.

Now (11) and the maximum principle show that $e^{-cv}(n+\Delta v) \leqslant C_3$. For otherwise, at a suitable sequence x_i, the left side of (11) will have a negative lim inf, while the right side will remain large and positive. Here we pick $c \geqslant 1 + \|R_{i\bar{j}k\bar{l}}\|$; recall that we already have an a priori bound on $|v|$, and note that the $(n+\Delta v)^{n/(n-1)}$ term will dominate all the other terms on the right. So we know that $n + \Delta v \leqslant C$ a priori. Recalling that $0 \leqslant (g_{j\bar{k}} + v_{j\bar{k}})$, we obtain at once $\|v_{j\bar{k}}\| \leqslant C$, the norm taken in $g_{j\bar{k}}$. In view of (M-A3) we obtain also a lower bound on the eigenvalues of $(g_{j\bar{k}} + v_{j\bar{k}})$.So we know a priori the following:

$$|v| \leqslant C, \qquad c(g_{jk}) \leqslant (g_{j\bar{k}} + v_{j\bar{k}}) \leqslant C(g_{j\bar{k}}). \tag{12}$$

These estimates in a ball of radius 1 in $(g_{j\bar{k}})$ yield also $\|v'\| \leqslant C$ in a ball of radius $1/2$, the norm being taken in $(g_{j\bar{k}})$. (This follows from elliptic theory.) Consequently, we have also $\|v_j\| \leqslant C$. Observe that we have not yet estimated derivatives of the type v_{jk} or $v_{\bar{j}\bar{k}}$.

Next we estimate the mixed third derivatives of v, using the quantity

$$S = \sum g'^{i\bar{r}}g'^{j\bar{s}}g'^{k\bar{t}}v_{,i\bar{j}k}\,v_{,r\bar{s}t}. \tag{13}$$

Here a comma denotes covariant differentiation in the g' metric. The analogue of S for real Monge-Ampère equations was introduced by Calabi. One calculates $\Delta' S$ and then applies the maximum principle to S. In principle, $\Delta' S$ involves derivatives of v up to order 5. However, by differentiating (M-A3) three times, we can re-express $\Delta' S$ in terms of derivatives of v up to order 4. Remarkably, the fourth-order derivatives now enter $\Delta' S$ as a sum of squares, and it follows that

(14) $$\Delta' S \geqslant c_1 S - c_2, \qquad c_1 \text{ and } c_2 \text{ positive constants.}$$

To make the estimate, one uses (12) and (13) to bound junk terms. The algebra is complicated, but again one needs no further calculus once the fifth derivatives have been eliminated from $\Delta' S$. Now (14) and the maximum principle show at once that S is bounded a priori, $S \leqslant C_4$. In view of (12), (13) we now know that $\|v_{,j\bar{k}l}\| \leqslant C_5$, where the norm and covariant derivatives are now taken in terms of $(g_{j\bar{k}})$.

So we have estimates for v, v', and the terms of mixed type in v'', v'''. Finally, we can use the elliptic theory of Chapter 3 in each ball of radius one in $(g_{j\bar{k}})$, to conclude that $v \in \mathrm{Lip}(3 - \varepsilon)$ with a priori bounds. So we can use Schauder theory (note that the equation is elliptic because of the lower bound for $(g_{j\bar{k}} + v_{j\bar{k}})$ in (12)) to conclude that $v \in C^\infty(D)$ with a priori bounds $\|v_{,j\bar{k}\cdots p}\| \leqslant C$ on each of its covariant derivatives. As usual the norm and the covariant differentiation are in terms of $(g_{j\bar{k}})$. This concludes the discussion of boundary regularity.

We now know enough estimates to solve (M-A3) by the continuity method, and to prove interior regularity. But so far we know little about the boundary behavior of u because our estimates are defined in terms of a metric which degenerates at the boundary.

3. Proof of boundary regularity. Recall that v solves (M-A3) with $F = O(\psi^{n+1})$, ψ a defining function for D. In this section, we shall show that

(15) $$\|v_{,j\bar{k}\cdots p}\| \leqslant C_\varepsilon \psi^{n+(3/2)-\varepsilon}.$$

Because of relations such as $|v_j| \leqslant C\|v_{,j}\|/\psi$ ($|\cdot| =$ Euclidean length, $\|\cdot\| =$ length in $g_{j\bar{k}}$), the estimate (15) implies $v \in C^{n+(3/2)-\varepsilon}(\bar{D})$, which is the boundary regularity asserted by the theorem. We may assume that $-\psi$ is strictly plurisubharmonic.

Consider now the linearized form of (M-A3):

$$(n+1)v + F = \int_0^1 \frac{d}{ds}\left[\log\det(g_{j\bar{k}} + s v_{j\bar{k}})\right] ds = \sum_{pq}\left(\int_0^1 A^{p\bar{q}}\, ds\right) v_{p\bar{q}},$$

where $(A^{p\bar{q}})$ is the inverse of $(g_{j\bar{k}} + s v_{j\bar{k}})$. From our earlier estimates we know that as matrices $(A^{p\bar{q}}) \sim (g^{jk})$, so the linearized equation takes the form

$$\Delta_A v - (n+1)v = F.$$

This is a well-behaved elliptic equation on each ball of unit radius with respect to $g_{j\bar{k}}$. Therefore all the estimates (15) for derivatives of v follow by standard elliptic theory if we can prove

$$|v| \leqslant C_\varepsilon \psi^{n+(3/2)-\varepsilon}. \tag{16}$$

(If a harmonic function is bounded on a ball, then all derivatives are bounded on the inner half of the ball.) So the Cheng-Yau theorem is reduced to proving (16).

LEMMA 1. $|v| \leqslant C\psi$.

PROOF. Since $-\psi$ is plurisubharmonic, $-\Delta_A \psi > 0$, and so

$$\Delta_A(v - C\psi) - (n+1)(v - C\psi) \geqslant 0.$$

By the maximum principle we can pick $\{z_i\}$ so that

$$(v - C\psi)(z_i) \to \sup_D (v - C\psi), \qquad \limsup_{i\to\infty} \Delta_A(v - C\psi)(z_i) \leqslant 0.$$

Comparing the last two inequalities gives $v - C\psi \leqslant 0$, i.e., $v \leqslant C\psi$. Repeating the argument with $v - C\psi$ replaced by $-v - C\psi$ yields $-v \leqslant C\psi$, so the lemma is proved. $\square$

LEMMA 2. $n + \Delta v = n + O(\psi)$, *where* Δ *is defined using the* $g_{j\bar{k}}$ *metric.*

PROOF. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $(v_{j\bar{k}})$ with respect to $(g_{j\bar{k}})$. Using (M-A3) and Lemma 1 we have

$$\prod_1^n (1 + \lambda_k) = e^{(n+1)v} e^F = 1 + O(\psi).$$

Since $(1/n)\Sigma_1^n(1 + \lambda_k) \geqslant \Pi_1^n(1 + \lambda_k)^{1/n}$, we have

$$(n + \Delta v) \geqslant n + O(\psi), \tag{17}$$

establishing half of the desired estimate. Consider now (11) with $c = 1$,

$$\Delta'[e^{-v}(n + \Delta v)] \geqslant e^{-v}(n + \Delta v) - ne^{-v} - e^{-v}\Delta F - n^2\Big(1 + \inf_{i\neq l} R_{i\bar{i}l\bar{l}}\Big)e^{-v}$$
$$+ \Big(1 + \inf_{i\neq l} R_{i\bar{i}l\bar{l}}\Big)e^{-v}\exp\left(\frac{-(n+1)v + F}{n-1}\right)(n + \Delta u)^{n/(n-1)}.$$

Near the boundary, we expect $(g_{j\bar{k}})$ to look like the Poincaré metric for the ball, for which $\inf_{i\neq l} R_{i\bar{i}l\bar{l}} = -1$. In fact, a computation gives $(1 + \inf_{i\neq l} R_{i\bar{i}l\bar{l}}) = O(\psi)$. Therefore the last two terms in the previous inequality are $O(\psi)$; here we make use of (12) from §2. So we have

$$\Delta'[e^{-v}(n + \Delta v)] \geqslant e^{-v}(n + \Delta v) - ne^{-v} - C\psi, \tag{18}$$

and now (17) yields

$$\Delta'[e^{-v}(n + \Delta v)] \geqslant -C'\psi.$$

Since $-\psi$ is strictly plurisubharmonic, we have $-\Delta'\psi \approx -\Delta\psi \geqslant c\psi$ in view of the way $(g_{j\bar{k}})$ degenerates. Therefore $\Delta'[e^{-v}(n + \Delta v) - C''\psi] \geqslant 0$ for large C''. We conclude that $e^{-v}(n + \Delta v) - C''\psi$ cannot take its maximum in the interior of D. Now apply the maximum principle to $e^{-v}(n + \Delta v) - C''\psi$. For a sequence $\{z_i\}$ which we now know approaches the boundary, $e^{-v}(n + \Delta v) - C''\psi$ approaches its sup, while

$$\limsup_{i\to\infty} \Delta'\{e^{-v}(n+\Delta v) - C''\psi\} \leqslant 0.$$

Since

$$\Delta'\{e^{-v}(n+\Delta v) - C''\psi\} \geqslant \{e^{-v}(n+\Delta v) - C''\psi\} - ne^{-v} + O(\psi),$$

by (18), it follows that $0 \geqslant \sup_D\{e^{-v}(n+\Delta v) - C''\psi\} - n$. (Here we used Lemma 1 and the fact that $\{z_i\} \to \partial D$.) Another application of Lemma 1 yields $n \geqslant (n + \Delta v) - C'''\psi$ in D, which together with (17) proves Lemma 2. □

Lemma 2 immediately gives an estimate of the mixed second derivatives $v_{i\bar{j}}$. In fact, retaining the notation of the proof of Lemma 2, we saw that $\prod_1^n(1+\lambda_k) = 1 + O(\psi)$, while Lemma 2 says that $(1/n)\Sigma_1^n(1+\lambda_k) = 1 + O(\psi)$. Thus the numbers $(1+\lambda_k)$ have practically the same arithmetic and geometric mean, so they must be almost equal. That is, $1 + \lambda_k = 1 + O(\psi)$, so

$$\|v_{,j\bar{k}}\| = O(\psi), \tag{19}$$

where the covariant derivatives and norm are taken in terms of $(g_{j\bar{k}})$. Together with Lemma 1, this yields

$$\|v'\| = O(\psi), \tag{20}$$

where the norm is taken in terms of $(g_{j\bar{k}})$. We get this using elliptic theory on a ball of radius one in $(g_{j\bar{k}})$.

LEMMA 3. *If* $\|v'\|^2 = \Sigma_{jk}\, g^{j\bar{k}} v_j \bar{v}_k$ *is the square of the norm of the gradient of* v *in the* $(g_{j\bar{k}})$*-metric, then*

$$\|v'\|^2 \leqslant C_\varepsilon \psi^{2n+1-\varepsilon}.$$

PROOF. A long calculation using (19) shows that

$$\left\{\Delta' \frac{\|v'\|^2}{\psi^\beta} - C'\psi^\alpha\right\} \geqslant C_{n,\beta}\left\{\frac{\|v'\|^2}{\psi^\beta} - C'\psi^\alpha\right\} + \tilde{C}_{n,\alpha}\psi^\alpha - C\|F'\|^2\psi^{-1-\beta} - C\psi^{4n},$$

where, for certain ranges of α and β, the constants $C_{n,\beta}$ and $\tilde{C}_{n,\alpha}$ are positive and $\tilde{C}_{n,\alpha}\psi^\alpha$ swamps the last two terms near ∂D. We claim that for suitable α and β,

$$\|v'\|^2 \leqslant C\psi^\beta \text{ implies } \|v'\|^2 \leqslant C'\psi^{\alpha+\beta}. \tag{21}$$

To see this, consider the "annulus" $\{0 \leqslant \psi(z) \leqslant \varepsilon\}$ and apply the maximum principle to the function $\{(\|v'\|^2)/\psi^\beta - C'\psi^\alpha\}$, where C' is large, depending on ε. If the maximum occurs on the inner boundary $\{\psi(z) = \varepsilon\}$, then that minimum is negative, so (21) is obvious. Otherwise, the maximum principle gives us a sequence on which $\{(\|v'\|^2)/\psi^\beta - C'\psi^\alpha\}$ approaches its sup over the annulus, while

$$0 \geqslant C_{n\beta}\left\{\left(\|v'\|^2\right)/\psi^\beta - C'\psi^\alpha\right\} + \left\{\tilde{C}_{n\alpha}\psi^\alpha - C\|F'\|^2\psi^{-1-\beta} - C\psi^{4n}\right\},$$

with the second term in braces positive. Consequently, $(\|v'\|^2)/\psi^\beta - C'\psi^\alpha \leqslant 0$ is the annulus $\{0 \leqslant \psi \leqslant \varepsilon\}$, so again (21) holds. Thus for suitable α, β estimate (21) is now proved in all cases.

Now starting with (20) and using (21) repeatedly, we can prove $\|v'\|^2 \leqslant C_s\psi^s$ for ever larger s, until at last we arrive at the conclusion of Lemma 3. □

We now are ready to finish the proof of the Cheng-Yau theorem by verifying (16). By Lemmas 1 and 3,

$$\|v'\| \leqslant C_\varepsilon \psi^{n+(1/2)-\varepsilon}, \qquad |v| \leqslant C\psi;$$

therefore we may recover $v(z)$ by integrating v' over a path from z to ∂D, and the resulting estimate is (16). The Cheng-Yau theorem is proved. □

Further work on the complex Monge-Ampère equations (M-A1)–(M-A3) has been carried out recently by Lee and Melrose **[44]**.

THEOREM (LEE-MELROSE **[44]**). *The solution u of* (M-A2) *has the asymptotic expansion*

$$(22)\quad u \sim \psi \sum_{k \geqslant 0} \eta_k(\psi^{n+1}\log\psi)^k \text{ near } \partial D = \{\psi = 0\} \subseteq \mathbf{C}^n \quad \text{with } \eta_k \in C^\infty(\overline{D}).$$

This means that for large N, $u - \psi\sum_{k=0}^{N}\eta_k(\psi^{n+1}\log\psi)^k$ has many continuous derivatives on $\overline{D}$ and vanishes to high order at ∂D.

We include only the most superficial sketch of the proof. The idea is reminiscent of the Schauder theory in Chapter 3. Starting with partial information on u (e.g. the Cheng-Yau theorem), we write down a linearized equation and hope to get out more than we put in by making a careful study of linear PDE with less-than-perfect coefficients. Then we use the new information on u to improve the coefficients of the linearized PDE and repeat the same argument over and over. It is easy enough to write down a linearized equation. In (M-A3), we just expand the determinant by minors to obtain

$$(23)\qquad L_v = \sum_{jk} a_{j\bar{k}} \frac{\partial^2 v}{\partial z_j \partial \bar{z}_k} = g \quad \text{where the } (a_{j\bar{k}}) \text{ are the minors.}$$

Of course $a_{j\bar{k}}$ and g depend on v.

The work comes in understanding (23). We have to assume u is given to high order by (22) near ∂D, then calculate the $(a_{j\bar{k}})$. Basically, L looks like the Laplacian in the Poincaré metric on the ball. In particular, L is elliptic in D but degenerates at the boundary. The heart of the proof is to understand how solutions of such degenerate elliptic linear equations behave near the boundary. See also R. Graham **[29]** for an analysis of (23) from a Heisenberg-group point of view.

4. Chains. We begin this section by showing how to use the Poincaré metric to get local invariants and distinguished curves on the boundary ∂D of a strictly pseudoconvex domain D. We know from the formal study of the first section that the volume form of the Poincarè metric on D is given by

$$d\mathrm{vol} = c_n \frac{dz_1 \wedge \cdots \wedge dz_n \wedge d\bar{z}_1 \wedge \cdots \wedge d\bar{z}_n}{(u(z))^{n+1}},$$

where $u(z)$ vanishes to exactly first order at ∂D. The procedure in the first section gives us the Taylor expansion of u to order $(n+1)$ at ∂D.

EXAMPLE. For the unit ball $D = \{1 - |z|^2 > 0\}$ we have $u(z) = 1 - |z|^2$ as the solution to (M-A2). If we introduce projective coordinates $\xi_0 = z_0$, $\xi_k/\xi_0 = z_k$ $(1 \leqslant k \leqslant n)$ on $\mathbf{C}^1 \times D$, then we can define the function

$$U(z_0, z_1, \ldots, z_n) = |z_0|^2 u(z) = |\xi_0|^2 - \sum_1^n |\xi_k|^2,$$

which gives us an indefinite metric

$$ds^2 = \sum_{j,k \geqslant 0} \frac{\partial^2 U}{\partial z_j \partial \bar{z}_k} dz_j d\bar{z}_k = |d\xi_0|^2 - \sum_1^n |d\xi_k|^2.$$

The linear fractional transformations of the ball arise as isometries of this metric.

In the general case we can follow the same procedure by introducing the function

$$U(z_0, z_1, \ldots, z_n) = |z_0|^2 u(z), \qquad (z_0, z_1, \ldots, z_n) \in (\mathbf{C} \setminus \{0\}) \times D \tag{24}$$

and defining a metric

$$ds^2 = \sum_{jk \geqslant 0} \frac{\partial^2 U}{\partial z_j \partial \bar{z}_k} dz_j d\bar{z}_k, \tag{25}$$

which turns out to be nondegenerate.

Now (24), (25) are invariantly defined. More precisely, every biholomorphic map $\Phi\colon D \to \tilde{D}$ lifts to a biholomorphic map $\Phi^{\#}\colon (\mathbf{C} \setminus \{0\}) \times D \to (\mathbf{C} \setminus \{0\}) \times \tilde{D}$ which preserves (24), (25). The rule for lifting is

$$\Phi^{\#}\colon (z_0, z) \to (\tilde{z}_0, \tilde{z}) = \left(z_0[\det \Phi'(z)]^{1/(n+1)}, \Phi(z)\right),$$

and the verification that (24), (25) are preserved by $\Phi^{\#}$ is trivial.

Consequently, we have a list of tensors invariantly attached to points of $(\mathbf{C} \setminus \{0\}) \times D$, namely the curvature for (25) and its covariant derivatives. Although we cannot compute (25) inside the domain, we do know its Taylor series to high order at ∂D. So the tensors $R_{j\bar{k}l\bar{m},p,\ldots,\bar{q}}$ can be computed on $(\mathbf{C} \setminus \{0\}) \times \partial D$, as long as the defining formulas do not involve too many differentiations of the metric $(g_{j\bar{k}})$.

EXAMPLE. $\| R_{j\bar{k}l\bar{m}} \|^2$ is a scalar invariant attached to points of $(\mathbf{C} \setminus \{0\}) \times \partial D$; it can be computed explicitly in terms of the Taylor expansion of ∂D. If D is in Moser's normal form,

$$D = \left\{\operatorname{Im} z_1 > |z'|^2 + \sum A_{p\bar{q}r\bar{s}} z_p \bar{z}_q z_r \bar{z}_s + \text{higher order terms}\right\},$$

then at $(z_0, z_1, \ldots, z_n) = (z_0, 0, \ldots, 0) \in (\mathbf{C} \setminus \{0\}) \times \partial D$ we can calculate that

$$\|R_{j\bar{k}l\bar{m}}\|^2 = c_n |z_0|^4 \sum |A_{p\bar{q}r\bar{s}}|^2.$$

More generally, we can write down scalar invariants

$$(26) \quad \Omega = \operatorname{Trace}\left(R_{a\bar{b}c\bar{d},p\bar{q},\ldots,r} \otimes \cdots \otimes R_{a'\bar{b}'c'\bar{d}',p'q',\ldots,\bar{r}'}\right) = |z^0|^{2s} P\left(A^l_{\alpha\bar{\beta}}\right),$$

for $(z_0, z_1, \ldots, z_n) = (z_0, 0)$ and ∂D in normal form. Here s is a positive integer and P is a polynomial in the coefficients $A^l_{\alpha\bar{\beta}}$ in Moser's normal form. Ratios of two Ω's with the same "weight" s will be scalar invariants attached to a boundary point. More generally, if $\Omega_1, \ldots, \Omega_m$ are finitely many expressions of the form (26) with weights $s_1, \ldots, s_m$, then to each boundary point $z \in \partial D$ we pick any z^0 and set $F(z) = (\Omega_1(z^0, z), \ldots, \Omega_m(z^0, z))$. $F(z)$ will be a vector in $\mathbf{C}^m$; if we identify $(\xi_1, \ldots, \xi_m) \in \mathbf{C}^m$ with $(\lambda^{s_1}\xi_1, \ldots, \lambda^{s_m}\xi_m)$ for $\lambda > 0$, then $F(z)$ is an invariant taking values in the quotient space $\mathbf{C}^m/\mathbf{R}^\times$.

Returning to (26), the invariance of Ω means that $P(A^l_{\alpha\bar{\beta}})$ has a transformation law under biholomorphic maps. If $\phi\colon \partial D \to \partial\tilde{D}$ is biholomorphic, where

$$\partial D = \left\{\operatorname{Im} z_1 = |z'|^2 + \sum A^l_{\alpha\bar{\beta}} (\operatorname{Re} z_1)^l z'^\alpha \bar{z}'^\beta\right\}$$

and

$$\partial \tilde{D} = \left\{\operatorname{Im} z_1 = |z'|^2 + \sum B^l_{\alpha\bar{\beta}} (\operatorname{Re} z_1)^l z'^\alpha \bar{z}'^\beta\right\}$$

are in normal form, then

$$(27) \qquad P\left(A^l_{\alpha\bar{\beta}}\right) = |\det \Phi'(0)|^{2s/(n+1)} P\left(B^l_{\alpha\bar{\beta}}\right).$$

CONJECTURE. Every polynomial satisfying (27) arises from some Ω as in (26).

This problem seems natural when we try to attach boundary invariants to domains. It is also what is needed for a precise asymptotic expansion for the Bergman kernel; see Chapter 12.

In fact we have to be careful in stating the conjecture, since $R_{j\bar{k}l\bar{m},\ldots,\bar{q}}\,\bar{q}$ of very high rank cannot be computed locally. One finds that the Ω of (26) can be computed locally basically if $s \leqslant n$. So we really have a prescription to write down polynomials satisfying (27) only when $s < n$. For general s we cannot even write down nontrivial examples, let alone classify all possibilities. The algebra in Chapter 12 proves our conjecture when $s \leqslant n - 20$.

We have seen how to attach invariant tensors and scalars to a boundary point using (M-A2). Now we shall see how to find distinguished curves. We use polar coordinates $z_0 = \rho e^{i\theta}$ so that $(\mathbf{C} \setminus \{0\}) \times D \simeq \mathbf{R}^+ \times S^1 \times \partial D$. Let us restrict the indefinite metric (24) to $\mathbf{R}^+ \times S^1 \times \partial D$; this is natural since (24) cannot be computed in the interior. The result after some small calculation is

$$(28) \qquad \partial\bar{\partial} U\big|_{\mathbf{R}^+ \times S^1 \times \partial D} = \rho^2 Q$$

with

$$Q = \frac{1}{(n+1)}\left(\frac{\partial u - \bar{\partial} u}{i}\right) d\theta + \sum_{j,k=1}^{n} \frac{\partial^2 u}{\partial z_j \partial \bar{z}_k} dz_j d\bar{z}_k. \tag{29}$$

One checks that Q is a nondegenerate Lorentz metric on $S^1 \times \partial D$. Now the left-hand side of (28) is a biholomorphic invariant, and ρ is a scalar factor. Therefore the conformal class of Q is a biholomorphic invariant. In other words, the following is true: Let $\Phi: D \to \tilde{D}$ be a biholomorphic map and let Q, $\tilde{Q}$ be the Lorentz metrics defined on $S^1 \times \partial D$, $S^1 \times \partial \tilde{D}$ by (29). Then Φ lifts to a map $\Phi^{\#}$: $S^1 \times \partial D \to S^1 \times \partial \tilde{D}$ which carries Q to another Lorentz metric that differs from $\tilde{Q}$ by a scalar factor.

Using the Lorentz metric (29), we can easily find distinguished curves in $S^1 \times \partial D$.

PROPOSITION 1. *Two conformally equivalent Lorentz metrics have the same light rays.*

PROOF. Recall that a light ray of a Lorentz metric $ds^2 = \sum g_{jk} dx_j dx_k$ is just a Hamiltonian path for the Hamiltonian $H = \frac{1}{2}\sum g^{jk}(x) p_j p_k$ satisfying $H = 0$, where p_j and x_j are conjugate variables. Under a conformal change of metric $\widetilde{ds}^2 = \rho(x) \sum g_{jk}(x) dx_j dx_k$, the new Hamiltonian becomes $\tilde{H} = \rho^{-1}(x) H$. We already saw in Chapter 2 that multiplying a Hamiltonian H by a nonvanishing factor changes the zero-energy Hamiltonian trajectories only by a change of clock. Therefore the unparametrized light rays of ds^2, $\widetilde{ds}^2$ are the same. □

Because of this proposition, we can take the light rays as distinguished curves on $S^1 \times \partial D$. If we define *chains* as the image of light rays under the projection $S^1 \times \partial D \to \partial D$, then these too become biholomorphic invariants.

Note that the calculation (29) requires that we know u to second order at ∂D, i.e., we need u up to errors $O(\psi^3)$. Recall that the formal procedure of §1 defined inductively

$$u_1 = [J(\psi)]^{-1/(n+1)} \psi, \qquad u_2 = u_1 \cdot [1 + c_n(1 - J(u_1))],$$

which involves derivatives of ψ up to order 4. Since $u = u_2 + O(\psi^3)$, we can use u_2 in place of u to calculate Q (29) and find the chains.

Consider the Hamiltonian $H(x_j, \Theta; p_j, p_\Theta)$ for light rays on $\partial D \times S^1$, where (x_j, Θ) are coordinates on $\partial D \times S^1$ and (p_j, p_Θ) are the conjugate variables. Since θ does not appear explicitly in (29), H is independent of Θ, and it follows at once from Hamilton's equations that p_Θ remains constant along a light ray. Consequently, the chains in ∂D are Hamiltonian trajectories for

$$H = \sum_{ij} a^{ij}(x) p_i p_j + \sum_j b^j(x) p_\Theta p_j + V(x) p_\Theta^2. \tag{30}$$

Here $(a^{ij}(x))$ is positive semidefinite with a one-dimensional nullspace, $p_\Theta =$ arbitrary constant, and the chains are precisely the trajectories on which $H = 0$.

In principle, H can be computed. Using H we can get global information about chains.

EXAMPLE 1. The chains are well behaved globally on a boundary of the form

$$\partial D = \{\operatorname{Im}(z_1) = \phi(z')\}.$$

The point is that ∂D is translation-invariant in $\operatorname{Re} z_1$. Therefore, using coordinates $x_1, \ldots, x_{2n-1}$ on ∂D with $x_1 = \operatorname{Re} z_1$, we obtain the chains from a Hamiltonian (30) with no x_1-dependence. The conjugate momentum p_1 is therefore conserved along a chain. From $p_1 = \text{const.}$, $H = 0$ and the form (30) of H, one can easily bound all the momenta, $|p_j| \leqslant \text{const.}$ (This amounts to checking that $\xi_1 \neq 0$, where $(\xi_1, \ldots, \xi_{2n-1})$ generates the nullspace of $a^{ij}(x)$ in (30).) Since the Hamiltonian vector field blows up only as $p \to \infty$, the trajectories are globally well behaved. See Example 4 to find what can go wrong if ∂D isn't translation-invariant.

EXAMPLE 2. If $\partial D = \{\operatorname{Im} z_1 = \phi(|z'|)\}$, then the chains on ∂D can be completely understood. In fact the ordinary differential equations for chains may be uncoupled and reduced to the calculation of integrals. To see this, use $y = \operatorname{Re} z_1$, $r = |z_1|$, $w = z'/|z'|$ as coordinates in ∂D. All the coordinates except r are cyclic for the Hamiltonian (30) so the problem is reduced to Hamilton's equations in one-space variable, which is trivial.

EXAMPLE 3 (BURNS-SCHNEIDER [7]). Take the hyperquadric $Q = \{\operatorname{Im} z_1 = |z'|^2\}$ with the origin removed, and identify (z_1, z') with $(\lambda^{2k} z_1, \lambda^k z')$ for all integers $-\infty < k < +\infty$ and a fixed $\lambda > 1$. In this way we obtain a C-R manifold M which looks locally like Q. It is easy to see that the two points (z_1, z') and $(-\bar{z}_1, -z')$ cannot be connected by a chain on M, simply because the chain which joins them in Q passes through the origin. M can be identified with the boundary of a domain by the mapping $(z_1, z') \to (z_1^{(\pi i)/(\ln \lambda)}, z_1^{-(1/2)} z')$; the powers of z_1 make sense, since $\operatorname{Im} z_1 > 0$ on M.

So we see that the analogy between geodesics and chains cannot be carried too far.

EXAMPLE 4. Chains may also have local pathologies. For instance, on $\partial D = \{\operatorname{Im} z_1 = |z'|^2 + (\operatorname{Re} z_1)|z'|^8\}$ there are chains which spiral in towards the origin. Note that $\partial D \subseteq \mathbf{C}^2$ is the simplest domain in normal form not covered by Example 1. To understand how the chains look, we introduce coordinates $y = \operatorname{Re} z_1$, $re^{i\theta} = z_2$ on ∂D, and then switch to new variables $\theta, s, v, p_\theta, p_s, p_v$ by means of the canonical transformation

$$\begin{aligned} &r = p_v^{-1/2} s, \qquad y = -\tfrac{1}{2} p_v^{-1} p_s s + v, \qquad \theta \text{ unchanged}, \\ &p_r = p_v^{1/2} p_s, \qquad p_y = p_v, \qquad p_\theta \text{ unchanged}. \end{aligned} \tag{31}$$

This is somewhat analogous to Sundmann's regularization of the 3-body problem in Chapter 2. Calculating the Hamiltonian (30) and composing with (31), we find that

$$(32)\qquad H = \left[p_s^2 + s^2 + \frac{p_\theta^2}{s^2} + (2p_\theta - 6p_\Theta)\right] + vp_v^{-3}\left[16s^8 + 16p_\theta s^6 + \left(192p_\theta p_\Theta - 432p_\Theta^2\right)s^4\right] + O\left(p_v^{-4}\right).$$

(To arrive at (32), we multiplied H by a nonvanishing factor p_v^{-1}; this of course has no effect on the unparametrized chains.)

What is important here is the behavior of orbits with p_v large. One checks by starting at (31) that a spiral chain on ∂D corresponds to a Hamiltonian path for (32) on which $H = 0$, s and p_s remain bounded, v tends to zero, and p_v tends to infinity. Now for large p_v, the first term in brackets gives the main contribution to H in (32); the rest is a small perturbation. So, approximately,

$$H \approx p_s^2 + s^2 + p_\theta^2/s^2 + (2p_\theta - 6p_\Theta).$$

(The formula would be exact if ∂D were the hyperquadric.) This Hamiltonian is trivial to understand: v, p_v, p_θ, p_Θ remain constant, while s and p_s undergo a simple periodic motion corresponding to a particle in a one-dimensional potential well. Adding in the small corrections in (32) to obtain the exact Hamiltonian, we can perform a first-order perturbation expansion to see more accurately how v, p_v, etc. really behave. One finds that p_v, rather than remaining constant, increases to infinity very slowly for large time, while s and p_s remain bounded. On most orbits $v \to \infty$, but by picking the right initial conditions one gets $v \to 0$. So spiral chains exist on ∂D. No one knows how prevalent spiral chains are for generic domains.

For more details of the calculations in this section see **[22]**.

CHAPTER 12. THE BERGMAN KERNEL

1. The asymptotic expansion. In this section we will briefly discuss some in the ideas in the proof of

THEOREM 1. *If $D = \{\psi(z) > 0\} \subseteq \mathbf{C}^n$ is a strongly pseudoconvex domain with $\psi' \neq 0$ on ∂D, then the Bergman kernel $K(z, w)$ may be written on the diagonal in the form*

$$(1)\qquad K(z, z) = \phi(z)/\psi^{n+1}(z) + \tilde{\phi}(z)\log\psi(z)$$

where ϕ, $\tilde{\phi} \in C^\infty(\overline{D})$ and $\phi\,|_{\partial D} \neq 0$. Off the diagonal, we may extend ψ, ϕ, $\tilde{\phi}$ to functions $\phi(z, w)$, $\tilde{\phi}(z, w)$ with the properties

1. *$\psi(z, w)$ is almost analytic in z, $\overline{w}$ in the sense that $\bar{\partial}_z\psi(z, w)$ and $\partial_w\psi(z, w)$ vanish to infinite order at $z = w$.*
2. $\psi(z, z) = \psi(z)$.
3. *Similarly for $\phi(z, w)$ and $\tilde{\phi}(z, w)$.*

The Bergman kernel has the form

$$K(z,w) = \phi(z,w)/\psi^{n+1}(z,w) + \tilde{\phi}(z,w)\log\psi(z,w)$$

for $(z,w) \in \Omega_\varepsilon = \{|z-w| < \varepsilon,\ \mathrm{dist}(z,\partial D) < \varepsilon\}$ *and* $\varepsilon > 0$ *small. Outside* Ω_ε, $K(z,w)$ *is smooth up to the boundary as a function of* (z,w). *There is an analogous expansion for the Szegö kernel.*

We give here a brief sketch of the original proof in **[21]** and of the alternate proof in **[5]**. Later we shall look more closely at the Bergman kernel by trying to compute bthe functions ϕ and $\tilde{\phi}$. The proof of Boutet-Sjöstrand **[5]** for the Szegö kernel is roughly as follows: fix $\partial D \subseteq \mathbf{C}^n$ and define the operator $T_{\partial D}$: $L^2(\partial D) \to L^2(\partial D)$ by $f \mapsto \bar{\partial}_b f$. The Szegö kernel gives the orthogonal projection from $L^2(\partial D)$ to the kernel of $T_{\partial D}$. Let $B \subseteq \mathbf{C}^n$ be the unit ball. Then microlocally in $T^*(\partial D)$ there exists a canonical change of coordinates Φ: $T^*(\partial D) \to T^*(\partial B)$ so that the induced Fourier integral operator U: $L^2(\partial D) \to L^2(\partial B)$ satisfies

$$T_{\partial D} = \varepsilon \cdot U^{-1} T_{\partial B} U, \tag{2}$$

where $T_{\partial B}$ is the $\bar{\partial}_b$-operator on the sphere, and ε is an elliptic $(n-1) \times (n-1)$ matrix of pseudodifferential operators of order zero. This reduces the study of the Szegö kernel for ∂D to the known simple case $D =$ unit ball.

To achieve (2), we first pick Φ and the principal symbol of ε so that (2) holds modulo lower-order errors. In view of Ergorov's theorem, this is purely a question about the symplectic geometry of the principal symbol of $T_{\partial D}$. Once (2) holds modulo lower-order errors, we may then add lower-order corrections to ε so that (2) holds modulo a smoothing error. The smoothing error has to be carried along in the proof, but it never causes real trouble, so we shall ignore it.

Now in principle, (2) gives the Szegö projection $K_{\partial D}$ as $U^{-1}K_{\partial B}U$. In practice one has to work to compose these operators explicitly to get a recognizable result. The calculations may be put in the context of Fourier integral operators with complex phase functions. For the ball, $\psi_B(z) = 1 - |z|^2$, $\psi_B(z,w) = 1 - z \cdot \bar{w}$,

$$K_{\partial B}(z,w) = \frac{C_n}{(1 - z\cdot\bar{w})^n} = \tilde{c}_n \int_0^\infty \lambda^{n-1} e^{-\lambda\psi_B(z,w)}\, d\lambda. \tag{3}$$

Thus, $K_{\partial B}$ is a Fourier integral operator with phase function $-\lambda\psi_B(z,w)$. Since U and U^{-1} are also Fourier integral operators, the calculus of operators with complex phase expresses $K_{\partial D}$ as a Fourier integral operator. The result is

$$\begin{aligned} &K_{\partial D}(z,w) = \tilde{c}_n \int_0^\infty a(z,w,\lambda) e^{-\lambda\psi(z,w)}\, d\lambda, \text{ where} \\ &a(z,w,\lambda) \sim \sum_{l \geqslant 0} \lambda^{n-1-l} a_l(z,w) \quad \text{and} \quad a_l(z,w) \in C^\infty(\bar{D} \times \bar{D}), \end{aligned} \tag{4}$$

which of course agrees with (3) for the unit ball.

Now we can simply read off from (4) that

$$\begin{aligned}K_{\partial D}(z, w) &\sim \sum_{l \geq 0}\left(\int_0^\infty \lambda^{n-1-l} e^{-\lambda\psi(z,w)}\, d\lambda\right) a_l(z, w)\\ &= \sum_{\substack{l\geq 0\\ l\neq n}} c_l(\psi(z, w))^{l-n} a_l(z, w) + (\text{const}) \log \psi(z, w) \cdot a_n(z, w)\\ &= \frac{\phi(z, w)}{(\psi(z, w))^n} + \tilde{\phi}(z, w) \log \psi(z, w),\end{aligned}$$

where $\phi = \sum_{l \neq n, l \geq 0} c_l \psi^l a_l$ and $\tilde{\phi} = (\text{const}) a_n$. A calculation of the principal symbol in (4) is enough to show that $\phi(z, z)|_{\partial D} \neq 0$. So the analogue of Theorem 1 for the Szegö kernel is proved.

In fact, the proof as presented in **[5]** is slightly different from our explanation. Rather than reducing $\bar{\partial}_b$ on D to $\bar{\partial}_b$ on the sphere, Boutet-Sjöstrand exhibit Fourier integral operators to reduce both these operators to the same "model" equation. The model is a variant of (57a) of Chapter 4.

To adapt the argument of Boutet-Sjöstrand to the Bergman kernel, one has to deal with the $\bar{\partial}$-Neumann problem instead of the $\bar{\partial}_b$-equation. However, we learned in Chapter 3 how a boundary-value problem on D reduces to a pseudodifferential equation on ∂D. The pseudodifferential equation arising from the $\bar{\partial}$-Neumann problem may be analyzed by the same methods we just explained for $\bar{\partial}_b$, and Theorem 1 for the Bergman kernel again results from a calculation with Fourier-integral operators with complex phase.

The original proof of Theorem 1 **[21]** starts with the following observation: if D_1, D_2 are strictly pseudoconvex domains which share a piece of a common boundary, then on a small neighborhood of a point in the common boundary, the two Bergman kernels differ by a function of (z, w) which is smooth up to the boundary.

This may be easily read off from the smoothness of solutions to the $\bar{\partial}$-Neumann problem in D_2. We form $u = \partial_z^\beta\{K_{D_1}(z, \cdot)\chi_{D_1} - K_{D_2}(z, \cdot)\}$, which is clearly orthogonal to analytic functions on D_2, while $\bar{\partial} u = \alpha$ is supported in $\partial D_1 \setminus \partial D_2$. So u solves a $\bar{\partial}$-Neumann problem, and it follows that u is smooth up to the boundary away from the singular support of α. The argument is essentially due to Kerzman **[37]**, who was also the first to prove the smoothness of the Bergman kernel off the boundary diagonal.

Returning to the proof of Theorem 1, let H_D denote the Hilbert space of analytic functions in $L^2(D)$ and $H_D^\perp$ the space of all distributions f on D satisfying $\int_D f(z)\overline{g(z)}\, dz = 0$ for all $g \in H_D$. Suppose that we have found a decomposition of the Dirac delta function δ_w at w:

$$\delta_w = F_w(z) + G_w(z) + \varepsilon_w(z), \qquad F_w \in H_D, G_w \in H_D^\perp, \tag{5}$$

where ε_w is an appropriately small error. Thinking now of a function $H(z, w)$ on $D \times D$ as an operator $Hf(z) = \int_D H(z, w) f(w)\, dw$ on $L^2(D)$, we can rewrite the last equation as

$$f(z) = (Ff)(z) + (Gf)(z) + (\varepsilon f)(z),$$

where $Ff \in H_D$ and $Gf \in H_D^{\perp}$. Since εf is small, we can interpret $F(z, w)$ as an approximation to the Bergman kernel $K(z, w)$. If for example $\|\varepsilon f\|_{L^2(D)} \leqslant \frac{1}{2}\|f\|_{L^2(D)}$, then we can recover the exact Bergman kernel by iteration, using a sequence of successive approximations.

We now explain how to make the decomposition (5). Take a point $w \in D$ near the boundary, and let w^* be the boundary point closest to w. Now, D is biholomorphically equivalent to the unit sphere to third order at w^*. Therefore we can find a strictly pseudoconvex comparison domain $D_{\text{small}} \subseteq D$ which is tangent to ∂D to third order at w^*, and which locally near w^* is biholomorphic to a piece of the unit ball. By cutting down D_{small}, we may assume that D_{small} is biholomorphic to a domain $B_{\text{small}} \subseteq$ Unit Ball whose boundary agrees with the unit sphere in a neighborhood of the image of w^*. It is easy to write down D_{small} and the biholomorphic map explicitly. Note that $w \in D_{\text{small}}$ if initially w was close to ∂D.

Next we look at the Bergman kernel for D_{small}. We already know the Bergman kernel for the ball, so the observation at the start of the proof shows that in effect we know the Bergman kernel for B_{small}. In view of the explicit biholomorphic mapping of D_{small} to B_{small}, we have the Bergman kernel $K_{D_{\text{small}}}(z, w)$ given as

$$K_{D_{\text{small}}}(z, w) = K_{\text{Explicit}}(z, w) + \tilde{\varepsilon}(z, w),$$

where the error $\tilde{\varepsilon}$ is smooth up to the boundary of D_{small}.

To make the decomposition (5), we now write

$$\begin{aligned} \delta_w &= \left[K_{\text{Explicit}}(\cdot, w)\right] + \left[\delta_w - K_{D_{\text{small}}}(\cdot, w)\chi_{D_{\text{small}}}\right] \\ &\quad + \left[\tilde{\varepsilon}(\cdot, w)\chi_{D_{\text{small}}} - K_{\text{Explicit}}(\cdot, w)\chi_{D - D_{\text{small}}}\right] \\ &\equiv F_w + G_w + \varepsilon_w. \end{aligned}$$

The definition of the Bergman kernel gives $K_{D_{\text{small}}}(\cdot, w)$ only as an analytic function on D_{small}, but inspection of the explicit formula shows at once that F_w continues analytically into all of D. So $F_w \in H_D$. Evidently $G_w \in H_D^{\perp}$ by the reproducing property of $K_{D_{\text{small}}}$. Finally (except for the smoothing error $\tilde{\varepsilon}$) the kernel ε_w is supported in the thin region $D - D_{\text{small}}$, and it follows that ε is a small operator on L^2.

Thus, we have our decomposition (5), and we can read off the Bergman projection K_D by successive approximation: $K_D = \Sigma_{l \geqslant 0} \pm F\varepsilon^l$, where F, G are the operators in (5). It remains to calculate $F\varepsilon^l$, which amounts to doing complicated multiple integrals. We omit the details.

2. Christoffers' Theorem. How accurately can we compute the functions ϕ, $\tilde{\phi}$ in Theorem 1? We may as well work on the diagonal $z = w$, since e.g. $\phi(z, w)$ is given to infinite order about $z = w$ by the expansion

$$\phi(z, w) \sim \Sigma_\alpha (1/\alpha!)(\bar{\partial}^\alpha \phi)(z)\, \overline{(w - z)}^\alpha.$$

Now, if D_1 and D_2 are strictly pseudoconvex domains whose boundaries coincide near a point p, we saw in §1 that their two Bergman kernels $K_{D_1}(z, z)$, $K_{D_2}(z, z)$ differ by a smooth function near p. So in the expansion

$$K_D(z, z) = \frac{\phi(z)}{\psi^{n+1}(z)} + \tilde{\phi}(z)\log\psi(z), \tag{1}$$

we see that ϕ is determined up to nth order at the boundary by local information, while the rest of ϕ is determined globally.

Similarly, $\tilde{\phi}$ is determined to infinite order by local information; again, since $K_D(z, z)$ is determined locally modulo smooth functions on $\bar{D}$. By looking carefully at either of the two proofs of Theorem 1 we can get more: for $p \in \partial D$ and $K_D(z, z)$ given by (1), we have $\partial^\alpha\phi(p)$ ($|\alpha| \leq n$) and $\partial^\beta\tilde{\phi}(p)$ (arbitrary β) determined uniquely by the Taylor series of ∂D at p. For fixed α, β, only finitely many derivatives of ∂D enter into the formula.

Thus, the maps $S\colon \psi \to \phi$ and $T\colon \psi \to \tilde{\phi}$ may be regarded as nonlinear differential operators. It would be interesting to write them down explicitly.

Two basic properties of S and T are immediate from the definition:

(A) If $\eta \neq 0$ on $\{\psi = 0\}$, then $S(\eta\psi) = \eta^{n+1}S(\psi) + O(\psi^{n+1})$ and $T(\eta\psi) = T(\psi) + O(\psi^M)$, for any M.

Indeed, $\eta\psi$ and ψ define the same domain and so lead to the same Bergman kernel.

(B) If Φ is a biholomorphic map, then $S(\psi \circ \Phi) = |\det \Phi'|^{-2}S(\psi) \circ \Phi + O(\psi^{n+1} \circ \Phi)$, while $T(\psi \circ \Phi) = |\det \Phi'|^{-2}T(\psi) \circ \Phi + O(\psi^M \circ \Phi)$ for any M.

These invariance properties are very strong. No one knows how to write down a single nonzero operator T with the required properties. However, we shall see below how to write down all possible operators S satisfying (A) and (B). It appears that among all possible S, the choice which appears in the Bergman kernel is rather typical.

So we now are trying to write down expressions which transform like the Bergman kernel. The first example is $K_{\text{Poincaré}}(z) = C_n/(u(z))^{n+1}$, where u is the formal solution of the Monge-Ampère equation (M-A2) of the preceding chapter. We have already seen that the map which takes ψ to u on $\{\psi \geq 0\}$ is a nonlinear differential operator and that u changes by $O(\psi^{n+1})$ when ψ is multiplied by a nonvanishing factor η.

So we might guess that $\phi/\psi^{n+1} = K_{\text{Poincaré}}$. This is unfortunately false. In fact, suppose that the boundary ∂D is in Moser's normal form,

$$\operatorname{Im} z_n = |z|^2 + \sum_{|\alpha|,|\beta| \geq 2} \sum_{l \geq 0} A^l_{\alpha\bar{\beta}}(\operatorname{Re} z_n)^l z^\alpha \bar{z}^\beta,$$

where $\operatorname{Trace}(A_{22}) = 0$, $\operatorname{Trace}^2(A_{32}) = 0$, $\operatorname{Trace}^3(A_{33}) = 0$. Let us compare the Bergman kernel with $K_{\text{Poincaré}} = C_n/u^{n+1}$ as we approach $0 \in \partial D$ from above. Set $z_t = (0, 0, \ldots, 0, it)$ so that $z_t \in D$ for $t > 0$.

THEOREM 2 (CHRISTOFFERS [12]). *As $t \to 0^+$ we have*

$$K_D(z_t, z_t) = C_n \frac{1}{t^{n+1}} + \frac{\gamma_n}{t^{n-1}} \sum_{pqrs} \left|A^0_{p\bar{q}r\bar{s}}\right|^2 + o(t^{1-n}), \tag{6}$$

$$K_{\text{Poincaré}}(z_t, z_t) = \frac{C_n}{u^{n+1}(z)} = C_n \frac{1}{t^{n+1}} + \frac{\gamma_n'}{t^{n-1}} \sum_{pqrs} \left|A^0_{p\bar{q}r\bar{s}}\right|^2 + O(t^{1-n}), \tag{7}$$

where γ_n, γ_n' are constants depending on the dimension, and $\gamma_n' \neq \gamma_n$.

Thus, $\phi/\psi^{n+1} \neq C_n/u^{n+1}$. The proof of Christoffers' theorem involves hard calculations. In principle we already know how to calculate u, so (7) is just a matter of hard work. To prove (6), we use Theorem 1 to express $K_D(z, w)$ in terms of $\phi(z)$, $\tilde{\phi}(w)$. (Recall, we can express $\phi(z, w)$ in terms of $\phi(z)$ and similarly for $\tilde{\phi}$.) We then write down the reproducing formula $F_\varepsilon(z) \int_D K_D(z, w) F_\varepsilon(w) dw$ for explicit F_ε such as $F_\varepsilon(z) = (z_n + i\varepsilon)^{-(n+k)}$. Letting $\varepsilon \to 0^+$, we obtain from the reproducing formula some information about the derivatives of $\phi(z)$ at the origin. By picking enough F_ε, we get enough information to recover ϕ, ϕ', ϕ'', at the origin, from which we read off (6). The details are very complicated.

Christoffers' calculations give $K_D(z, z)$ up to an error $o(\psi^{-n+1})$ when $z = z_t$ and D is in normal form. By putting Theorem 2 in an invariant form, we can read off $K_D(z, z)$ modulo $o(\psi^{-n+1})$ for general D and z. For the invariant version of Theorem 2, we recall from the preceding chapter the function

$$U(z_0, z_1, \ldots, z_n) = |z_0|^2 u(z_1, \ldots, z_n) \quad \text{on } (\mathbf{C} \setminus \{0\}) \times D$$

and the indefinite metric

$$ds^2 = \sum_{j,k \geq 0} \frac{\partial^2 U}{\partial z_j \partial \bar{z}_k} dz_j\, \overline{dz_k} \quad \text{with curvature } R_{p\bar{q}r\bar{s}}.$$

A calculation shows that $\|R_{p\bar{q}r\bar{s}}\|^2 U^2 = (\text{const}) t^2 \Sigma_{pqrs} |A^0_{p\bar{q}r\bar{s}}|^2$, where the norm on the left is taken in ds^2. So Christoffers' theorem shows that when $z = z_t$ and D is in normal form, the Bergman kernel is given by

$$K_D(z, z) = \frac{C_n}{u^{n+1}} \left\{1 + \gamma_n'' \|R_{p\bar{q}r\bar{s}}\| U^2 + o(\psi^2)\right\}. \tag{8}$$

As the two sides of (8) both transform invariantly, we know that (8) holds for general strictly pseudoconvex domains, since it holds in normal form. After $K_{\text{Poincaré}} = C_n/u^{n+1}$, formula (8) is our second example of a nonlinear differential operator which transforms like the Bergman kernel.

After seeing formula (8) and recalling Weyl's invariant theory for the heat equation, we can make a conjecture on how to calculate the Bergman kernel.

CONJECTURE. Let $R_{p\bar{q}r\bar{s},\sigma\cdots\bar{\tau}}$ be the various covariant derivatives of the curvature in the metric ds^2 as above. Then the Bergman kernel is given by

$$K(z, z) = \frac{C_n}{u^{n+1}}\left[1 + \sum_{s\geq 2} P_s\left(R_{\alpha\bar{\beta}\gamma\bar{\delta},\rho\cdots\bar{\sigma}}\right)U^s\right] \tag{9}$$

$$\text{modulo } C^\infty(\bar{D}) + (\log\psi)\cdot C^\infty(\bar{D}),$$

where P_s is a linear combination of traces of tensor products of the $R_{\alpha\bar{\beta}\gamma\bar{\delta},\rho\cdots\bar{\sigma}}$.

Let us see what this means for the U^2 coefficient, which we would need to compute $K(z, z)$ modulo $o(\psi^{1-n})$. For each expression $\Omega = \mathrm{Trace}[R_{\alpha\bar{\beta}\gamma\bar{\delta},\rho\cdots\bar{\sigma}} \otimes \cdots \otimes R_{p\bar{q}r\bar{s},\mu\cdots\bar{\lambda}}]$ there is natural degree of homogeneity, namely the power of $|z_0|^2$ which appears in it. (Here z_0 is the extra variable in $\mathbf{C}\setminus\{0\}$.) This has to balance the homogeneity of U^2 so that $|z_0|^2$ does not appear in $K(z, z)$. The only invariant Ω with the proper degree to balance U^2 turns out to be $\| R_{p\bar{q}r\bar{s}}\|^2$. So the conjecture implies formula (8), thought it does not predict the value of the constant γ_n''. The conjecture is true—we shall sketch a proof in the next section. Before getting involved in the proof, we should note some important differences between the Bergman kernel and the heat kernel, for which of course the analogue of our conjecture is true.

For the heat kernel, one can simply follow its construction step by step and check explicitly that we never depart from the desired form

$$K_t(x, x) \sim \frac{C_n}{t^{n/2}}\left\{1 + \sum_{s\geq 1} P_s\left(R_{ijkl,p\cdots q}\right)t^s\right\},$$

with P_s linear combinations of traces of tensor products of the $R_{ijkl,p\cdots q}$. If we try the same direct aproach for the Bergman kernel, we come immediately to a serious difficulty. Both the known proofs of Theorem 1 depend strongly on making arbitrary choices; either a canonical transformation Φ in the Boutet-Sjöstrand proof, or an approximating domain D_{small} in the original argument. Therefore, if we look at an intermediate step in the construction of the Bergman kernel, we shall see many irrelevant terms which fail to transform correctly under biholomorphic maps. It is only at the end that the terms all combine in the right way to produce an invariant Bergman kernel. Since the intermediate steps bear little resemblance to the final result, it is hard to imagine how to use the "direct approach" to prove the conjecture on the Bergman kernel. Perhaps one should look for an invariant proof of Theorem 1. The phase function $\lambda\psi(z, w)$ in the Boutet-Sjöstrand proof bears a tantalizing similarity to our $U = |z_0|^2 u(z)$.

The asymptotics of the heat kernel can also be understood by using Weyl's invariant theory for the group $O(n)$. Unfortunately, if we try using invariant theory for the Bergman kernel, then we find that the relevant group is H, the group of linear fractional transformations of the ball which fix a given point on the boundary. This group is not semisimple so Weyl's invariant theory does not apply. In the next section we shall study invariant theory for H, exploiting the fact that H sits inside the semisimple group $U(n, 1)$. Although the algebra is not completely understood, we shall get far enough to prove the conjecture (9).

3. Parabolic invariant theory. In this section we prove some results of invariant theory from **[23]** that can be used to attack the conjecture of the previous section. Let H be a subgroup of semisimple Lie group G. Typically G will consist of matrices preserving an indefinite quadratic form, and H will be the set of matrices in G which fix a given vector e in the light cone. There are several different settings for invariant theory.

Problem 1. Given an action of G on a vector space X, find all G-invariant polynomials on X. This case was discussed in Chapter 5. For our purposes, it is well understood.

Problem 2. Given an action of H on X, find all H-invariant polynomials on X. This problem is too hard.

Problem 3. Given an action of G on X find all H-invariant polynomials on X. This may be reduced back to Problem 1 using a little algebraic geometry, which we shall explain in a moment. The results are due to Weizenbock **[65]**, Seshadri **[56]**, and Hochschild-Mostow **[31]**.

Problem 4. Given an action of G on X and a vector subspace $Y \subseteq X$ which is H-invariant but not G-invariant, find all H-invariant polynomials on Y. This is what we need for the Bergman kernel.

Let us start with an example of Problem 3. Take $G = O(n, m)$ acting on R^{n+m}, fix a vector e in the light cone, and set $H = \{T \in G \mid Te = e\}$. Recall from Chapter 5 that the G-invariant polynomials in the components of vectors $v_1, \ldots, v_N$ are already generated by the obvious examples $\langle v_j, v_k \rangle$. The analogous result for H is as follows.

PROPOSITION 1. *All the H-invariant polynomials in the components of vectors $v_1, \ldots, v_N$ are generated by the obvious examples $\langle v_j, v_k \rangle$ and $\langle v_j, e \rangle$.*

Here, we are in the setting of Problem 3, where X is a direct sum of N copies of R^{n+m} and G acts on X by $T(v_1, \ldots, v_N) = (Tv_1, \ldots, Tv_N)$. To prove the proposition, we introduce the concept of a *normal variety* $V \subseteq \mathbf{C}^n$. The variety V is called normal if whenever a rational function $F = P/Q$ is holomorphic on $V \setminus$ (singular points of V) we can find a polynomial S whose restriction to V agrees with F.

EXAMPLE 1. $V = \{(z, w) \mid z^2 = w^3\}$ is not normal. The points of V are given by $z = t^3$, $w = t^2$; $t = z/w = w^2/z$ is evidently a rational function of (z, w) holomorphic on V away from the singular point $(0, 0)$. However, any polynomial $S(z, w)$ takes the form $S(t^3, t^2)$ on V, and we certainly cannot write t in this form.

EXAMPLE 2. The variety defined by setting a quadratic form $\langle r, r \rangle$ equal to zero is normal. This is a standard result in elementary algebraic geometry.

PROOF OF PROPOSITION 1. Let $P(v^1, \ldots, v^N)$ be an H-invariant polynomial. Using P we shall define a function $F(v^0, v^1, \ldots, v^N)$ for v^0 in the light cone and $v^1, \ldots, v^N$ arbitrary. We first pick a matrix $T_{v^0} \in G$ which carries v^0 back to e,

$(T_{v^0})v^0 = e$. This can be done since G acts transitively on the light cone, but the choice of T_{v^0} is determined only modulo left multiplication by matrices in H. Now set

$$F(v^0, v^1, \ldots, v^N) = P(T_{v^0}v^1, \ldots, T_{v^0}v^N).$$

This is independent of the choice of T_{v^0} since P is H-invariant. We can easily check that F is a G-invariant function defined on

$$V = \{(v^0, v^1, \ldots, v^N) \mid \|v^0\|^2 = 0\} \subset R^{n+m} \oplus X.$$

If F were a polynomial, then by Weyl's invariant theory we would know that F agrees on V with a polynomial generated by $\langle v^j, v^k \rangle$, $\langle v^j, v^0 \rangle$, $\langle v^0, v^0 \rangle = 0$. However, when $v^0 = e$ we can take $T_{v^0} =$ Identity in the definition of F, so that $F(e, v^1, \ldots, v^N) = P(v^1, \ldots, v^N)$ and consequently P is generated by $\langle v^j, v^k \rangle$, $\langle v^j, e \rangle$, proving the proposition. We are now going to show that F is a polynomial.

It is easy to write down explicit matrices $T_{v^0} \in G$ taking v^0 to e, whose entries are of the form (Polynomial in v^0)/$(v_l^0)^{\text{power}}$, where v_l^0 is any component of v^0. From the definition of F we see therefore that

$$F(v^0, v^1, \ldots, v^N) = \frac{P_l(v^0, v^1, \ldots, v^N)}{(v_l^0)^{\text{power}}} \quad \text{on } V \tag{10}$$

for suitable polynomials P_l. While we do not yet know that F is a polynomial, we see at once that F is a rational function with no singularities away from $\{v^0 = 0\}$. The different right-hand sides of (10) agree on V, and it follows by an easy analytic continuation that they still agree on the complex variety

$$V^c = \{(v^0, v^1, \ldots, v^N) \in \mathbf{C}^{(n+m)N} \mid \langle v^0, v^0 \rangle = 0\}.$$

Thus F extends from V to a rational function on V^c, and furthermore this rational function is holomorphic on V^c away from the singular set $\{v^0 = 0\}$. Since V^c is defined by the vanishing of one quadratic form, it is normal and so F agrees on V with a polynomial. The proposition is proved. $\square$

The ideas just explained give a framework for solutions of Problem 4. We suppose that G is a group of linear transformations on a vector space $V = \mathbf{R}^n$ or $\mathbf{C}^n$ while H is the subgroup of those $T \in G$ which fix a vector $e \in V$. An obvious recipe to make H-invariant polynomials $P(y)$ on Y is to start with a G-invariant polynomial $Q(v, x)$ defined on $V \oplus X$ and then set $P(y) = Q(e, y)$. Such a $P(y)$ will be called a *Weyl invariant*. Weyl invariants can be understood by the methods of Chapter 5, since we are dealing with G-invariant polynomials. For instance if $G = O(n, m)$ or $U(n, m)$ and X consists of tensors, then Q must be a linear combination of traces of tensor products. This is just what we need to write down the Bergman kernel. We now have to face the main issue: is every H-invariant polynomial on Y necessarily a Weyl invariant? The method of proof of the proposition shows that the answer is yes, provided the following variety is normal:

$$V = \{(Te, Ty) \mid T \in G,\ y \in Y\}. \tag{11}$$

(Strictly speaking we have to complexify V. Let us ignore this point.) If $Y = X$ (Problem 3) then again V will be defined by the vanishing of a single quadratic form and so must be normal. For Y such as we need to study the Bergman kernel, the analysis of the singularity of V is very complicated, and V is not normal.

To see what actually happens when we deal with Problem 4, we present a nontrivial example. Take $G = O(n, 1)$ the group of matrices preserving the quadratic form $\|v\|^2 = 2v_0v_1 - \Sigma_{k\geq 2} v_k^2$, set $e = (1, 0, \ldots, 0) \in$ Light Cone, and set $H = \{T \in G \mid Te = e\}$. We fix a positive integer g, and define for $N > g$ the vector spaces $X_N = \{$symmetric N-tensors $R = (R_{j_1 j_2, \ldots, j_N})\}$, $Y_N = \{R \in X_N \mid (R_{j_1, j_2, \ldots j_g, 0, \ldots, 0}) = 0\}$. G acts in the obvious way on X_N and $Y_N \subseteq X_N$ is an H-invariant subspace. So we are in the setting of Problem 4, and we ask whether every H-invariant polynomial P on Y_N is a Weyl invariant. The best answer we know depends on the degree of P.

To state the result, we have to bring in a new ingredient, namely the relation of the different Y_N. There is a natural projection Π: $X_{N+1} \to X_N$ defined by setting $(\pi R)_{j_1 \cdots j_N} = R_{j_1 \cdots j_N 0}$. Note that π commutes with the action of H and carries Y_{N+1} to Y_N. By composing, we get projections $\pi_N^{N'}$: $Y_{N'} \to Y_N$ for $N' > N$, and these commute with the action of H. So if P is an H-invariant polynomial on Y_N, then $P \circ \pi_N^{N'}$ is an H-invariant polynomial on $Y_{N'}$ $(N' > N)$. Furthermore, it is easy to check that if P is a Weyl invariant, then $P \circ \pi_N^{N'}$ is a Weyl invariant. Now we can state our result.

THEOREM 3. *Let P be a homogeneous H-invariant polynomial on Y_N.*

(A) *If* $\deg P < n - 1$, *then P is a Weyl invariant.*

(B) *If* $\deg P > n - 1$, *then $P \circ \pi_N^{N'}$ is a Weyl invariant for large $N' > N$.*

To prove the theorem, we pass to the varieties V_N given by (11). Explicitly

$$V_N = \left\{ (v, R) \in R^{n+1} \oplus X_N \;\middle|\; \begin{array}{l} \text{(i) } \|v\|^2 = 0 \text{ and} \\ \text{(ii) } \displaystyle\sum_{k_1, \ldots, k_{N-g}} v_{k_1} \cdots v_{k_{N-g}} R_{j_1 \cdots j_g k_1 \cdots k_{N-g}} = 0 \end{array} \right\}.$$

Requirement (ii) means Trace$[v \otimes \cdots \otimes v \otimes R] = 0$. There are G-invariant projections π: $V_{N+1} \to V_N$ defined by sending (v, R) to $(v, \tilde{R})$ with $\tilde{R}_{j_1 \cdots j_N} = \Sigma_k v_k R_{j_1 \cdots j_N k}$, i.e. $\tilde{R} =$ Trace$[v \otimes R]$. The main step in proving Theorem 3 is to show the following:

STABLE NORMALITY OF V_N. *Let $F(v, R)$ be a rational function on V_N holomorphic off the singular set. Suppose $F(v, R)$ is homogeneous of degree d in R.*

(A) *If $d < n - 1$ then F is the restriction of a polynomial to V_N.*

(B) *If $d > n - 1$ then F need not be the restriction of a polynomial. However, for large $N' > N$, $F \circ \pi_N^{N'}$ is the restriction of a polynomial to $V_{N'}$.*

The projections $\pi_N^{N'}$ collapse $V_{N'}$ near the singular set, so there is a chance of making a nice transcendental proof of stable normality. So far, the only known proof is rather nasty and technical.

At last we can discuss the example needed for the Bergman kernel. Here $G = U(n, 1)$, the group of complex matrices preserving $\|v\|^2 = v_0\overline{v_1} + v_1\overline{v_0} - \sum_{k\geq 2} v_k\overline{v_k}$, while H consists of all $T \in G$ which fix the vector $e = (1, 0, 0, \ldots, 0)$. We set $X_N = \sum_{0\leq p,q\leq N} \oplus X_{pq}$, where X_{pq} consists of all tensors $(R_{i\bar{j}k\bar{l},\mu_1\cdots\mu_p\bar{\nu}_1\cdots\bar{\nu}_q})$ of rank $(p+2, q+2)$, and

$$\varepsilon_N = \{R \in X_N \mid R_{i\bar{j}k\bar{l},\text{etc.}} \text{ arise as the covariant derivatives at } (1, 0, \ldots, 0) \text{ of the metric}$$

$$ds^2 = \sum_{jk\geq 0} \frac{\partial^2 U}{\partial z_j \partial \bar{z}_k} dz_j\, \overline{dz_k}$$

associated to some domain in Moser's normal form$\}$.

The main problem for the Bergman kernel is to understand the H-invariant polynomials on ε_N. It turns out that ε_N sits in X_N as a submanifold, and that the study of H-invariant polynomials on ε_N can be easily reduced to $Y_N =$ Tangent space to ε_N at the origin. Roughly speaking,

$$Y_N = \left\{\left(R_{j_1\cdots j_N\bar{k}_1\cdots\bar{k}_N}\right) \mid R_{j_10\cdots 0\bar{k}_1\cdots\bar{k}_N} = R_{j_1\cdots j_N\bar{k}_1\bar{0}\cdots\bar{0}} = 0\right\},$$

so that our problem is rather analogous to the setting of Theorem 3. Although we expect an analogous result here, the study of H-invariants on Y_N turns out to be much harder than Theorem 3. After much work, one can derive the following partial analogoue of Theorem 3.

THEOREM 4. *Every H-invariant polynomial of degree $n \leq 20$ on Y_N is a Weyl invariant.*

To prove this requires the analogue of part (A) of stable normality. The details are extraordinarily hard. We next show how Theorem 4 yields the Bergman kernel. With the notation of §§1, 2, we have

THEOREM 5. $K_D(z, z) = (C_k/u^{n+1})\{1 + \sum_s P_s U^s\} + O(\psi^{-20})$, *where P_s is a linear combination of traces of tensor products of the* $(R_{\alpha\bar{\beta}\gamma\bar{\delta},p\cdots\bar{\sigma}})$.

PROOF. By induction on $l \leq n - 20$, we shall prove that

$$K_D(z, z) = \frac{C_n}{u^{n+1}}\left\{1 + \sum_{s\leq l} P_s U^s + O\left(\psi^{l+1}\right)\right\}. \tag{12}$$

For $l = 0$ this is Theorem 1, while for $l = 2$ it is Theorem 2. So assume (12) for l and try to prove it for $l + 1$. From the known case of (12) and since $K_D(z, z) = \phi/u^{n+1} + O(\log\psi)$ with $\phi \in C^\infty(\overline{D})$ by Theorem 1, we obtain trivially

$$K_D(z, z) = \frac{C_n}{u^{n+1}}\left\{1 + \sum_{s\leq l} P_s U^s + \tilde{P}_{l+1} U^{l+1} + O\left(\psi^{l+2}\right)\right\} \tag{13}$$

for some smooth function $\tilde{P}_{l+1}$. The problem is to show that $\tilde{P}_{l+1}$ can be taken as a linear combination of traces of tensor products of $(R_{\alpha\bar{\beta}\gamma\bar{\delta},\text{etc.}})$.

Now, equation (13) shows that $\tilde{P}_{l+1}$ is uniquely defined on the boundary ∂D. Moreover, if ∂D is in Moser's normal form with coefficients $(A^l_{\alpha\bar{\beta}})$, then by going carefully through either of the two proofs of Theorem 1, we can check that the value of $\tilde{P}_{l+1}$ at the origin is a polynomial in the $A^l_{\alpha\bar{\beta}}$. Let us write $\tilde{P}_{l+1}(0) = Q(A^l_{\alpha\bar{\beta}})$. Furthermore, (13) and the transformation laws for the Bergman kernel and the Monge-Ampère equation show that the polynomial Q has its own transformation law under the action of the isotropy group H on normal forms. More precisely, suppose D and $\tilde{D}$ are in normal form with coefficients $(A^l_{\alpha\bar{\beta}})$ and $(\tilde{A}^l_{\alpha\bar{\beta}})$ respectively, and suppose D and $\tilde{D}$ are equivalent by a biholomorphic map Φ fixing the origin. Then $Q(A^l_{\alpha\bar{\beta}}) = |\det \Phi'(0)|^s Q(\tilde{A}^l_{\alpha\bar{\beta}})$ for a suitable s depending on l. We shall say that Q is a *boundary invariant of weight* s.

Using Theorem 4, we can understand arbitrary boundary invariants of weight s for a range of s. The point is simply to switch over from the $A^l_{\alpha\bar{\beta}}$ to the covariant derivatives $(R_{\alpha\bar{\beta}\gamma\bar{\delta},\rho\cdots\bar{\sigma}})$ evaluated at $(1,0)$ (which in projective space lies over the origin). One checks that $A^l_{\alpha\bar{\beta}}$ may be expressed as a polynomial in the $R_{\alpha\bar{\beta}\gamma\bar{\delta},\text{etc.}}$, so that our boundary invariant Q may now be written as $Q = P(R_{\alpha\bar{\beta}\gamma\bar{\delta},\text{etc.}})$. The transformation law for boundary invariants shows that P has a certain homogeneity and is H-invariant on

$$\varepsilon_N = \left\{\text{tensors}\left(R_{\alpha\bar{\beta}\gamma\bar{\delta},\text{etc.}}\right) \text{ arising from a domain in normal form}\right\}.$$

The homogeneity of P imposes a bound on its degree, and Theorem 4 can be used to show that P is a linear combination of traces of tensor products of the $(R_{\alpha\bar{\beta}\gamma\bar{\delta},\text{etc.}})$. Therefore Q is a boundary invariant of weight s if and only if Q is a linear combination of traces of $(R_{\alpha\bar{\beta}\gamma\bar{\delta},\text{etc.}})$ arising from the domain with coefficients $(A^l_{\alpha\bar{\beta}})$.

Now we can apply our discussion of boundary invariants to the polynomial Q determined by P_{l+1} in (13). For a suitable P_{l+1} = linear combination of traces of tensor products of $(R_{\alpha\bar{\beta}\gamma\bar{\delta},\text{etc.}})$, we have $P_{l+1}(0) = P_{l+1}$ for domains in normal form. Equation (13) shows that

$$K_D(z,z) = \frac{C_n}{u^{n+1}}\left\{1 + \sum_{s \leqslant l+1} P_s U^s + O\left(\psi^{l+2}\right)\right\} \quad \text{as } z \to 0$$

for D in normal form. Since both sides transform invariantly under biholomorphic maps, the same is true when z approaches an arbitrary boundary point of an arbitrary strictly pseudoconvex domain. This is exactly equation (12) with l replaced by $l + 1$, so the induction is complete. □

Essentially, we have found all the nonlinear differential operators of the form S in §2. The discussion fails completely when applied to T (the log term), both because our Monge-Ampère function carries an ambiguity $O(\psi^{n+2})$ and because the invariant theory will now involve an analogue of the hard case $\deg P = n - 1$ in Theorem 3. R. Graham has recently made discoveries related to the problem of writing down T.

We should mention that an analogue of Theorem 5 holds also for the Szegö kernel.

Throughout this chapter, the analogy between the heat kernel and the Bergman kernel has shown itself ever more clearly. Is there an analogue of the Gauss-Bonnet theorem for Chern-Moser invariants?

REFERENCES

1. V. I. Arnold, *Mathematical methods of classical mechanics*, Springer-Verlag, Berlin and New York, 1978.

2. S. Bell and E. Ligocka, *A simplification and extension of Fefferman's theorem on biholomorphic mappings*, Invent. Math. **57** (1980), 283–289.

3. M. Berger, P. Gauduchon and E. Mazet, *Le spectre d'une variété Riemannienne*, Springer-Verlag, Berlin and New York, 1971.

4. L. Boutet de Monvel, *Integration des équations de Cauchy-Riemann induites formelles*, Séminaire Goulaouic-Lions-Schwartz (1974–1975).

5. L. Boutet de Monvel and J. Sjöstrand, *Sur la singularité des noyaux de Bergman et de Szegö*, Soc. Math. de France Astérisque **34–35** (1976), 123–164.

6. L. Boutet de Monvel, to appear.

7. D. Burns and S. Schneider, personal communication.

8. E. Cartan, *Sur la géométrie pseudo-conforme des hypersurfaces de deux variables complexes*. I, II, Oeuvres II, **2**, 1231–1304; Oeuvres III, **2**, 1217–1238.

9. D. Catlin, *Necessary conditions for subellipticity and hypoellipticity for the ∂-Neumann problem on pseudoconvex domains*, in [**26**].

10. S. Y. Cheng and S-T. Yau, *On the regularity of the Monge-Ampère equation* $\det(\partial^2 u/\partial x^i \partial x^j) = F(x, u)$, Comm. Pure Appl. Math. **30** (1977), 41–68.

11. S. S. Chern and J. Moser, *Real hypersurfaces in complex manifolds*, Acta Math. **133** (1974), 219–271.

12. H. Christoffers, Thesis, Univ. of Chicago, 1980.

13. J. P. D'Angelo, *Orders of contact of real and complex subvarieties*, Illinois J. Math. **26** (1982), 41–51.

14. K. Diederich and P. Pflug, *Necessary conditions for hypoellipticity of the $\bar\partial$-problem*, in [**26**].

15. J. J. Duistermaat and V. Guillemin, *The spectrum of positive elliptic operators and periodic geodesics*, Invent. Math. **29** (1975), 39–79.

16. J. J. Duistermaat and J. Sjöstrand, to appear.

17. Yu. V. Egorov, *Subelliptic pseudodifferential operators*, Dokl. Akad. Nauk SSSR **188** (1969), 20–22.

18. ______, *On canonical transformations of pseudodifferential operators*, Uspehi Mat. Nauk **25** (1969), 235–236.

19. ______, *Subelliptic operators*, Uspehi Mat. Nauk **30:2** (1975), 71–114 = Russian Math. Surveys **30:2** (1975), 59–118.

20. ______, *Subelliptic operators*, Uspehi. Mat. Nauk **30:3** (1975), 57–104 = Russian Math. Surveys **30:3** (1975), 55–105.

21. C. Fefferman, *The Bergman kernel and biholomorphic mappings of pseudoconvex domains*, Invent. Math. **26** (1974), 1–65.

22. ______, *Monge-Ampère equations, the Bergman kernel, and geometry of pseudoconvex domains*, Ann. of Math. (2) **103** (1976), 395–416; erratum **104** (1976), 393–394.

23. ______, *Parabolic invariant theory in complex analysis*, Adv. in Math. **31** (1979), 131–262.

24. G. B. Folland and J. J. Kohn, *The Neumann problem for the Cauchy-Riemann complex*, Princeton Univ. Press, Princeton, N. J., 1972.

25. G. B. Folland and E. M. Stein, *Estimates for the $\bar\partial_b$ complex and analysis on the Heisenberg group*, Comm. Pure Appl. Math. **27** (1974), 429–522.

26. J. E. Fornaess (editor), *Recent developments in several complex variables*, Princeton Univ. Press, Princeton, N. J., 1981.

27. P. Gilkey, *The index theorem and the heat equation*, Publish or Perish, Berkeley, Calif., 1974.

28. H. Goldstein, *Classical mechanics*, Addison-Wesley, Reading, Mass., 1950.

29. R. Graham, to appear.

30. V. Guillemin, *Some classical theorems in spectral theory revisited*, in [**35**].

31. G. Hochschild and G. D. Mostow, *Unipotent groups in invariant theory*, Proc. Nat. Acad. Sci. USA **70** (1973), 646–648.

32. L. Hörmander, *Hypoelliptic second order differential equations*, Acta Math. **119** (1967), 141–171.

33. ______, *The spectral function of an elliptic operator*, Acta Math. **121** (1968), 193–218.

34. ______, *Fourier integral operators*. I, Acta Math. **127** (1971), 79–183.

35. L. Hörmander (editor), *Seminar on singularities of solutions of linear partial differential equations*, Princeton Univ. Press, Princeton, N. J., 1979.

36. L. Hörmander, *Subelliptic operators*, in [**35**].

37. N. Kerzman, *The Bergman kernel function: differentiability at the boundary*, Math. Ann. **195** (1972), 149–158.

38. S. Kobayashi and K. Nomizu, *Foundations of differential geometry*, Wiley, New York, 1969.

39. J. J. Kohn, *Harmonic integrals on strongly pseudoconvex manifolds*. I, II, Ann. of Math. (2) **78** (1963), 112–148; ibid **79** (1964), 450–472.

40. ______, *Subellipticity of the $\bar\partial$-Neumann problem on pseudoconvex domains: sufficient conditions*, Acta Math. **142** (1979), 79–122.

41. S. G. Krantz, *Function theory of several complex variables*, Wiley, New York, 1982.

42. M. Kuranishi, *Strongly pseudoconvex CR structures over small balls*. I, II, Ann. of Math. (2) **115** (1982), 451–500; ibid **116** (1982), 1–64.

43. S. Lang, *Algebra*, Addison-Wesley, Reading, Mass., 1971.

44. S. Lee and R. Melrose, personal communication.

45. V. P. Maslov, *Theory of perturbations and asymptotic methods*, Moskov. Gos. Univ., Moscow, 1965. (Russian)

46. A. Melin and J. Sjöstrand, *Fourier integral operators with complex valued phase functions*, Lecture Notes in Math., vol. 459, Springer-Verlag, Berlin and New York, pp. 120–223.

47. S. Minakshisundaram and A. Pleijel, *Some properties of the eigenfunctions of the Laplace-operator on Riemannian manifolds*, Canad. J. Math. **1** (1949), 242–256.

48. J. Moser, *Holomorphic equivalence and normal forms of hypersurfaces*, Proc. Sympos. Pure Math., vol. 27, part 2, Amer. Math. Soc., Providence, R. I., 1975, pp. 109–112.

49. A. Nagel and E. M. Stein, *Lectures on psudo-differential operators*, Princeton Univ. Press, Princeton, N. J., 1979.

50. L. Nirenberg, *On a problem of Hans Lewy*, Uspehi. Mat. Nauk **292** (176) (1974), 241–251.

51. L. Nirenberg, S. Webster and P. Yang, *Local boundary regularity of holomorphic mappings*, Comm. Pure Appl. Math. **33** (1980), 305–338.

52. V. K. Patodi, *Curvature and eigenforms of the Laplace operator*, J. Differential Geom. **5** (1971), 233–249.

53. D. H. Phong, *On integral representations of the Neumann operator*, Proc. Nat. Acad. Sci. USA **76** (1979), 1554–1558.

54. L. P. Rothschild and E. M. Stein, *Hypoelliptic differential operators and nilpotent groups*, Acta Math. **137** (1976), 247–320.

55. M. Sato, T. Kawai and M. Kashiwera, *Microfunctions and pseudodifferential equations*, Lecture Notes in Math., vol. 287, Springer-Verlag, Berlin and New York, 1973.

56. C. Seshadri, *On a theorem of Weizenböck in invariant theory*, J. Math. Kyoto Univ. **1** (1962), 403–409.

57. E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Univ. Press, Princeton, N. J., 1970.

58. N. Tanaka, *On the pseudo-conformal geometry of hypersurfaces of the space of n complex variables*, J. Math. Soc. Japan **14** (1962), 397–429.

59. ______, *Graded Lie algebras and geometric structures*, Proc. U. S.-Japan Seminar in Differential Geometry, Nippon Hyoronsha, 1966, pp. 147–150.

60. D. S. Tartakoff, *A survey of some recent results in C^∞ and real analytic hypoellipticity for partial differential equations with applications to several complex variables, in* [**26**].

61. M. E. Taylor, *Pseudodifferential operators*, Princeton Univ. Press, Princeton, N. J., 1981.

62. F. Trèves, *Analytic hypoellipticity of a class of pseudodifferential operators with double characteristics and applications to the* $\bar{\partial}$*-Neumann problem*, Comm. Partial Differential Equations **3** (1978), 475–642.

63 ______, *Introduction to pseudodifferential and Fourier integral operators*, 2 vols., Plenum Press, New York, 1980.

64. S. M. Webster, *On the mapping problem for algebraic real hypersurfaces*, Invent. Math. **43** (1977), 53–68.

65. R. Weizenböck, *Über die invarianten von linearen gruppen*, Acta Math. **58** (1932), 230–250.

66. R. O. Wells, Jr., *The Cauchy-Reimann equations and differential geometry*, Bull. Amer. Math. Soc. (N.S.) **6** (1982), 187–199.

67. H. Weyl, *Classical groups*, Princeton Univ. Press, Princeton, New Jersey 1946.

DEPARTMENT OF MATHEMATICS, RUTGERS UNIVERSITY, NEW BRUNSWICK, NEW JERSEY 08903

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY, PRINCETON, NEW JERSEY 08544

PROGRAM IN APPLIED MATHEMATICS, PRINCETON UNIVERSITY, PRINCETON, NEW JERSEY 08544