CHAPTER 1. THE HEISENBERG GROUP AND ITS REPRESENTATIONS

The Heisenberg group and its Lie algebra were born long before they were christened.

The Heisenberg Lie algebra is so named because its structure equations are Heisenberg's canonical commutation relations in quantum mechanics. These relations, however, are merely the quantized version of the Poisson bracket relations for canonical coordinates in Hamiltonian mechanics, and the importance of the latter was recognized as long ago as 1843 in Jacobi's lectures on dynamics ([82, Vorlesung 35]). Moreover, the Heisenberg group and its discrete variants have long played an implicit role in the theory of theta functions and related parts of analysis and number theory. But the names "Heisenberg group" and "Heisenberg algebra" did not come into common usage until the 1970's, and only since that time has the Heisenberg group received the recognition it deserves.

In an abstract sense, the Heisenberg group has only one locally faithful irreducible unitary representation. More precisely, up to unitary equivalence it has a one-parameter family of such representations, all of which are related to one another via automorphisms of the Heisenberg group. There are, however, several quite different "natural" ways of realizing these representations concretely in particular Hilbert spaces, and the really interesting part of the representation theory of the Heisenberg group consists of studying these realizations, the relations among them, and the integral transforms and special functions derived from them, in detail. It is to this task that this chapter is largely devoted.

1. Background from Physics

Much of the material in this monograph is motivated or illuminated by ideas coming from physics. The relevant physics is on a very basic level: the classical and quantum kinematics of a single particle moving in n-dimensional space. (We reassure the reader, perhaps unnecessarily, that the case $n = 3$ is not the only physically meaningful one: for example, if $n = 3k$, a "particle in n-space" can be a mathematical model for k particles in 3-space.) In this section we provide a very brief review of the classical and quantum pictures and the relationship between them. For further information, the following are good references: for classical mechanics, Abraham–Marsden [1], Arnold [6], Goldstein [58]; for quantum mechanics, Landau–Lifshitz [94], Mackey [98], Messiah [103].

Hamiltonian Mechanics. According to Newton's second law, the motion of a particle is governed by a second-order ordinary differential equation involving the forces acting on the particle, so that once these forces are known, the motion of the particle is completely determined by its position and velocity at a particular time. In other words, position and velocity give a complete specification of the "state" of the particle. However, in the Hamiltonian description it is found to be preferable to replace velocity by momentum (= mass × velocity). Therefore, we shall take as the state space the so-called **phase space** $\mathbf{R}^{2n}$ with coordinates

$$(p,q) = (p_1, \dots, p_n, q_1, \dots, q_n)$$

where p is the momentum vector of the particle and q is its position vector. (From a coordinate-invariant point of view, one should regard momentum as a cotangent vector and $\mathbf{R}^{2n}$ as the cotangent bundle of $\mathbf{R}^n$.) The physical **observables** are real-valued functions on phase space. Thus every observable is a function of position and momentum—another version of the fact that the latter quantities specify the state of the system.

The time evolution of the system, and its various symmetries, are given by certain transformations of phase space. The characteristic feature of Hamiltonian mechanics is that these transformations are not arbitrary diffeomorphisms, but rather share the property of leaving invariant the differential form*

$$\Omega = \sum_1^n dp_j \wedge dq_j.$$

Ω is a translation-invariant bilinear form on tangent vectors; if we identify the tangent space to $\mathbf{R}^{2n}$ at any point with $\mathbf{R}^{2n}$ itself, Ω becomes the standard **symplectic form** on $\mathbf{R}^{2n}$, which we denote by brackets:

$$[(p,q), (p',q')] = pq' - qp'. \tag{1.1}$$

* The formalism of differential geometry will be used for the next couple of pages and then will disappear from the scene; the reader to whom it is unfamiliar is asked to be indulgent during this interlude.

The diffeomorphisms of $\mathbf{R}^{2n}$ that preserve Ω are called **canonical transformations** or **symplectomorphisms**. The group of linear canonical transformations, i.e., the group of all $T \in GL(2n, \mathbf{R})$ such that

$$[T(p,q),\, T(p',q')] = [(p,q),\, (p',q')]$$

is the **symplectic group** $Sp(n, \mathbf{R})$.

Since the form Ω is nondegenerate, it provides an identification of tangent vectors with cotangent vectors. Actually, it provides two equally good ones, differing from each other by a minus sign; we shall choose the one that assigns to each tangent vector X at a point (p,q) the cotangent vector ω_X at (p,q) whose action as a linear form on tangent vectors at (p,q) is given by

$$\omega_X(Y) = \Omega(Y, X).$$

If f is a smooth observable ($=$ function on $\mathbf{R}^{2n}$), the vector field associated to the one-form df under this correspondence is called the **Hamiltonian vector field** of f and is denoted by X_f. Explicitly, X_f is given by

$$\Omega(Y, X_f) = df(Y)$$

$$\text{or} \qquad X_f = \sum_1^n \left(\frac{\partial f}{\partial p_j}\frac{\partial}{\partial q_j} - \frac{\partial f}{\partial q_j}\frac{\partial}{\partial p_j} \right).$$

If f and g are smooth observables, their **Poisson bracket** $\{f, g\}$ is the observable defined by

$$\text{(1.2)} \qquad \{f,g\} = \Omega(X_f, X_g) = \sum \left(\frac{\partial f}{\partial p_j}\frac{\partial g}{\partial q_j} - \frac{\partial f}{\partial q_j}\frac{\partial g}{\partial p_j} \right).$$

It is easily verified that the Poisson bracket is skew-symmetric and satisfies the Jacobi identity

$$\{f, \{g,h\}\} + \{h, \{f,g\}\} + \{g, \{h,f\}\} = 0,$$

so that it makes the space of smooth observables into a Lie algebra. It is also easily verified that the map $f \to X_f$ is a Lie algebra homomorphism,

$$[X_f, X_g] = X_{\{f,g\}},$$

whose kernel consists of the constant functions. The Poisson bracket of the coordinate functions p_j, q_k are given by

$$\text{(1.3)} \qquad \{p_j, p_k\} = \{q_j, q_k\} = 0, \qquad \{p_j, q_k\} = \delta_{jk}.$$

Any system of coordinates p'_j, q'_k that satisfies (1.3) is called **canonical**, and the canonical transformations of $\mathbf{R}^{2n}$ are characterized by the fact that they map each canonical coordinate system into another one.

The "infinitesimal automorphisms" of the system are the vector fields X such that the flow generated by X consists of (local) canonical transformations. This holds precisely when the Lie derivative of Ω along X vanishes, i.e.,

$$i_X(d\Omega) + d(i_X\Omega) = 0$$

where, for any k-form ϕ, $i_X\phi$ is the contraction of X with ϕ:

$$i_X\phi(Y_1, ..., Y_{k-1}) = \phi(X, Y_1, ..., Y_{k-1}).$$

But $d\Omega = 0$ and $i_X\Omega = -\omega_X$ in the terminology introduced above, so X is an infinitesimal automorphism iff ω_X is closed. Since $\mathbf{R}^{2n}$ is simply connected, all closed one-forms are exact, so the closedness of ω_X means that X is a Hamiltonian vector field. In short, the map $f \to X_f$ establishes a one-to-one correspondence between smooth observables modulo constants and infinitesimal automorphisms. The latter comprise (in a rough sense) the "Lie algebra" of the group of canonical transformations, and the Poisson bracket is the pullback of this Lie algebra structure to the observables. If H is a smooth observable, the canonical transformations generated by X_H are obtained by integrating Hamilton's equations:

$$\frac{dp_j}{dt} = -\frac{\partial H}{\partial q_j}, \qquad \frac{dq_j}{dt} = \frac{\partial H}{\partial p_j}.$$

Quantum Mechanics. In classical mechanics, once the state of the system is specified, the value of every observable is completely determined. In the quantum world this is no longer true: in a given state, observables have only probability distributions of values, which may be, but usually are not, concentrated at a single point. The mathematical setup is as follows. The state space for a quantum system is a projective Hilbert space $\mathbf{P}\mathcal{H}$, that is, the set of all (complex) lines through the origin in a Hilbert space $\mathcal{H}$. Normally we think of states as being given by unit vectors in $\mathcal{H}$, with the understanding that two unit vectors define the same state when they are scalar multiples of one another. The observables are projection-valued Borel measures on $\mathbf{R}$, that is, mappings Π from the Borel sets in $\mathbf{R}$ to the orthogonal projections on $\mathcal{H}$ such that $\Pi(\mathbf{R}) = I$ and if $E_1, E_2, ...$ are disjoint Borel sets,

$$(1.4) \qquad \Pi(E_j)\Pi(E_k) = 0 \quad \text{for} \quad j \neq k, \qquad \sum_j \Pi(E_j) = \Pi\Big(\bigcup_j E_j\Big).$$

If $\mathcal{R}(E)$ is the range of $\Pi(E)$, (1.4) is equivalent to

$$\mathcal{R}(E_j)\perp\mathcal{R}(E_k) \quad \text{for} \quad j \neq k, \qquad \bigoplus_j \mathcal{R}(E_j) = \mathcal{R}(\bigcup_j E_j).$$

If Π is such a measure and $u \in \mathcal{H}$ is a unit vector, then $E \to \langle \Pi(E)u, u\rangle$ is an ordinary probability measure on $\mathbf{R}$, and this is the probability distribution of the observable Π in the state u.

By the spectral theorem, projection-valued measures Π are in one-to-one correspondence with self-adjoint operators A on $\mathcal{H}$ (cf. Appendix B):

$$A = \int \lambda \, d\Pi(\lambda), \qquad \Pi(E) = \chi_E(A).$$

Thus one can, and generally does, think of observables as self-adjoint operators. We note that if A is a self-adjoint operator and u is a unit vector in the domain of A, the mean or expectation of the observable A in the state u is

$$\int \lambda \langle d\Pi(\lambda)u, u\rangle = \langle Au, u\rangle.$$

The probability distribution of A in the state u is concentrated at a single point λ precisely when u is an eigenvector of a with eigenvalue λ. Moreover, if the spectrum of A is purely discrete, so that A has an orthonormal eigenbasis $\{e_j\}$ with eigenvalues $\{\lambda_j\}$, the probability distribution of A in any state is given by

$$E \longrightarrow \sum_{\lambda_j \in E} |\langle u, e_j\rangle|^2. \tag{1.5}$$

For the system we are interested in, a particle moving in n-space, the Hilbert space $\mathcal{H}$ is taken to be $L^2(\mathbf{R}^n)$. If $f \in L^2(\mathbf{R}^n)$ is a unit vector, $|f|^2$ is interpreted as the probability density of the position of the particle in the state f; that is, the probability that the particle will be found in a set $B \subset \mathbf{R}^n$ is $\int_B |f|^2$. From this, we can easily identify the self-adjoint operators $Q_1, ..., Q_n$ corresponding to the classical coordinate functions $q_1, ..., q_n$. Namely, if $E \subset \mathbf{R}$, the probability that the jth coordinate x_j of the particle will lie in E is

$$\int_{x_j \in E} |f(x)|^2 \, dx. \tag{1.6}$$

Thus the projection-valued measure Π_j for this observable is given by

$$\Pi_j(E) = \text{multiplication by the characteristic function of } \{x\colon\, x_j \in E\},$$

and it follows easily that the operator

$$Q_j = \int \lambda \, d\Pi_j(\lambda)$$

is multiplication by the jth coordinate function, which we generally denote by X_j:

$$Q_j f(x) = X_j f(x) = x_j f(x), \tag{1.7}$$

defined on the domain of all $f \in L^2$ such that $x_j f \in L^2$.

We observe that there are no states $f \in L^2$ for which the observables Q_j have definite values: the simultaneous eigenfunctions of the Q_j's are the delta functions $\delta_{x_0}(x) = \delta(x - x_0)$. These, however, can be considered as a set of "idealized states" that form a "continuous orthonormal basis":

$$\langle \delta_{x_1}, \delta_{x_2} \rangle = \int \delta(x - x_1)\delta(x - x_2) \, dx = \delta(x_1 - x_2),$$

$$f = \int f(x)\delta_x \, dx = \int \langle f, \delta_x \rangle \delta_x \, dx,$$

all integrals being interpreted in the sense of distributions. The formula (1.6) for the distribution of Q_j is then the analogue of (1.5).

What about momentum? According to the principles of wave mechanics (cf. Messiah [103]), the eigenfunctions for momentum are the plane waves $e_\xi(x) = e^{2\pi i x \xi}$: the momentum of e_ξ is $h\xi$, where h is Planck's constant. Like the delta functions, the e_ξ's are not in L^2 but form a "continuous orthonormal basis":

$$\langle e_{\xi_1}, e_{\xi_2} \rangle = \int e^{2\pi i(\xi_1 - \xi_2)x} dx = \delta(\xi_1 - \xi_2),$$

$$f = \int \widehat{f}(\xi) e_\xi \, d\xi = \int \langle f, e_\xi \rangle e_\xi \, d\xi.$$

Here $\widehat{f}$ is the Fourier transform of f, and these equations are restatements of the Fourier inversion formula. By analogy with (1.6) and (1.7), we deduce that the probability in the state f that the jth component of the momentum will lie in E is

$$\int_{h\xi_j \in E} |\widehat{f}(\xi)|^2 \, d\xi$$

and hence that the self-adjoint operator P_j corresponding to the classical observable p_j is given by

$$(P_j f)^\frown(\xi) = h\xi_j \widehat{f}(\xi), \qquad \text{or} \qquad P_j = h\mathcal{F}^{-1} Q_j \mathcal{F}.$$

In other words,

$$P_j = \frac{h}{2\pi i}\frac{\partial}{\partial x_j} = hD_j. \tag{1.8}$$

The "automorphisms" of a quantum mechanical system are the bijections on the state space $\mathbf{P}\mathcal{H}$ that preserve the projective inner product $|\langle u, v\rangle|^2$. (Here u and v are unit vectors in $\mathcal{H}$, but $|\langle u, v\rangle|^2$ depends only on the lines containing u and v.) By a theorem of Wigner (see Bargmann [12]), these are precisely the maps of projective space induced by unitary or anti-unitary operators on $\mathcal{H}$. For our purposes we may ignore the anti-unitary operators and regard the automorphism group as the unitary group on $\mathcal{H}$ modulo scalar multiples of the identity. The "infinitesimal automorphisms" are the generators of one-parameter groups of unitary operators, and by Stone's theorem (cf. Nelson [114], Reed–Simon [122]) these are precisely the skew-adjoint operators. If $B = -B^*$, the unitary group generated by B is e^{tB}; this is projectively trivial precisely when B is an imaginary multiple of I. If we disregard the (very substantial) difficulties involved in algebraically manipulating unbounded operators, we have a natural Lie algebra structure on the set of infinitesimal automorphisms, given by the commutator

$$[B_1, B_2] = B_1B_2 - B_2B_1.$$

Just as in classical mechanics, there is a correspondence between observables, modulo scalar multiples of I, and infinitesimal automorphisms, for we can convert any self-adjoint operator into a skew-adjoint one by multiplying by an imaginary constant. We shall find it convenient to take this constant to be $2\pi i$. Thus, the one-parameter unitary group associated to an observable A is $e^{2\pi i t A}$, and the induced (formal) Lie algebra structure on the set of observables is

$$(A_1, A_2) \longrightarrow \frac{1}{2\pi i}[2\pi i A_1, 2\pi i A_2] = 2\pi i[A_1, A_2].$$

(Remark: the "physically correct" choice of constant is not $2\pi i$ but $-2\pi i/h$. This is irrelevant for our purposes; it would result in the relabeling of some parameters when we describe certain unitary representations, but it does not ultimately affect the quantization procedures discussed in Chapter 2.) In particular, the basic observables Q_j and P_j satisfy the **canonical commutation relations**

$$[P_j, P_k] = [Q_j, Q_k] = 0, \qquad [P_j, Q_k] = \frac{h\delta_{jk}}{2\pi i}I. \tag{1.9}$$

Quantization. By the "quantization problem" we shall mean the problem of setting up a correspondence $f \to A_f$ between classical and quantum observables, i.e., between functions on $\mathbf{R}^{2n}$ and self-adjoint operators on $L^2(\mathbf{R}^n)$, such that the properties of the classical observables are reflected as much as possible in their quantum counterparts in a way consistent with the probabilistic interpretation of quantum observables. Since this discussion is intended only to provide motivation, we shall ignore all technical difficulties associated with unbounded operators. On the formal level, then, a quantization procedure $f \to A_f$ ideally should have the following properties.

(i) The quantum counterparts of the position and momentum coordinates q_j and p_j should be the operators Q_j and P_j defined by (1.7) and (1.8). Moreover, if f is a constant function c, the probability that $f = c$ is one no matter which state the system is in, whence it follows that the quantum counterpart of f must be the operator cI. Thus:

$$A_{q_j} = Q_j, \qquad A_{p_j} = P_j, \qquad A_c = cI. \tag{1.10}$$

(ii) If f, g are classical observables, the expectation of A_{f+g} in any state should be the sum of the expectations of A_f and A_g, that is,

$$\langle A_{f+g}u, u\rangle = \langle A_f u, u\rangle + \langle A_g u, u\rangle.$$

But if A is a self-adjoint operator, the diagonal matrix elements $\langle Au, u\rangle$ determine all matrix elements $\langle Au, v\rangle$, and hence the operator A, by polarization. Therefore,

$$A_{f+g} = A_f + A_g. \tag{1.11}$$

(iii) Suppose $\phi: \mathbf{R} \to \mathbf{R}$ is a Borel function. If $E \subset \mathbf{R}$ and $\mathcal{Q}$ is a probabilistically determined quantity, the probability that $\phi(\mathcal{Q}) \in E$ is the probability that $\mathcal{Q} \in \phi^{-1}(E)$. Thus, if f is a classical observable and $A_f = \int \lambda\, d\Pi_f(\lambda)$, the spectral projections for $A_{\phi\circ f}$ should be $\Pi_{\phi\circ f}(E) = \Pi_f(\phi^{-1}(E))$. But these are the spectral projections for the operator $\phi(A_f)$ defined by the spectral functional calculus (cf. Appendix B). Thus:

$$A_{\phi\circ f} = \phi(A_f). \tag{1.12}$$

(iv) A much weaker requirement than (1.12) is that $A_{cf} = cA_f$ (c constant) and $A_{f^2} = (A_f)^2$, and this together with (1.11) implies that the quantum counterpart of a product fg should be the Jordan product of A_f and A_g. Indeed,

$$(A_f + A_g)^2 = (A_{f+g})^2 = A_{(f+g)^2} = (A_f)^2 + 2A_{fg} + (A_g)^2,$$

so that

$$A_{fg} = \tfrac{1}{2}(A_f A_g + A_g A_f). \tag{1.13}$$

(v) Finally, there should be a correspondence between the Lie algebra structures of classical and quantum observables: $[A_f, A_g]$ should be a constant multiple of $A_{\{f,g\}}$. In view of (1.9) and (1.10), the constant must be $h/2\pi i$:

$$[A_f, A_g] = \frac{h}{2\pi i} A_{\{f,g\}}. \tag{1.14}$$

Now, how much of this can be accomplished? We shall insist on the basic position-momentum correspondence (1.10), the additivity (1.11), and the very special case $A_{cf} = cA_f$ ($c \in \mathbf{R}$) of (1.12): in other words, we require $f \to A_f$ to be a *linear* map satisfying (1.10). However, there is no such map that satisfies either (1.13) or (1.14) (or (1.12), which is stronger than (1.13)). That (1.13) is impossible can easily be seen as follows. Let $f(p,q) = p_1$ and $g(p,q) = q_1$. If (1.13) were true we would have

$$\tfrac{1}{4}(P_1Q_1 + Q_1P_1)^2 = (A_{fg})^2 = A_{f^2g^2} = \tfrac{1}{2}(P_1^2Q_1^2 + Q_1^2P_1^2).$$

But a simple calculation shows that

$$\tfrac{1}{4}(P_1Q_1 + Q_1P_1)^2 = \frac{-h^2}{4\pi^2}\left[x_1^2\frac{\partial^2}{\partial x_1^2} + 2x_1\frac{\partial}{\partial x_1} + \frac{1}{4}\right],$$

$$\tfrac{1}{2}(P_1^2Q_1^2 + Q_1^2P_1^2) = \frac{-h^2}{4\pi^2}\left[x_1^2\frac{\partial^2}{\partial x_1^2} + 2x_1\frac{\partial}{\partial x_1} + 1\right].$$

The proof that (1.14) cannot be satisfied is a bit more involved; we shall present it in Section 4.4 (Groenewold's theorem).

We should not be too surprised or disappointed at these negative results: life would be dull if things were so simple! We shall, however, keep (1.12), (1.13), and (1.14) in mind as guidelines, and in Chapter 2 we shall construct quantization procedures which satisfy them in an approximate sense for large classes of observables, the approximation being good when Planck's constant h is small. We shall also investigate the extent to which the (pointwise) boundedness or positivity of a classical observable can be reflected in the (operator-theoretic) boundedness or positivity of its quantum counterpart.

2. The Heisenberg Group

The Poisson bracket relations (1.3) for canonical coordinates in Hamiltonian mechanics and the commutation relations (1.9) for their quantum analogues are formally identical, and the abstract algebraic structure underlying them is the following. We consider $\mathbf{R}^{2n+1}$ with coordinates

$$(p_1, \ldots, p_n, q_1, \ldots, q_n, t) = (p, q, t),$$

and we define a Lie bracket on $\mathbf{R}^{2n+1}$ by

$$[(p,q,t),\,(p',q',t')] = (0,\,0,\,pq'-qp') = \big(0,\,0,\,[(p,q),\,(p',q')]\big), \tag{1.15}$$

where the bracket on the right is the symplectic form on $\mathbf{R}^{2n}$. It is easily verified that the bracket (1.15) makes $\mathbf{R}^{2n+1}$ into a Lie algebra, called the **Heisenberg Lie algebra** and denoted by $\mathfrak{h}_n$. If $P_1,\dots,P_n,Q_1,\dots,Q_n,T$ is the standard basis for $\mathbf{R}^{2n+1}$, the Lie algebra structure is given by

$$[P_j,P_k] = [Q_j,Q_k] = [P_j,T] = [Q_j,T] = 0, \qquad [P_j,Q_k] = \delta_{jk}T. \tag{1.16}$$

Thus, (1.3) and (1.9) say that in both classical and quantum mechanics, the momentum, position, and constant observables span a Lie algebra isomorphic to $\mathfrak{h}_n$.

In order to identify the Lie group corresponding to $\mathfrak{h}_n$, it is convenient to use a matrix representation. Given $(p,q,t)\in\mathbf{R}^{2n+1}$, we define the matrix $m(p,q,t)\in M_{n+2}(\mathbf{R})$ by

$$m(p,q,t) = \begin{pmatrix} 0 & p_1 & \dots & p_n & t \\ 0 & 0 & \dots & 0 & q_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & q_n \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

(where all entries are zero except on the first row and last column). Moreover, we define

$$M(p,q,t) = I + m(p,q,t).$$

It is easily verified that

$$m(p,q,t)m(p',q',t') = m(0,0,pq'), \tag{1.17}$$

$$M(p,q,t)M(p',q',t') = M(p+p',\,q+q',\,t+t'+pq'). \tag{1.18}$$

From (1.17) it follows that

$$[m(p,q,t),m(p',q',t')] = m(0,\,0,\,pq'-qp'),$$

where the bracket now denotes the commutator. Hence the correspondence $X\to m(X)$ is a Lie algebra isomorphism from $\mathfrak{h}_n$ to $\{m(X) : X\in\mathbf{R}^{2n+1}\}$, and to obtain the corresponding Lie group we can simply apply the matrix exponential map. From (1.17) we have

$$m(p,q,t)^2 = m(0,0,pq) \qquad \text{and} \qquad m(p,q,t)^k = 0 \quad \text{for} \quad k\ge 3,$$

so

$$e^{m(p,q,t)} = I + m(p,q,t) + \tfrac{1}{2}m(0,0,pq) = M(p,q,t+\tfrac{1}{2}pq). \tag{1.19}$$

Thus the exponential map is a bijection from $\{m(X) : X \in \mathbf{R}^{2n+1}\}$ to $\{M(X) : X \in \mathbf{R}^{2n+1}\}$, and the latter is a group with group law (1.18). We could take this to be the Lie group corresponding to $\mathfrak{h}_n$, but we prefer to use a slightly different model. It is easily verified that

$$\exp m(p,q,t)\exp m(p',q',t') = \exp m\big(p+p',\, q+q',\, t+t'+\tfrac{1}{2}(pq'-qp')\big).$$

Therefore, if we identify $X \in \mathbf{R}^{2n+1}$ with the matrix $e^{m(X)}$, we make $\mathbf{R}^{2n+1}$ into a group with group law

$$(p,q,t)(p',q',t') = \big(p+p',\, q+q',\, t+t'+\tfrac{1}{2}(pq'-qp')\big). \tag{1.20}$$

We call this group the **Heisenberg group** and denote it by $\mathbf{H}_n$. The exponential map from $\mathfrak{h}_n$ to $\mathbf{H}_n$ is then merely the identity, and the inverse of the element (p,q,t) is simply $(-p,-q,-t)$.

Occasionally it is better to identify (p,q,t) with the matrix $M(p,q,t)$, which by (1.18) yields the group law

$$(p,q,t)(p',q',t') = (p+p',\, q+q',\, t+t'+pq').$$

We call $\mathbf{R}^{2n+1}$ with this group law the **polarized Heisenberg group** and denote it by $\mathbf{H}_n^{\mathrm{pol}}$. By (1.19), the map

$$(p,q,t) \longrightarrow (p,q,t+\tfrac{1}{2}pq)$$

is an isomorphism from $\mathbf{H}_n$ to $\mathbf{H}_n^{\mathrm{pol}}$, and it is also the exponential map from $\mathfrak{h}_n$ to $\mathbf{H}_n^{\mathrm{pol}}$. The inverse of (p,q,t) in $\mathbf{H}_n^{\mathrm{pol}}$ is $(-p,-q,-t+pq)$.

We observe that

$$\mathcal{Z} = \{(0,0,t) : t \in \mathbf{R}\}$$

is the center, and also the commutator subgroup, of both $\mathbf{H}_n$ and $\mathbf{H}_n^{\mathrm{pol}}$. Moreover, Lebesgue measure on $\mathbf{R}^{2n+1}$ is a bi-invariant Haar measure on both $\mathbf{H}_n$ and $\mathbf{H}_n^{\mathrm{pol}}$.

The Automorphisms of the Heisenberg Group. We denote by $\mathrm{Aut}(\mathbf{H}_n)$ and $\mathrm{Aut}(\mathfrak{h}_n)$ the automorphism groups of $\mathbf{H}_n$ and $\mathfrak{h}_n$ (as a topological group and a Lie algebra, respectively). Since the underlying set of both $\mathbf{H}_n$ and $\mathfrak{h}_n$ is $\mathbf{R}^{2n+1}$, $\mathrm{Aut}(\mathbf{H}_n)$ and $\mathrm{Aut}(\mathfrak{h}_n)$ are both sets of mappings from $\mathbf{R}^{2n+1}$ to itself. In fact, they are equal. This is an instance of a general theorem about automorphisms of simply connected Lie groups and their Lie algebras, but we can give a simple direct proof.

(1.21) Proposition. $\mathrm{Aut}(\mathbf{H}_n) = \mathrm{Aut}(\mathfrak{h}_n)$.

Proof: By (1.15) and (1.20), the Heisenberg group and algebra structures on $\mathbf{R}^{2n+1}$ are related by

$$XY = X + Y + \tfrac{1}{2}[X,Y] \qquad (X, Y \in \mathbf{R}^{2n+1}).$$

From this it is clear that if $\alpha \in \mathrm{Aut}(\mathfrak{h}_n)$ then $\alpha \in \mathrm{Aut}(\mathbf{H}_n)$. On the other hand, suppose $\alpha \in \mathrm{Aut}(\mathbf{H}_n)$. If $[X,Y] = 0$ then

$$\begin{aligned}\alpha(X+Y) = \alpha(XY) &= \alpha(X)\alpha(Y) = \alpha(X) + \alpha(Y) + \tfrac{1}{2}[\alpha(X), \alpha(Y)] \\ = \alpha(YX) &= \alpha(Y)\alpha(X) = \alpha(X) + \alpha(Y) - \tfrac{1}{2}[\alpha(X), \alpha(Y)]\end{aligned}$$

and hence $\alpha(X+Y) = \alpha(X) + \alpha(Y)$. In particular, α is additive on every one-dimensional subspace, and it is continuous; hence it commutes with scalar multiplication. Therefore, if $X, Y \in \mathbf{R}^{2n+1}$ and $s \in \mathbf{R}$, we have

$$\begin{aligned}s\alpha\big(X + Y + \tfrac{1}{2}s[X,Y]\big) &= \alpha\big(sX + sY + \tfrac{1}{2}s^2[X,Y]\big) = \alpha\big((sX)(sY)\big) \\ &= \big(s\alpha(X)\big)\big(s\alpha(Y)\big) = s\alpha(X) + s\alpha(Y) + \tfrac{1}{2}s^2[\alpha(X), \alpha(Y)].\end{aligned}$$

If we divide through by s and let $s \to 0$ we obtain $\alpha(X+Y) = \alpha(X) + \alpha(Y)$. Thus α is linear; taking this into account, the above equation also shows that $\alpha([X,Y]) = [\alpha(X), \alpha(Y)]$. In short, $\alpha \in \mathrm{Aut}(\mathfrak{h}_n)$. ∎

We now identify the automorphisms of $\mathbf{H}_n$ explicitly. It is easy to write down several families of them:

(i) *Symplectic maps.* If $S \in Sp(n, \mathbf{R})$, the map

$$(p, q, t) \longrightarrow \big(S(p,q), t\big)$$

is clearly in $\mathrm{Aut}(\mathbf{H}_n)$. For the moment, we denote the group of such automorphisms of $\mathbf{H}_n$ by G_1.

(ii) *Inner Automorphisms.* It is easily checked that

$$(a, b, c)(p, q, t)(a, b, c)^{-1} = (p,\ q,\ t + aq - bp).$$

We denote the group of inner automorphisms of $\mathbf{H}_n$ by G_2.

(iii) *Dilations.* If $r > 0$, the map $\delta[r]$ defined by

$$\delta[r](p, q, t) = (rp, rq, r^2 t)$$

is obviously in $\mathrm{Aut}(\mathbf{H}_n)$; moreover, $\delta[rs] = \delta[r]\delta[s]$. We denote the group of dilations by G_3.

(iv) *Inversion.* The map i defined by

$$i(p, q, t) = (q, p, -t)$$

is in $\mathrm{Aut}(\mathbf{H}_n)$. We denote the two-element group consisting of i and the identity by G_4.

(1.22) Theorem. *With notation as above, every automorphism of* $\mathbf{H}_n$ *can be written uniquely as* $\alpha_1\alpha_2\alpha_3\alpha_4$ *with* $\alpha_j \in G_j$.

Proof: If $\alpha \in \mathrm{Aut}(\mathbf{H}_n)$, α maps the center $\mathcal{Z}$ to itself, and by Proposition (1.21), α is linear; hence α must be of the form

$$\alpha(p,q,t) = \big(T(p,q),\ ap + bq + st\big)$$

with $T \in GL(2n,\mathbf{R})$, $a, b \in \mathbf{R}^n$, and $s \in \mathbf{R} \setminus \{0\}$. By composing α with the inversion i if necessary, we can make $s > 0$; then by composing with the dilation $\delta[s^{-1/2}]$ we can make $s = 1$; finally by composing with a suitable inner automorphism we can make $a = b = 0$. What is left is a map of the form $(p,q,t) \to (S(p,q),t)$ where $S \in GL(2n,\mathbf{R})$, and clearly this is an automorphism of $\mathbf{H}_n$ iff $S \in Sp(n,\mathbf{R})$. ∎

3. The Schrödinger Representation

We recall that the quantum-mechanical position and momentum operators are $Q_j = X_j$ (multiplication by x_j) and $P_j = hD_j$ ($h/2\pi i$ times differentiation with respect to x_j). We may regard these operators as continuous operators on the Schwartz class $\mathcal{S}(\mathbf{R}^n)$. As such, they satisfy the commutation relations (1.9), and it follows that the map $d\rho_h$ from the Heisenberg algebra $\mathfrak{h}_n$ to the set of skew-Hermitian operators on $\mathcal{S}$ defined by

$$d\rho_h(p,q,t) = 2\pi i(hpD + qX + tI)$$

is a Lie algebra homomorphism. We wish to exponentiate this representation of $\mathfrak{h}_n$ to obtain a unitary representation of the Heisenberg group $\mathbf{H}_n$.

For the moment we take $h = 1$. The main point is to compute the operators $e^{2\pi i(pD+qX)}$. If $f \in L^2$, let

$$g(x,t) = [e^{2\pi it(pD+qX)}f](x).$$

g is the solution of the differential equation $\partial g/\partial t = 2\pi i(pD + qX)g$ subject to the initial condition $g(x,0) = f(x)$, that is,

$$\frac{\partial g}{\partial t} - \sum p_j \frac{\partial g}{\partial x_j} = 2\pi i q x g, \qquad g(x,0) = f(x).$$

The expression on the left is just the directional derivative of g along the vector $(-p,1)$, so if we set

$$x(t) = x - tp, \qquad G(t) = g\big(x(t),t\big),$$

we obtain

$$G'(t) = 2\pi i q(x - tp)G(t), \qquad G(0) = f(x).$$

This ordinary differential equation is easily solved:

$$g(x - tp, t) = G(t) = f(x)e^{2\pi i t q x - \pi i t^2 p q}.$$

Setting $t = 1$ and replacing x by $x + p$, we obtain the desired result:

$$e^{2\pi i(pD+qX)} f(x) = e^{2\pi i q x + \pi i p q} f(x + p). \tag{1.23}$$

From this formula it is evident that $e^{2\pi i(pD+qX)}$ is a unitary operator on L^2, and it is easily checked that

$$e^{2\pi i(pD+qX)} e^{2\pi i(rD+sX)} = e^{\pi i(ps-qr)} e^{2\pi i[(p+r)D+(q+s)X]}. \tag{1.24}$$

It follows therefore that the map ρ from $\mathbf{H}_n$ to the group of unitary operators on L^2 defined by

$$\rho(p, q, t) = e^{2\pi i(pD+qX+tI)} = e^{2\pi i t} e^{2\pi i(pD+qX)},$$

that is,

$$\rho(p, q, t) f(x) = e^{2\pi i t + 2\pi i q x + \pi i p q} f(x + p), \tag{1.25}$$

is a unitary representation of $\mathbf{H}_n$. Moreover, the operators $\rho(p, q, t)$, besides being unitary on L^2, are continuous on $\mathcal{S}$ and extend to continuous operators on $\mathcal{S}'$; and we shall frequently regard them as such.

At this point we can put Planck's constant back in. Namely, we set

$$\rho_h(p, q, t) = \rho(hp, q, ht) = e^{2\pi i h t} e^{2\pi i(hpD+qX)},$$

or

$$\rho_h(p, q, t) f(x) = e^{2\pi i h t + 2\pi i q x + \pi i h p q} f(x + hp). \tag{1.26}$$

Then for any real number h, ρ_h is a unitary representation of $\mathbf{H}_n$ on $L^2(\mathbf{R}^n)$, and the corresponding representation $d\rho_h$ of $\mathfrak{h}_n$ is given by (1.23). Moreover, ρ_h and $\rho_{h'}$ are inequivalent for $h \neq h'$. (It suffices to observe that the central characters $e^{2\pi i h t}$ and $e^{2\pi i h' t}$ are inequivalent for $h \neq h'$.) We shall see in the next section that ρ_h is irreducible for $h \neq 0$.

We call ρ_h the **Schrödinger representation** of $\mathbf{H}_n$ with parameter h. Generally we shall take $h = 1$ and restrict attention to the representation

$\rho = \rho_1$; the generalization to other (nonzero) values of h is an easy exercise which we shall omit except when it leads to something of particular interest.

As we pointed out in the Prologue, in some ways it is more natural to replace ρ by the representation

$$\rho'(p,q,t) = \rho(-q,p,t) = e^{2\pi it}e^{2\pi i(pX-qD)}, \tag{1.27}$$

in which the symplectic form rather than the Euclidean inner product is used to pair (p,q) with (D,X) in the exponent. Indeed, we have

$$e^{2\pi i(pX-qD)}f(x) = e^{2\pi ipx-\pi ipq}f(x-q)$$
$$[e^{2\pi i(pX-qD)}f]^{\frown}(\xi) = e^{-2\pi iq\xi+\pi ipq}\widehat{f}(\xi-p).$$

Thus if the mean values of the position and momentum of f are x_0 and ξ_0, the mean values of the position and momentum of $e^{2\pi i(pX-qD)}f$ are x_0+q and ξ_0+p. The operator $e^{2\pi i(pX-qD)}$ therefore represents a translation in position space by q and a translation in momentum space by p, which accords with the usual interpretation of p and q as momentum and position variables. However, ρ' also has its disadvantages, and we shall generally stick to ρ. In any case, it is easy to go from one representation to the other. On the Heisenberg group side, it is just a matter of composing with the automorphism $(p,q,t) \to (-q,p,t)$ of $\mathbf{H}_n$; and on the L^2 side, ρ and ρ' are intertwined by the Fourier transform, as one can easily check:

$$\mathcal{F}\rho(p,q,t)\mathcal{F}^{-1} = \rho'(p,q,t).$$

The kernel of ρ is $\{(0,0,k) : k \in \mathbf{Z}\}$. For some purposes it is better to throw away this kernel, so we define the **reduced Heisenberg group** $\mathbf{H}_n^{\mathrm{red}}$ to be the quotient

$$\mathbf{H}_n^{\mathrm{red}} = \mathbf{H}_n \;/\; \{(0,0,k) : k \in \mathbf{Z}\}.$$

We still write elements of $\mathbf{H}_n^{\mathrm{red}}$ as (p,q,t), with the understanding that t is taken to be a real number mod 1, and we regard ρ as a representation of $\mathbf{H}_n^{\mathrm{red}}$, which is now *faithful.* In fact, since the central variable t always acts in a simple-minded way, as multiplication by the scalar $e^{2\pi it}$, it is often convenient to disregard it entirely; we therefore define

$$\rho(p,q) = \rho(p,q,0) = e^{2\pi i(pD+qX)}.$$

The Integrated Representation. The unitary representation ρ of $\mathbf{H}_n^{\mathrm{red}}$ determines a representation of the convolution algebra $L^1(\mathbf{H}_n^{\mathrm{red}})$, still denoted by ρ, in the usual way: if $\Phi \in L^1(\mathbf{H}_n^{\mathrm{red}})$,

$$\rho(\Phi) = \int_{\mathbf{H}_n^{\mathrm{red}}} \Phi(X)\rho(X)\,dX = \iiint \Phi(p,q,t)\rho(p,q,t)\,dp\,dq\,dt.$$

The integral here is a Bochner integral, and $\rho(\Phi)$ is an operator on $L^2(\mathbf{R}^n)$ satisfying $\|\rho(\Phi)\| \le \|\Phi\|_1$.

Given $\Phi \in L^1(\mathbf{H}_n^{\mathrm{red}})$, we can expand it in a Fourier series in the central variable t:

$$\Phi(p,q,t) = \sum_{-\infty}^{\infty} \Phi_k(p,q)e^{2\pi ikt}.$$

(This series can be interpreted, for example, as the limit in the L^1 norm of its Cesàro means.) Since $\rho(p,q,t) = e^{2\pi it}\rho(p,q)$, we have

$$\begin{aligned}\rho(\Phi) &= \sum_{-\infty}^{\infty} \iiint \Phi_k(p,q)\rho(p,q)e^{2\pi i(k+1)t}dp\,dq\,dt \\ &= \iint \Phi_{-1}(p,q)\rho(p,q)\,dp\,dq.\end{aligned}$$

Thus, the only part of Φ that contributes to $\rho(\Phi)$ is the (-1)th Fourier component Φ_{-1}, so we might as well consider ρ as a representation of $L^1(\mathbf{R}^{2n})$ (with a nonstandard convolution structure, to be discussed below) rather than of $L^1(\mathbf{H}_n^{\mathrm{red}})$. Accordingly, for $F \in L^1(\mathbf{R}^{2n})$ let us define

$$\rho(F) = \iint F(p,q)\rho(p,q)\,dp\,dq = \iint F(p,q)e^{2\pi i(pD+qX)}dp\,dq. \tag{1.28}$$

($\rho(F)$ is sometimes called the "Weyl transform" of F, but this is historically inaccurate. In fact, the "Weyl transform" of F should be $\rho(\widehat{F})$, as we shall explain in Chapter 2.) The explicit formula for the operator $\rho(F)$ is as follows:

$$\begin{aligned}\rho(F)f(x) &= \iint F(p,q)e^{2\pi iqx+\pi ipq}f(x+p)\,dp\,dq \\ &= \iint F(y-x,q)e^{\pi iq(x+y)}f(y)\,dy\,dq.\end{aligned}$$

In other words, $\rho(F)$ is an integral operator with kernel

$$\begin{aligned}K_F(x,y) &= \int F(y-x,q)e^{\pi iq(x+y)}dq \\ &= (\mathcal{F}_2^{-1}F)\left(y-x, \frac{y+x}{2}\right),\end{aligned} \tag{1.29}$$

where $\mathcal{F}_2$ denotes Fourier transformation in the second variable. From this we easily deduce:

(1.30) Theorem. *The map ρ from $L^1(\mathbf{R}^{2n})$ to the space of bounded operators on $L^2(\mathbf{R}^n)$, defined by (1.28), extends uniquely to a bijection from $\mathcal{S}'(\mathbf{R}^{2n})$ to the space of continuous linear maps from $\mathcal{S}(\mathbf{R}^n)$ to $\mathcal{S}'(\mathbf{R}^n)$. Moreover, ρ maps $L^2(\mathbf{R}^{2n})$ unitarily onto the space of Hilbert-Schmidt operators on $L^2(\mathbf{R}^n)$, and $\rho(F)$ is a compact operator on $L^2(\mathbf{R}^n)$ for all $F \in L^1(\mathbf{R}^{2n})$.*

Proof: The kernel K_F is obtained from F by partial Fourier transformation followed by an invertible and measure-preserving change of variable. These operations make sense when F is an arbitrary tempered distribution and define K_F as a tempered distribution. In this case the operation $f \to \int K_F(\cdot, y)f(y)\,dy$ defines a continuous linear map from $\mathcal{S}(\mathbf{R}^n)$ to $\mathcal{S}'(\mathbf{R}^n)$. Explicitly, if $f, g \in \mathcal{S}(\mathbf{R}^n)$,

$$\langle \int K_F(\cdot, y)f(y)\,dy, g\rangle = \langle K_F, g \otimes \overline{f}\rangle = \langle F, h\rangle$$

where $h \in \mathcal{S}(\mathbf{R}^{2n})$ is defined by

$$h(p,q) = \int e^{-2\pi i qx - \pi i pq}\overline{f(x+p)}g(x)dx.$$

By the Schwartz kernel theorem (see Treves [138]), every continuous linear map from $\mathcal{S}(\mathbf{R}^n)$ to $\mathcal{S}'(\mathbf{R}^n)$ is of this form. Moreover, the map $F \to K_F$ is clearly unitary on $L^2(\mathbf{R}^{2n})$, which shows that $\rho\big(L^2(\mathbf{R}^n)\big)$ is the set of Hilbert-Schmidt operators. (For background on Hilbert-Schmidt operators, see Reed–Simon [122].) In particular, $\rho(F)$ is compact for $F \in L^1 \cap L^2$, and hence for all $F \in L^1$ since $\|\rho(F)\| \leq \|F\|_1$ and the norm limit of compact operators is compact. ∎

Remark. For conditions for the operator $\rho(F)$ to be bounded on $L^p(\mathbf{R}^n)$, $1 < p < \infty$, see Mauceri [102].

For future reference, we record how the original representation $\rho(p,q)$ combines with the integrated representation $\rho(F)$. The following proposition is an easy consequence of the definitions.

Proposition (1.31). *If $F \in \mathcal{S}'(\mathbf{R}^{2n})$ and $a, b \in \mathbf{R}^n$, we have*

$$\rho(a,b)\rho(F) = \rho(G) \qquad \text{and} \qquad \rho(F)\rho(a,b) = \rho(H)$$

where

$$G(p,q) = e^{\pi i(bp-aq)}F(p-a,\ q-b) \qquad \text{and} \qquad H(p,q) = e^{\pi i(aq-bp)}F(p-a,\ q-b).$$

Twisted Convolution. We return for a moment to $L^1(\mathbf{H}_n^{\mathrm{red}})$. This space is a Banach algebra under convolution,

$$\Phi * \Psi(X) = \int \Phi(Y)\Psi(Y^{-1}X)\,dY = \int \Phi(XY^{-1})\Psi(Y)\,dY,$$

and the representation ρ is an algebra homomorphism:

$$\rho(\Phi)\rho(\Psi) = \iint \Phi(X)\Psi(Y)\rho(XY)\,dX\,dY = \rho(\Phi * \Psi).$$

We wish to transfer this algebra structure to $L^1(\mathbf{R}^{2n})$. For $F \in L^1(\mathbf{R}^{2n})$ we have

$$\rho(F) = \rho(F^0) \qquad \text{where} \qquad F^0 \in L^1(\mathbf{H}_n^{\mathrm{red}}), \quad F^0(p,q,t) = F(p,q)e^{-2\pi it},$$

and if $F, G \in L^1(\mathbf{R}^{2n})$,

$$\begin{aligned}
&F^0 * G^0(p,q,t)\\
&= \int_0^1 \iint_{\mathbf{R}^{2n}} F(p',q')e^{-2\pi it'}G(p-p',\,q-q')e^{-2\pi i(t-t')+\pi i(p'q-q'p)}dp'dq'dt'\\
&= e^{-2\pi it}\iint F(p',q')G(p-p',\,q-q')e^{\pi i(p'q-q'p)}dp'dq'.
\end{aligned}$$

That is,

$$F^0 * G^0 = (F \natural G)^0,$$

where

$$\begin{aligned}
F \natural G(p,q) &= \iint F(p',q')G(p-p',\,q-q')e^{\pi i(p'q-q'p)}dp'dq'\\
&= \iint F(p-p',\,q-q')G(p',q')e^{\pi i(pq'-qp')}dp'dq'.
\end{aligned} \tag{1.32}$$

We call $F \natural G$ the **twisted convolution** of F and G. Its definition is set up so that

$$\rho(F \natural G) = \rho(F)\rho(G).$$

Twisted convolution enjoys most of the properties of ordinary convolution on $\mathbf{R}^{2n}$ except that it is not commutative. Like ordinary convolution, it extends from L^1 to other L^p spaces and satisfies Young's inequality:

$$\|F \natural G\|_r \le \|F\|_p\|G\|_q \quad \text{when} \quad \frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1.$$

But with respect to L^p estimates, twisted convolution is even better than ordinary convolution:

(1.33) Proposition. *If $F, G \in L^2(\mathbf{R}^{2n})$ then $F \natural G \in L^2(\mathbf{R}^{2n})$ and $\|F \natural G\|_2 \leq \|F\|_2\|G\|_2$.*

Proof: This follows from the fact, observed above, that the map $F \to K_F$ defined by (1.29) is an isometry on L^2:

$$\begin{aligned}\|F \natural G\|_2 = \|K_{F \natural G}\|_2 &= \left\| \int K_F(x,y)K_G(y,z)\,dy \right\|_{L^2(x,z)} \\ &\leq \|K_F\|_2\|K_G\|_2 = \|F\|_2\|G\|_2.\end{aligned}$$

The second equality is a restatement of the fact that $\rho(F \natural G) = \rho(F)\rho(G)$, and the next estimate follows from the Schwarz inequality. ∎

One can obtain other L^p estimates for twisted convolution by interpolating between Proposition (1.33) and Young's inequality.

The Uncertainty Principle. The uncertainty principle in its general form states that if A and B are quantum observables (i.e., self-adjoint operators), the probability distributions of A and B cannot both be concentrated near single points in any state u such that $\langle (AB - BA)u, u\rangle \neq 0$. To make this precise, when μ is a probability measure on $\mathbf{R}$ we shall adopt the second moment of μ about $a \in \mathbf{R}$,

$$\left[\int (\lambda - a)^2\,d\mu(\lambda)\right]^{1/2},$$

as a measure of how much μ fails to be concentrated at a. When μ is the distribution of the observable $A = \int \lambda\,d\Pi(\lambda)$ in the state u, i.e., $\mu(E) = \langle \Pi(E)u, u\rangle$, we have

$$\left[\int (\lambda - a)^2\,d\mu(\lambda)\right]^{1/2} = \langle (A-a)^2 u,\, u\rangle^{1/2} = \|(A-a)u\|.$$

The general uncertainty principle can then be enunciated as follows.

(1.34) Theorem. *If A and B are self-adjoint operators on a Hilbert space $\mathcal{H}$, then*

$$\|(A-a)u\|\,\|(B-b)u\| \geq \tfrac{1}{2}|\langle (AB - BA)u, u\rangle|$$

for all $u \in \mathrm{Dom}(AB) \cap \mathrm{Dom}(BA)$ and all $a, b \in \mathbf{R}$. Equality holds precisely when $(A-a)u$ and $(B-b)u$ are purely imaginary scalar multiples of one another.

Proof: We have

$$\begin{aligned}\langle (AB - BA)u,\, u\rangle &= \langle [(A-a)(B-b) - (B-b)(A-a)]u,\, u\rangle \\ &= \langle (B-b)u,\, (A-a)u\rangle - \langle (A-a)u,\, (B-b)u\rangle \\ &= 2i\,\mathrm{Im}\langle (B-b)u,\, (A-a)u\rangle\end{aligned}$$

and hence

$$\langle (AB - BA)u,\, u\rangle \le 2\big|\langle (B-b)u,\, (A-a)u\rangle\big| \le 2\|(A-a)u\|\, \|(B-b)u\|.$$

The first inequality is an equality precisely when $\langle (B-b)u,\, (A-a)u\rangle$ is imaginary, and the second one is an equality precisely when $(A-a)u$ and $(B-b)u$ are linearly dependent. ∎

If we apply this result to the position and momentum operators X and D on $L^2(\mathbf{R})$, we obtain:

(1.35) Corollary. *If $u \in L^2(\mathbf{R})$ and $a, b \in \mathbf{R}$ we have*

$$\|(X-a)u\|_2 \|(D-b)u\|_2 \ge \frac{1}{4\pi}\|u\|_2, \tag{1.36}$$

with equality if and only if

$$u(x) = ce^{2\pi ibx}e^{-\pi r(x-a)^2} \qquad \textit{for some} \qquad c \in \mathbf{C}, \quad r > 0.$$

Proof: The inequality is valid by Theorem (1.34) since $[D, X] = (2\pi i)^{-1}I$. (The preceding proof works when $u \in \mathrm{Dom}(DX) \cap \mathrm{Dom}(XD)$, but the result remains valid for all $u \in L^2$, with the understanding that, for example, $\|(D-b)u\|_2 = \infty$ if $u \notin \mathrm{Dom}(D)$. This may be established by an approximation argument which we leave to the reader.) Equality holds iff

$$u'(x) - 2\pi ibu(x) = 2\pi r(x-a)u(x)$$

for some real r, and the solutions of this differential equation are the Gaussians described above. ∎

Another interesting variant of the uncertainty principle is the following:

(1.37) Corollary. *If $u \in L^2(\mathbf{R})$, we have*

$$\|Xu\|_2^2 + \|Du\|_2^2 \ge \frac{1}{2\pi}\|u\|_2^2, \tag{1.38}$$

with equality if and only if $u(x) = ce^{-\pi x^2}$.

Proof: (1.38) follows from (1.36) (with $a = b = 0$) together with the numerical inequality $\alpha\beta \le \frac{1}{2}(\alpha^2 + \beta^2)$ $(\alpha, \beta \ge 0)$. Equality holds here iff $\alpha = \beta$, which forces $r = 1$ above. ∎

(1.36) and (1.38) are actually equivalent. To deduce (1.36) from (1.38), apply (1.38) to the functions $u_\alpha(x) = \alpha^{1/2}u(\alpha x)$, $\alpha > 0$. Since $\|Xu_\alpha\|_2 = \alpha^{-1}\|Xu\|_2$ and $\|Du_\alpha\|_2 = \alpha\|Du\|_2$, the result is

$$\alpha^{-2}\|Xu\|_2^2 + \alpha^2\|Du\|_2^2 \geq \frac{1}{2\pi}\|u\|_2^2.$$

Minimizing the left side over all $\alpha > 0$ yields (1.36) with $a = b = 0$. Applying the latter result to the function $v(x) = e^{-2\pi i b(x+a)}u(x+a)$, one obtains (1.36) in general.

These results generalize in the obvious way to n dimensions. Namely, if $u \in L^2(\mathbf{R}^n)$,

$$\|(X_j - a_j)u\|_2\|(D_j - b_j)u\|_2 \geq \frac{1}{4\pi}\|u\|_2^2 \qquad (a_j, b_j \in \mathbf{R}), \tag{1.39}$$

$$\|X_ju\|_2^2 + \|D_ju\|_2^2 \geq \frac{1}{2\pi}\|u\|_2^2. \tag{1.40}$$

Equality holds in (1.39) [resp. (1.40)] for a fixed j iff

$$u(x) = v(x)\exp(2\pi i b_j x_j - \pi r(x_j - a_j)^2) \qquad [\text{resp.} \quad u(x) = v(x)\exp(-\pi x_j^2)\,]$$

where v is independent of x_j, and it holds in (1.39) [resp. (1.40)] for all j iff

$$u(x) = c\exp\sum_1^n(2\pi i b_j x_j - \pi r_j(x_j - a_j)^2) \qquad [\text{resp.} \quad u(x) = c\exp(-\pi x^2)\,].$$

There are a number of other versions of the uncertainty principle in the literature: see de Bruijn [37] and the papers [34], [35], [119], [120], and [121] of Cowling, Price, and Sitaram (in various combinations).

One of the recurring themes of this monograph is the beauty and importance of the Gaussian functions

$$f(x) = e^{xAx+bx+c}.$$

(Here A is an $n \times n$ complex matrix with $\operatorname{Re} A$ negative definite, $b \in \mathbf{C}^n$, and $c \in \mathbf{C}$.) We have just seen the first indication of their fundamental nature, in the fact that when A is real and diagonal they are precisely the extremal functions for the uncertainty inequalities. In fact, every Gaussian is an extremal for the uncertainty inequalities for some set of operators $\{D_j', X_j'\}_{j=1}^n$ obtained from $\{D_j, X_j\}_{j=1}^n$ by a symplectic linear transformation. (We shall explain this in detail in Section 4.5.) We shall see that the Gaussians play a special role in a number of other contexts. For the moment, we merely point out two sobriquets that Gaussians have acquired from their scientific applications. In the one-dimensional case, the functions

$$f(t) = e^{2\pi i\omega t}e^{-\pi a(t-\tau)^2} \qquad (t,\omega,\tau \in \mathbf{R}, \quad a > 0)$$

are known as **Gabor functions**, after a paper of Gabor [53] in which their utility as simple components for building electrical signals was demonstrated. (See Section 3.4.) Also, the functions

$$f(x) = \rho(p,q)\big[2^{n/4}e^{-\pi x^2}\big] = e^{2\pi iqx+\pi ipq}e^{-\pi(x+p)^2},$$

obtained by translating the basic Gaussian $2^{n/4}e^{-\pi x^2}$ in phase space, are known in quantum physics, and especially quantum optics, as **coherent states**.

4. The Fourier-Wigner Transform

In this section we study the matrix coefficients of the representation ρ. If $f, g \in L^2(\mathbf{R}^n)$, the matrix coefficient of ρ at (f,g) is the function M on $\mathbf{H}_n$ (or $\mathbf{H}_n^{\text{red}}$) defined by

$$M(p,q,t) = \langle \rho(p,q,t)f, g\rangle.$$

Clearly $M(p,q,t) = e^{2\pi it}M(p,q,0)$, so the t dependence carries no information and can best be ignored. Accordingly, for $f,g \in L^2(\mathbf{R}^n)$, we define the function $V(f,g)$ on $\mathbf{R}^{2n}$ by

$$\begin{aligned} (1.41) \qquad V(f,g)(p,q) &= \langle \rho(p,q)f, g\rangle = \langle e^{2\pi i(pD+qX)}f, g\rangle \\ &= \int e^{2\pi iqx+\pi ipq} f(x+p)\overline{g(x)}\,dx \\ &= \int e^{2\pi iqy} f(y+\tfrac{1}{2}p)\overline{g(y-\tfrac{1}{2}p)}\,dy. \end{aligned}$$

The map V has no standard name; we shall call it the **Fourier-Wigner transform**, for reasons that will become clear in Section 1.8. It is clear from the Schwarz inequality that $V(f,g)$ is always a bounded, continuous function on $\mathbf{R}^{2n}$ satisfying $\|V(f,g)\|_\infty \le \|f\|_2\|g\|_2$.

V can be extended in an obvious way from a sesquilinear map defined on $L^2(\mathbf{R}^n) \times L^2(\mathbf{R}^n)$ to a linear map $\widetilde{V}$ defined on the tensor product $L^2(\mathbf{R}^n) \otimes L^2(\mathbf{R}^n)$, which is naturally isomorphic to $L^2(\mathbf{R}^{2n})$. Namely, if $F \in L^2(\mathbf{R}^{2n})$ we define

$$\widetilde{V}(F)(p,q) = \int e^{2\pi iqy}F(y+\tfrac{1}{2}p, y-\tfrac{1}{2}p)\,dy.$$

We then have $V(f,g) = \widetilde{V}(f \otimes \overline{g})$, where $f \otimes \overline{g}(x,y) = f(x)\overline{g}(y)$. $\widetilde{V}$ is the composition of the measure-preserving change of variables $(y,p) \to (y + \frac{1}{2}p,\ y - \frac{1}{2}p)$ with inverse Fourier transformation in the first variable. Therefore it is unitary on $L^2(\mathbf{R}^{2n})$, maps $\mathcal{S}(\mathbf{R}^{2n})$ onto itself, and extends to a continuous bijection of $\mathcal{S}'(\mathbf{R}^{2n})$ onto itself. Transferring these results back to V, we obtain the following:

(1.42) Proposition. *V maps $\mathcal{S}(\mathbf{R}^n) \times \mathcal{S}(\mathbf{R}^n)$ into $\mathcal{S}(\mathbf{R}^{2n})$ and extends to a map from $\mathcal{S}'(\mathbf{R}^n) \times \mathcal{S}'(\mathbf{R}^n)$ into $\mathcal{S}'(\mathbf{R}^{2n})$. Moreover, V is "sesqui-unitary" on L^2; that is, for all $f_1, g_1, f_2, g_2 \in L^2(\mathbf{R}^n)$,*

$$\langle V(f_1,g_1), V(f_2,g_2)\rangle = \langle f_1, f_2\rangle\overline{\langle g_1, g_2\rangle}.$$

In the language of representation theory, this proposition says that ρ is square-integrable (modulo the center). The irreducibility of ρ is an easy corollary:

(1.43) Proposition. *The representation ρ_h is irreducible for any $h \in \mathbf{R}\backslash\{0\}$.*

Proof: Suppose $\mathcal{M} \subset L^2(\mathbf{R}^n)$ is a nonzero closed invariant subspace and $f \neq 0 \in \mathcal{M}$. If $g \perp \mathcal{M}$ then $g \perp e^{2\pi i(hpD+qX)}f$ for all $p, q \in \mathbf{R}^n$; in other words, $V(f,g) = 0$. But this implies that $\|f\|_2\|g\|_2 = 0$, whence $g = 0$ and $\mathcal{M} = L^2(\mathbf{R}^n)$. ∎

Here is what happens to $V(f,g)$ when f and g are transformed by the operators $\rho(a,b)$.

(1.44) Proposition. *For any $a, b, c, d \in \mathbf{R}^n$ we have*

$$\text{(a)} \quad \begin{aligned} &V\big(\rho(a,b)f, \rho(c,d)g\big)(p,q) \\ &\qquad = e^{\pi i(dp+da+pb-cq-cb-qa)}V(f,g)(p+a-c,\ q+b-d). \end{aligned}$$

In particular,

$$\text{(b)} \quad V\big(\rho(a,b)f,\ g\big)(p,q) = e^{\pi i(pb-qa)}V(f,g)(p+a,\ q+b)$$

$$\text{(c)} \quad V\big(f,\ \rho(c,d)g\big)(p,q) = e^{\pi i(dp-cq)}V(f,g)(p-c,\ q-d)$$

$$\text{(d)} \quad V\big(\rho(a,b)f,\ \rho(a,b)g\big)(p,q) = e^{2\pi i(pb-qa)}V(f,g)(p,q)$$

Proof: We have

$$V\big(\rho(a,b)f,\ \rho(c,d)g\big)(p,q) = \langle\rho(-c,-d)\rho(p,q)\rho(a,b)f,\ g\rangle$$

and, in $\mathbf{H}_n$,

$$(-c,-d,0)(p,q,0)(a,b,0) = \big(p+a-c,\ q+b-d,\ \tfrac{1}{2}(dp+da+pb-cq-cb-qa)\big).$$

(a) follows easily from these equations, and (b), (c), and (d) are special cases of (a). ∎

The matrix elements of the integrated representation can also be expressed in terms of the Fourier-Wigner transform. Indeed, we have

(1.45)
$$\langle\rho(F)f,g\rangle = \iint F(p,q)\langle\rho(p,q)f,g\rangle\,dp\,dq = \iint F(p,q)V(f,g)(p,q)\,dp\,dq.$$

An interesting thing happens when we use the conjugate of a Fourier-Wigner transform as input for the representation ρ:

(1.46) Proposition. *If $\phi,\psi\in L^2(\mathbf{R}^n)$ and $\Phi=\overline{V(\phi,\psi)}$ then*

$$\rho(\Phi)f=\langle f,\phi\rangle\psi \qquad \text{for} \quad f\in L^2(\mathbf{R}^n).$$

Proof: By (1.45) and Proposition (1.42), we have

$$\begin{aligned}\langle\rho(\Phi)f,g\rangle &= \int \overline{V(\phi,\psi)}V(f,g) = \langle V(f,g),V(\phi,\psi)\rangle\\ &= \langle f,\phi\rangle\overline{\langle g,\psi\rangle} = \langle f,\phi\rangle\langle\psi,g\rangle,\end{aligned}$$

whence the result is immediate. ∎

In other words, the operators $\rho(\Phi)$ where $\overline{\Phi}$ is a Fourier-Wigner transform are precisely the operators on L^2 with one-dimensional range. This leads to a nice formula for the twisted convolution of Fourier-Wigner transforms:

(1.47) Proposition. $\overline{V(\phi_1,\psi_1)}\natural\overline{V(\phi_2,\psi_2)}=\langle\psi_2,\phi_1\rangle\overline{V(\phi_2,\psi_1)}$.

Proof: Let $\Phi_j=\overline{V(\phi_j,\psi_j)}$ and $\Psi=\overline{V(\phi_2,\psi_1)}$. Then

$$\rho(\Phi_1)\rho(\Phi_2)f=\rho(\Phi_1)\langle f,\phi_2\rangle\psi_2=\langle f,\phi_2\rangle\langle\psi_2,\phi_1\rangle\psi_1=\langle\psi_2,\phi_1\rangle\rho(\Psi)f.$$

But $\rho(\Phi_1)\rho(\Phi_2)=\rho(\Phi_1\natural\Phi_2)$ and ρ is faithful, so $\Phi_1\natural\Phi_2=\langle\psi_2,\phi_1\rangle\Psi$. ∎

We conclude this discussion with some calculations of Fourier-Wigner transforms of Gaussians that we shall need later.

(1.48) Proposition. *Let*

$$\phi(x)=2^{n/4}e^{-\pi x^2}, \qquad \Phi=V(\phi,\phi), \qquad \Phi^{ab}=V\big(\phi,\rho(a,b)\phi\big).$$

Then

(a) $$\Phi(p,q)=e^{-(\pi/2)(p^2+q^2)}.$$

(b) $$\Phi^{ab}(p,q)=e^{\pi i(bp-aq)}e^{-(\pi/2)[(p-a)^2+(q-b)^2]}.$$

(c) $$\rho(\Phi)\rho(a,b)\rho(\Phi)=e^{-(\pi/2)(a^2+b^2)}\rho(\Phi).$$

(d) $$\Phi\natural\Phi^{ab}=e^{(-\pi/2)(a^2+b^2)}\Phi.$$

Proof: (a) follows from the Fourier transform formula for Gaussians (Appendix A):

$$\Phi(p,q) = 2^{n/2}\int e^{2\pi i q y} e^{-\pi[(y+(p/2)]^2 - \pi[(y-(p/2))]^2}\,dy$$
$$= 2^{n/2} e^{(-\pi/2)p^2}\int e^{2\pi i q y} e^{-2\pi y^2}\,dy = e^{-(\pi/2)(p^2+q^2)}.$$

(b) follows from (a) and Proposition (1.44c). As for (c), by Proposition (1.46),

$$\rho(\Phi)\rho(p,q)\rho(\Phi)f = \rho(\Phi)\big[\langle f,\phi\rangle\rho(a,b)\phi\big]$$
$$= \langle f,\phi\rangle\langle\rho(a,b)\phi,\phi\rangle\phi = \langle f,\phi\rangle\Phi(a,b)\phi = e^{-(\pi/2)(a^2+b^2)}\rho(\Phi)f.$$

Finally, by Proposition (1.31) and (b) we have $\rho(a,b)\rho(\Phi) = \rho(\Phi^{ab})$, so by (c), $\rho(\Phi \natural \Phi^{ab}) = e^{-(\pi/2)(a^2+b^2)}\rho(\Phi)$. Since ρ is faithful, (d) follows. ∎

Radar Ambiguity Functions. We shall give a brief account of how Fourier-Wigner transforms turn up in the theory of radar.

A radar apparatus transmits an electromagnetic signal that reflects off a target and returns to the apparatus. The signal may be represented by a complex function of time, $f(t)$. [Technically, $f(t) = u(t) + iv(t)$ where u is the amplitude of the physical signal and v is the Hilbert transform of u. The energy of the signal is $\frac{1}{2}\|f\|_2^2$.] We assume that the frequencies of f are concentrated around some (large) number ω, so that we can write $f(t) = f_0(t)e^{2\pi i\omega t}$ where $f_0(t)$ is slowly varying in comparison to $e^{2\pi i\omega t}$. Let r be the distance of the target from the apparatus and $v = dr/dt$ its radial velocity. We assume that the signal f is essentially limited to a time interval Δt which is large in comparison to the period ω^{-1} but small enough so that r and v may be considered as constant in this interval. The reflected signal then arrives back at the apparatus after a time delay $\tau = 2r/c$ (where c is the propagation speed of the signal) and with frequencies dilated by a factor $1-(2v/c)$ because of the Doppler effect. Since the frequencies of the transmitted signal are mostly near ω, we may assume instead that the frequencies are shifted by the amount $-\phi = -2\omega v/c$. In short, when these approximations have been made, the returning signal is

$$f_{\tau\phi}(t) = f(t-\tau)e^{-2\pi i\phi t}.$$

Now suppose there are two targets that produce returning signals $f_{\tau_1\phi_1}$ and $f_{\tau_2\phi_2}$. If these signals are similar to each other there will be a difficulty in distinguishing the two targets, so we are concerned with the mean squared difference,

$$\text{(1.49)}\qquad \int |f_{\tau_1\phi_1} - f_{\tau_2\phi_2}|^2\,dt = 2\int |f|^2\,dt - 2\mathrm{Re}\langle f_{\tau_1\phi_1}, f_{\tau_2\phi_2}\rangle.$$

Only the second term on the right depends on the targets, and since $f(t) = f_0(t)e^{2\pi i\omega t}$ we have

$$\langle f_{\tau_1\phi_1}, f_{\tau_2\phi_2}\rangle = e^{2\pi i\omega(\tau_2-\tau_1)} \int f_0(t-\tau_1)\overline{f_0(t-\tau_2)}e^{2\pi i(\phi_2-\phi_1)t}\,dt.$$

The integral varies slowly with τ_1 and τ_2, but the exponential in front is rapidly oscillating; so if we want (1.49) to be large in a way that is stable under small perturbations of τ_1 and τ_2, $|\langle f_{\tau_1\phi_1}, f_{\tau_2\phi_2}\rangle|$ must be near zero. If we set $\tau = \tau_1 - \tau_2$ and $\phi = \phi_1 - \phi_2$, we have

$$\langle f_{\tau_1\phi_1}, f_{\tau_2\phi_2}\rangle = e^{-2\pi i\phi\tau_1} \int f(t)\overline{f(t+\tau)}e^{-2\pi i\phi t}\,dt.$$

Since only the absolute value is important, and since nothing essential is changed by switching the two targets, we could equally well consider

$$A_1(\tau,\phi) = \int f(t)\overline{f(t+\tau)}e^{-2\pi i\phi t}\,dt$$

or $$A_2(\tau,\phi) = e^{-\pi i\phi\tau}A_1(\tau,\phi)$$

or $$A_3(\tau,\phi) = A_2(-\tau,-\phi) = \int f(t)\overline{f(t-\tau)}e^{2\pi i\phi t-\pi i\phi\tau}dt,$$

and we have

$$A_3 = V(f,f).$$

A_1, A_2, and A_3 and the squares of their absolute values are all referred to in one place or another as the **ambiguity function** of the signal f. Whichever variant is used, the intuitive significance is that if $|A_j(\tau,\phi)|$ is large, two targets whose associated time and frequency shifts differ by τ and ϕ will be hard to distinguish. In this connection, we observe that by the Schwarz inequality,

$$|A_j(\tau,\phi)| \le A_j(0,0) = \|f\|_2^2,$$

and by Proposition (1.42),

$$\iint |A_j(\tau,\phi)|^2\,d\tau\,d\phi = \|f\|_2^4.$$

This last equation may be interpreted as saying that for a signal f of fixed energy $\frac{1}{2}\|f\|_2^2$, there is a fixed amount of ambiguity distributed over the (τ,ϕ) plane that cannot be eliminated. This is sometimes called "conservation of ambiguity" or the "radar uncertainty principle."

Ambiguity functions were introduced into radar theory by Woodward [157], and the connection with the Wigner transform (cf. Section 1.8) was noted by Klauder [90]. A detailed account of the use of ambiguity functions in radar design can be found in Cook–Bernfeld [31]. Recently there has been interest in analyzing ambiguity functions by explicit use of the connection with the Heisenberg group: see Auslander–Tolimieri [9] and Schempp [125].

5. The Stone–von Neumann Theorem

We have constructed a family $\{\rho_h : h \in \mathbf{R}\backslash\{0\}\}$ of irreducible unitary representations of $\mathbf{H}_n$. We now prove the classic theorem of Stone [133] and von Neumann [146], which says in effect that any irreducible unitary representation of $\mathbf{H}_n$ that is nontrivial on the center is equivalent to some ρ_h. Since the irreducible representations that are trivial on the center are easily described, as we shall see below, we shall obtain a complete classification of the irreducible unitary representations of $\mathbf{H}_n$.

Nowadays the Stone–von Neumann theorem is usually obtained as a corollary of the Mackey imprimitivity theorem. Here we present von Neumann's original proof, a pretty argument that does not deserve the obscurity into which it has fallen. It actually does more than classify the irreducible representations: it also shows that any primary representation of $\mathbf{H}_n$ is a direct sum of copies of an irreducible representation, and hence that $\mathbf{H}_n$ is a type I group.

(1.50) The Stone–von Neumann Theorem. *Let π be a unitary representation of $\mathbf{H}_n$ on a Hilbert space $\mathcal{H}$, such that $\pi(0,0,t) = e^{2\pi iht}I$ for some $h \in \mathbf{R}\backslash\{0\}$. Then $\mathcal{H} = \bigoplus \mathcal{H}_\alpha$ where the $\mathcal{H}_\alpha$'s are mutually orthogonal subspaces of $\mathcal{H}$, each invariant under π, such that $\pi \,|\, \mathcal{H}_\alpha$ is unitarily equivalent to ρ_h for each α. In particular, if π is irreducible then π is equivalent to ρ_h.*

Proof: We present the proof for $h = 1$; the argument in general is exactly the same. The crucial point is to identify the elements of $\mathcal{H}$ that correspond to the Gaussian $e^{-\pi x^2}$ in the Schrödinger representation. Concerning the latter, we adopt the following notation:

$$\phi(x) = 2^{n/4}e^{-\pi x^2}, \qquad \phi^{ab}(x) = \rho(a,b)\phi(x) = 2^{n/4}e^{2\pi ibx+\pi iab}e^{-\pi(x+a)^2},$$

$$\Phi = V(\phi,\phi), \qquad \Phi^{ab} = V(\phi,\phi^{ab}).$$

By Proposition (1.48), we then have

$$(1.51) \qquad \langle \phi^{pq}, \phi^{ab}\rangle = \Phi^{ab}(p,q) = e^{\pi i(bp-aq)}e^{-(\pi/2)[(p-a)^2+(q-b)^2]},$$

$$(1.52) \qquad \Phi \natural \Phi^{ab} = e^{-(\pi/2)(a^2+b^2)}\Phi.$$

Returning to the representation π, we mimic some constructions that we made with ρ in Section 1.3. First we set $\pi(p,q) = \pi(p,q,0)$, and we have

$$(1.53) \quad \pi(p,q)\pi(r,s) = \pi\big(p+r,\ q+s,\ \tfrac{1}{2}(ps-qr)\big) = e^{\pi i(ps-qr)}\pi(p+r,\ q+s).$$

We consider the integrated version of π,

$$\pi(F) = \iint F(p,q)\pi(p,q)\,dp\,dq \qquad (F \in L^1(\mathbf{R}^{2n})),$$

and just as with ρ, we have

(1.54) $\pi(F)\pi(G) = \pi(F \natural G)$,

(1.55) $\pi(F)\pi(a,b) = \pi(G)$ where $G(p,q) = e^{\pi i(aq-bp)}F(p-a,\, q-b)$,

(1.56) $\pi(a,b)\pi(F) = \pi(H)$ where $H(p,q) = e^{\pi i(bp-aq)}F(p-a,\, q-b)$.

Moreover, π is faithful on $L^1(\mathbf{R}^{2n})$. Indeed, if $\pi(F) = 0$ then, by (1.55) and (1.56), for any $u, v \in \mathcal{H}$ and $a, b \in \mathbf{R}^n$,

$$0 = \langle \pi(a,b)\pi(F)\pi(-a,-b)u, v\rangle$$
$$= \iint e^{2\pi i(bp-aq)}F(p,q)\langle \pi(p,q)u, v\rangle\, dp\, dq.$$

Thus by the Fourier inversion theorem,

$$F(p,q)\langle \pi(p,q)u, v\rangle = 0 \qquad \text{for a.e. } (p,q),$$

and since u and v are arbitary, $F = 0$ a.e.

Now let us take F to be the function Φ defined above. By (1.51), (1.52), (1.54), and (1.56),

$$\pi(\Phi)\pi(a,b)\pi(\Phi) = \pi(\Phi \natural \Phi^{ab}) = e^{-(\pi/2)(a^2+b^2)}\pi(\Phi). \tag{1.57}$$

In particular, taking $a = b = 0$ we obtain $\pi(\Phi)^2 = \pi(\Phi)$, and since Φ is even and real it is easily seen that $\pi(\Phi)$ is self-adjoint. Thus $\pi(\Phi)$ is an orthogonal projection which is nonzero since $\Phi \neq 0$ and π is faithful. Let $\mathcal{R}$ denote the range of $\pi(\Phi)$. If $u, v \in \mathcal{R}$ then $u = \pi(\Phi)u$ and $v = \pi(\Phi)v$, so by (1.53),

$$\begin{aligned} \langle \pi(p,q)u,\, \pi(r,s)v\rangle &= \langle \pi(-r,-s)\pi(p,q)\pi(\Phi)u,\, \pi(\Phi)v\rangle \\ &= e^{\pi i(ps-qr)}\langle \pi(\Phi)\pi(p-r, q-s)\pi(\Phi)u,\, v\rangle \\ &= e^{\pi i(ps-qr)}e^{-(\pi/2)[(p-r)^2+(q-s)^2]}\langle u, v\rangle. \end{aligned} \tag{1.58}$$

Let $\{v_\alpha\}$ be an orthonormal basis for $\mathcal{R}$, and let $\mathcal{H}_\alpha$ be the closed linear span of $\{\pi(p,q)v_\alpha : p, q \in \mathbf{R}^n\}$. By (1.58), $\mathcal{H}_\alpha \perp \mathcal{H}_\beta$ for $\alpha \neq \beta$, and $\mathcal{H}_\alpha$ is invariant under π by definition. Hence $\mathcal{N} = \left(\bigoplus \mathcal{H}_\alpha\right)^\perp$ is also invariant under π, and we have $\pi(\Phi)\,|\,\mathcal{N} = 0$. But this implies that $\mathcal{N} = \{0\}$, for otherwise we could apply the above reasoning to $\pi\,|\,\mathcal{N}$ to conclude that $\pi(\Phi)\,|\,\mathcal{N}$ were a nonzero orthogonal projection.

We claim that $\pi\,|\,\mathcal{H}_\alpha$ is equivalent to ρ for all α. Indeed, fix an α and let $v^{pq} = \pi(p,q)v_\alpha$. Then by (1.52) and (1.58),

$$\langle v^{pq}, v^{rs}\rangle = \langle \phi^{pq}, \phi^{rs}\rangle \qquad \text{for all } p, q, r, s.$$

It follows that if $u = \sum a_{jk}v^{p_j q_k}$ and $f = \sum a_{jk}\phi^{p_j q_k}$ then $\|u\|_{\mathcal{H}} = \|f\|_2$, and in particular $u = 0$ iff $f = 0$. Therefore the correspondence $v^{pq} \to \phi^{pq}$ extends by linearity and continuity to a unitary map from $\mathcal{H}_\alpha$ to $L^2(\mathbf{R}^n)$ that intertwines $\pi\,|\,\mathcal{H}_\alpha$ and ρ. ∎

We can now give a complete classification of the irreducible unitary representations of $\mathbf{H}_n$. Suppose π is such a representation. By Schur's lemma (cf. Appendix B), π must map the center $\mathcal{Z}$ of $\mathbf{H}_n$ homomorphically into the group $\{cI : |c| = 1\}$, so $\pi(0,0,t) = e^{2\pi iht}I$ for some $h \in \mathbf{R}$. If $h \neq 0$, the Stone–von Neumann theorem shows that π is equivalent to ρ_h. If $h = 0$, on the other hand, π factors through the quotient group $\mathbf{H}_n/\mathcal{Z}$, which is isomorphic to $\mathbf{R}^{2n}$. The irreducible representations of the latter are all one-dimensional (Schur's lemma again) and hence are just the homomorphisms from $\mathbf{R}^{2n}$ into the circle group, namely, $(p,q) \to e^{2\pi i(ap+bq)}$. We have therefore proved:

(1.59) Theorem. *Every irreducible unitary representation of* $\mathbf{H}_n$ *is unitarily equivalent to one and only one of the following representations:*
(a) ρ_h $(h \in \mathbf{R}\setminus\{0\})$, *acting on* $L^2(\mathbf{R}^n)$,
(b) $\sigma_{ab}(p,q,t) = e^{2\pi i(ap+bq)}$ $(a, b \in \mathbf{R}^n)$, *acting on* $\mathbf{C}$.

The Group Fourier Transform. If G is a locally compact group, let $\widehat{G}$ denote a collection of irreducible unitary representations of G containing exactly one member of each equivalence class. If $\pi \in \widehat{G}$ we denote by $\mathcal{H}_\pi$ the Hilbert space on which π acts. Given $f \in L^1(G)$ and $\pi \in \widehat{G}$, we define the operator $\widehat{f}(\pi)$ on $\mathcal{H}_\pi$ by

$$\widehat{f}(\pi) = \int_G f(x)\pi(x)^*\,dx = \int_G f(x)\pi(x^{-1})\,dx,$$

where dx denotes Haar measure. The map $f \to \widehat{f}$ is called the **group Fourier transform**. For a large class of groups G there exists a measure μ on $\widehat{G}$ (the "Plancherel measure") such that for all sufficiently nice functions f on G one has the Fourier inversion formula

$$f(x) = \int_{\widehat{G}} \operatorname{tr}\big(\widehat{f}(\pi)\pi(x)\big)\,d\mu(\pi) \tag{1.60}$$

and the Plancherel formula

$$\int_G |f(x)|^2\,dx = \int_{\widehat{G}} \operatorname{tr}\big(\widehat{f}(\pi)^*\widehat{f}(\pi)\big)\,d\mu(\pi) = \int_{\widehat{G}} \|\widehat{f}(\pi)\|_{HS}^2\,d\mu(\pi). \tag{1.61}$$

(Here tr denotes trace and $\|\cdot\|_{HS}$ denotes the Hilbert-Schmidt norm.)

We now compute the Plancherel measure for $\mathbf{H}_n$, using the parametrization of $\widehat{\mathbf{H}_n}$ given by Theorem (1.59). Our analysis will show that we need only consider the representations ρ_h; that is, the one-dimensional representations σ_{ab} form a set of Plancherel measure zero.

If $f \in L^1(\mathbf{H}_n)$, $h \in \mathbf{R}\backslash\{0\}$, and $\phi \in L^2(\mathbf{R}^n)$, $\widehat{f}(\rho_h)\phi$ is given by

$$\begin{aligned}
\widehat{f}(\rho_h)\phi(x) &= \iiint f(p,q,t)\rho_h(-p,-q,-t)\phi(x)\,dp\,dq\,dt \\
&= \iiint f(p,q,t)e^{-2\pi iqx+\pi ihpq-2\pi iht}\phi(x-hp)\,dp\,dq\,dt \\
&= |h|^{-n}\iiint f(h^{-1}(x-y),\,q,\,t)e^{-\pi i(y+x)q-2\pi iht}\phi(y)\,dy\,dq\,dt.
\end{aligned}$$

Thus, $\widehat{f}(\rho_h)$ is an integral operator with kernel

$$\begin{aligned}
K_f^h(x,y) &= |h|^{-n}\iint f(h^{-1}(x-y),\,q,\,t)e^{-\pi i(y+x)q-2\pi iht}\,dq\,dt \\
&= |h|^{-n}\mathcal{F}_{2,3}f\big(h^{-1}(x-y),\,\tfrac{1}{2}(x+y),\,h\big),
\end{aligned} \tag{1.62}$$

where $\mathcal{F}_{2,3}$ denotes Fourier transformation in the second and third variables. Moreover,

$$\begin{aligned}
\widehat{f}(\rho_h)\rho_h(p,q,t) &= \iiint f(p',q',t')\rho_h(-p',-q',-t')\rho_h(p,q,t)\,dp'dq'dt' \\
&= \iiint f(p',q',t')\rho_h\big(p-p',\,q-q',\,t-t'-\tfrac{1}{2}(q'p-p'q)\big)dp'dq'dt' \\
&= \widehat{g}(\rho_h)
\end{aligned}$$

where

$$g(p',q',t') = f(p-p',\,q-q',\,t-t')e^{\pi ih(p'q-q'p)}.$$

Hence, in view of (1.62), the integral kernel of $\widehat{f}(\rho_h)\rho_h(p,q,t)$ is

$$\begin{aligned}
&F(x,y) \\
&= |h|^{-n}\iint f(p-h^{-1}(x-y),\,q-q',\,t-t')e^{\pi i[(x-y)q-q'p]-\pi i(y+x)q'-2\pi iht'}\,dq'dt' \\
&= |h|^{-n}\iint f(p-h^{-1}(x-y),q',t')e^{\pi i[(x-y)q-(q-q')p]-\pi i(x+y)(q-q')-2\pi ih(t-t')}dq'dt'.
\end{aligned}$$

If f is such that all the integrals converge nicely, then, we have

$$\begin{aligned}
\operatorname{tr}\big(\widehat{f}(\rho_h)\rho_h(p,q,t)\big) &= \int F(x,x)dx \\
&= |h|^{-n}\iiint f(p,q',t')e^{\pi ip(q-q')-2\pi ix(q-q')-2\pi ih(t-t')}dq'\,dt'\,dx \\
&= |h|^{-n}\iint f(p,q',t')e^{\pi ip(q-q')-2\pi ih(t-t')}\delta(q-q')\,dq'dt \\
&= |h|^{-n}\int f(p,q,t')e^{-2\pi ih(t-t')}\,dt'.
\end{aligned}$$

But by the (ordinary) Fourier inversion formula,

$$f(p,q,t)=\iint f(p,q,t')e^{-2\pi ih(t-t')}\,dt'\,dh=\int \operatorname{tr}\big(\widehat{f}(\rho_h)\rho_h(p,q,t)\big)|h|^n\,dh.$$

Thus (1.60) holds if we define the Plancherel measure on $\widehat{\mathbf{H}_n}$ to be $|h|^n\,dh$ on the family $\{\rho_h\}$ and 0 on the family $\{\sigma_{ab}\}$. Moreover, by (1.62) and the (ordinary) Plancherel theorem,

$$\begin{aligned}\|\widehat{f}(\rho_h)\|_{HS}^2&=\int |K_f^h(x,y)|^2\,dx\,dy\\&=|h|^{-2n}\iint \big|\mathcal{F}_{2,3}f\big(h^{-1}(x-y),\ \tfrac12(x+y),\ h\big)\big|^2\,dx\,dy\\&=|h|^{-n}\iint |\mathcal{F}_{2,3}f(p,z,h)|^2\,dp\,dz\\&=|h|^{-n}\int |\mathcal{F}_3 f(p,q,h)|^2\,dp\,dq,\end{aligned}$$

so that (1.61) also holds:

$$\|f\|_2^2=\int |h|^n\|\widehat{f}(\rho_h)\|_{HS}^2\,dh.$$

There is much more that can be said about the group Fourier transform on $\mathbf{H}_n$; see Geller [55], [56], [57].

6. The Fock–Bargmann Representation

There is a particularly interesting realization of the infinite-dimensional irreducible unitary representations of $\mathbf{H}_n$ in a Hilbert space of entire functions. We shall carry out the analysis for the representation ρ and indicate at the end how to generalize to ρ_h.

Let

$$\phi_0(x)=2^{n/4}e^{-\pi x^2}$$

be the standard Gaussian on $\mathbf{R}^n$. Since $\|\phi_0\|_2=1$, by Proposition (1.42) the map $f\to V(f,\phi_0)$ is an isometry from $L^2(\mathbf{R}^n)$ into $L^2(\mathbf{R}^{2n})$. Explicitly, we have

$$\begin{aligned}V(f,\phi_0)(p,q)&=\langle f,\ \rho(-p,-q)\phi_0\rangle\\&=2^{n/4}\int f(x)e^{2\pi iqx-\pi ipq}e^{-\pi(x-p)^2}\,dx\\&=2^{n/4}e^{-(\pi/2)(p^2+q^2)}\int f(x)e^{2\pi x(p+iq)-\pi x^2-(\pi/2)(p+iq)^2}\,dx.\end{aligned}$$

For $z \in \mathbf{C}^n$ let us define

$$Bf(z) = 2^{n/4} \int f(x) e^{2\pi x z - \pi x^2 - (\pi/2) z^2} \, dx.$$

Then we have

$$V(f, \phi_0)(p, q) = e^{-(\pi/2)|z|^2} Bf(z), \qquad \text{with} \quad z = p + iq.$$

Bf is called the **Bargmann transform** of f. For $f \in L^2$, the integral defining $Bf(z)$ plainly converges uniformly for z in any compact subset of $\mathbf{C}^n$, so that Bf is an entire analytic function on $\mathbf{C}^n$. Moreover, since the map $f \to V(f, \phi_0)$ is an isometry on L^2, B is an isometry from $L^2(\mathbf{R}^n)$ into $L^2(\mathbf{C}^n,\ e^{-\pi|z|^2} dz)$. (Here and in the sequel, dz denotes Lebesgue measure on $\mathbf{C}^n$.) Hence B is an isometry from $L^2(\mathbf{R}^n)$ into the **Fock space**

$$\mathcal{F}_n = \left\{ F : F \text{ is entire on } \mathbf{C}^n \text{ and } \|F\|_{\mathcal{F}}^2 = \int |F(z)|^2 e^{-\pi|z|^2} dz < \infty \right\}.$$

We shall show below that B maps $L^2(\mathbf{R}^n)$ onto $\mathcal{F}_n$, and also explain the connection between $\mathcal{F}_n$ and the physicists' Fock space. First, we investigate the properties of $\mathcal{F}_n$. We denote the scalar product on $\mathcal{F}_n$ by $\langle\ ,\ \rangle_{\mathcal{F}}$.

(1.63) Theorem. *Let*

$$\zeta_\alpha(z) = \sqrt{\frac{\pi^{|\alpha|}}{\alpha!}} z^\alpha.$$

Then $\{\zeta_\alpha : |\alpha| \geq 0\}$ is an orthonormal basis for $\mathcal{F}_n$.

Proof: Orthonormality is easily proved by integrating in polar coordinates:

$$\begin{aligned}\sqrt{\frac{\alpha!\beta!}{\pi^{|\alpha|+|\beta|}}} \langle \zeta_\alpha, \zeta_\beta \rangle_{\mathcal{F}} &= \prod_1^n \int_{\mathbf{C}} z_j^{\alpha_j} \bar{z}_j^{\beta_j} e^{-\pi|z_j|^2} dz_j \\ &= \prod_1^n \int_0^\infty \int_0^{2\pi} e^{i\theta(\alpha_j - \beta_j)} r^{\alpha_j + \beta_j + 1} e^{-\pi r^2} d\theta \, dr.\end{aligned}$$

The θ-integral is zero unless $\beta = \alpha$, in which case we get

$$\frac{\alpha!}{\pi^{|\alpha|}} \|\zeta_\alpha\|_{\mathcal{F}}^2 = (2\pi)^n \prod_1^n \int_0^\infty r^{2\alpha_j + 1} e^{-\pi r^2} dr = \prod_1^n \int_0^\infty \left(\frac{s}{\pi}\right)^{\alpha_j} e^{-s} ds = \frac{\alpha!}{\pi^{|\alpha|}}.$$

This calculation shows more generally that if

$$B_R = \{z \in \mathbf{C}^n : |z| \leq R\} \qquad \text{and} \qquad c_{R,\alpha} = \pi^{-|\alpha|} \prod_1^n \int_0^R s^{\alpha_j} e^{-s} ds$$

then $\{c_{R,\alpha}^{-1/2} z^\alpha\}$ is an orthonormal set in $L^2(B_R,\, e^{-\pi|z|^2}dz)$. To prove completeness, then, suppose $F \in \mathcal{F}_n$, and let $\sum a_\alpha z^\alpha$ be the Taylor series of F about 0. For all $R > 0$ this series converges to F uniformly on B_R, hence in $L^2(B_R,\, e^{-\pi|z|^2}dz)$. From the preceding remark it follows that $c_{R,\alpha}^{-1/2} a_\alpha$ is the αth Fourier coefficient of F with respect to the set $\{c_{R,\alpha}^{-1/2} z^\alpha\}$,

$$a_\alpha = c_{R,\alpha}^{-1} \int_{B_R} F(z)\bar{z}^\alpha e^{-\pi|z|^2}\, dz,$$

and the Parseval formula holds:

$$\int_{B_R} |F(z)|^2 e^{-\pi|z|^2}\, dz = \sum c_{R,\alpha}|A_\alpha|^2.$$

Let $R \to \infty$: then $c_{R,\alpha} \to \pi^{-|\alpha|}\alpha!$, so these equations become

$$\|F\|_{\mathcal{F}}^2 = \sum |\langle f, \zeta_\alpha\rangle_{\mathcal{F}}|^2.$$

Therefore $\{\zeta_\alpha\}$ is a basis. ∎

(1.64) Corollary. *If $F \in \mathcal{F}_n$ then the Taylor series of F converges to F in the topology of $\mathcal{F}_n$.*

(1.65) Corollary. *If $F \in \mathcal{F}_n$ then $|F(z)| \le e^{(\pi/2)|z|^2}\|F\|_{\mathcal{F}}$ for all $z \in \mathbf{C}^n$.*

Proof: The preceding argument shows that the Fourier series of F with respect to the basis $\{\zeta_\alpha\}$ is the Taylor series of F. Thus, if $F = \sum a_\alpha\zeta_\alpha$, the Schwarz inequality yields

$$\begin{aligned} |F(z)| &= \left|\sum a_\alpha(\pi^{|\alpha|}/\alpha!)^{1/2} z^\alpha\right| \\ &\le \left(\sum |a_\alpha|^2\right)^{1/2} \left(\sum (\pi^{|\alpha|}/\alpha!)|z|^{2\alpha}\right)^{1/2} = \|F\|_{\mathcal{F}} e^{(\pi/2)|z|^2}. \qquad ∎ \end{aligned}$$

By Corollary (1.65), for each z the map $F \to F(z)$ is a bounded linear functional on $\mathcal{F}_n$, so there exists $E_z \in \mathcal{F}_n$ such that

$$F(z) = \langle F, E_z\rangle_{\mathcal{F}}.$$

It is easy to identify E_z; we have

$$\begin{aligned} E_z(w) &= \sum \langle E_z, \zeta_\alpha\rangle_{\mathcal{F}} \zeta_\alpha(w) = \sum \overline{\zeta_\alpha(z)}\zeta_\alpha(w) = \sum (\pi^{|\alpha|}\bar{z}^\alpha w^\alpha/\alpha!) \\ &= e^{\pi w\bar{z}}. \end{aligned} \tag{1.66}$$

Put in other terms, the function $K(z,\overline{w}) = e^{\pi z\overline{w}}$ is the reproducing kernel for the space $\mathcal{F}_n$:

$$F(z) = \int e^{\pi z\overline{w}} F(w) e^{-\pi|w|^2}\,dw, \qquad \text{for} \quad F \in \mathcal{F}_n, \quad z \in \mathbf{C}^n.$$

We observe also that

$$\|E_z\|_{\mathcal{F}}^2 = \sum \frac{\pi^{|\alpha|}}{\alpha!}|z^\alpha|^2 = e^{\pi|z|^2}. \tag{1.67}$$

An important consequence of the existence of a reproducing kernel is that every bounded operator on $\mathcal{F}_n$ can be written as an integral operator. More precisely, we have:

(1.68) Proposition. *If T is a bounded operator on $\mathcal{F}_n$, let $K_T(z,\overline{w}) = TE_w(z)$. Then K_T is an entire function on $\mathbf{C}^{2n}$ that satisfies*

(a) $K_T(\cdot, w) \in \mathcal{F}_n$ *for all* w *and* $K_T(z,\cdot) \in \mathcal{F}_n$ *for all* z,

(b) $|K_T(z,\overline{w})| \le e^{(\pi/2)(|z|^2+|w|^2)}\|T\|$,

(c) $TF(z) = \displaystyle\int K_T(z,\overline{w})F(w)e^{-\pi|w|^2}\,dw$ *for all* $F \in \mathcal{F}_n$ *and* $z \in \mathbf{C}^n$.

Proof: We have

$$TF(z) = \langle TF,\, E_z\rangle_{\mathcal{F}} = \langle F,\, T^*E_z\rangle_{\mathcal{F}} = \int \overline{T^*(E_z)(w)}F(w)e^{-\pi|w|^2}\,dw,$$

and

$$\overline{T^*E_z(w)} = \overline{\langle T^*E_z,\, E_w\rangle} = \langle TE_w,\, E_z\rangle = TE_w(z).$$

These formulas show that K_T is entire (since E_z depends antiholomorphically on z) and satisfies (a) and (c). As for (b), by (1.67),

$$|K_T(z,\overline{w})| \le \|TE_w\|_{\mathcal{F}}\|E_z\|_{\mathcal{F}} \le \|T\|\,\|E_w\|_{\mathcal{F}}\|E_z\|_{\mathcal{F}} = e^{(\pi/2)(|w|^2+|z|^2)}\|T\|. \qquad \blacksquare$$

In this connection the following observation is sometimes useful:

(1.69) Proposition. *An entire function $K(z,\overline{w})$ of z and $\overline{w}$ is uniquely determined by its restriction to the diagonal $z = w$.*

Proof: Let $u = \frac{1}{2}(z+\overline{w})$ and $v = -\frac{1}{2}i(z-\overline{w})$, so that $z = u+iv$ and $\overline{w} = u - iv$. Then $K(z,\overline{w}) = G(u,v)$ where G is entire. But G is determined (by Taylor's formula, say) by its values for u and v real, and u and v are real precisely when $z = w$. $\blacksquare$

(1.70) Corollary. *A bounded operator T on $\mathcal{F}_n$ is uniquely determined by the function $K_T(z, \bar{z}) = \langle TE_z, E_z \rangle$.*

We now return to consideration of the Heisenberg group. The representation ρ can be transfered via the Bargmann transform to a representation β of $\mathbf{H}_n$ on $B(L^2(\mathbf{R}^n))$ (which, as we shall shortly see, coincides with $\mathcal{F}_n$). To describe this representation, it will be convenient to identify the underlying manifold of $\mathbf{H}_n$ with $\mathbf{C}^n \times \mathbf{R}$:

$$(p, q, t) \longleftrightarrow (p + iq,\, t).$$

In this parametrization of $\mathbf{H}_n$ the group law is given by

$$(z, t)(z', t') = \big(z + z',\, t + t' + \tfrac{1}{2}\mathrm{Im}\,\bar{z}z'\big).$$

The transferred representation β is then defined by

$$\beta(p + iq,\, t)B = B\rho(p, q, t).$$

As with ρ, we set

$$\beta(w) = \beta(w, 0), \quad \text{i.e.,} \quad \beta(w, t) = e^{2\pi it}\beta(w).$$

We proceed to calculate β. Let $z = p + iq$, $w = r + is$. Then for $f \in L^2(\mathbf{R}^n)$,

$$\begin{aligned}
\big[\beta(w)Bf\big](z) &= \big[B\rho(r, s)f\big](z) \\
&= e^{(\pi/2)|z|^2} V(\rho(r, s)f,\, \phi_0)(p, q) \\
&= e^{(\pi/2)|z|^2} e^{\pi i(ps - qr)} V(f, \phi_0)(p + r,\, q + s) \qquad \text{[by Prop. (1.44b)]} \\
&= e^{(\pi/2)|z|^2} e^{-\pi i \mathrm{Im}\, z\bar{w}} e^{-(\pi/2)|z+w|^2} Bf(z + w) \\
&= e^{-(\pi/2)|w|^2 - \pi z\bar{w}} Bf(z + w).
\end{aligned}$$

In other words,

$$\beta(w, t)F(z) = e^{-(\pi/2)|w|^2 - \pi z\bar{w} + 2\pi it} F(z + w). \tag{1.71}$$

At this point we observe that

$$B\phi_0(z) = 2^{n/2} e^{-(\pi/2)|z|^2} \int e^{-2\pi zx - 2\pi x^2}\, dx = 1 = E_0(z),$$

and hence, if $w = r + is$,

$$B\big(\rho(r, s)\phi_0\big)(z) = \beta(w)(1)(z) = e^{-(\pi/2)|w|^2 - \pi z\bar{w}} = e^{-(\pi/2)|w|^2} E_{-w}(z). \tag{1.72}$$

Thus all the E_w's are in the range of B, and since $\langle F, E_w\rangle_{\mathcal{F}} = 0$ only when $F = 0$, it follows that $B(L^2(\mathbf{R}^n)) = \mathcal{F}_n$ as claimed.

Incidentally, in physicists' terminology, (1.72) says that the coherent states in the Fock model are just the functions E_z.

Next, we compute the infinitesimal representation of β, that is, the operators corresponding to X_j and D_j under B. Here again it will be more suitable to consider the complex linear combinations

$$(1.73)\qquad A_j = \sqrt{\pi}B(X_j + iD_j)B^{-1}, \qquad A_j^* = \sqrt{\pi}B(X_j - iD_j)B^{-1}.$$

We set

$$w = r + is, \qquad \frac{\partial}{\partial w_j} = \frac{1}{2}\left(\frac{\partial}{\partial r_j} - i\frac{\partial}{\partial s_j}\right), \qquad \frac{\partial}{\partial \overline{w}_j} = \frac{1}{2}\left(\frac{\partial}{\partial r_j} + i\frac{\partial}{\partial s_j}\right).$$

Since

$$X_j f = \frac{1}{2\pi i}\frac{\partial}{\partial s_j}\rho(r,s)f|_{r=s=0} \qquad D_j f = \frac{1}{2\pi i}\frac{\partial}{\partial r_j}\rho(r,s)f|_{r=s=0},$$

by (1.71) we have

$$(1.74)\qquad \begin{aligned} A_jF &= \frac{\sqrt{\pi}}{2\pi i}B\left(\frac{\partial}{\partial s_j} + i\frac{\partial}{\partial r_j}\right)\rho(r,s)B^{-1}F|_{r=s=0} \\ &= \frac{1}{\sqrt{\pi}}\frac{\partial}{\partial w_j}\beta(w)F|_{w=0} = \frac{1}{\sqrt{\pi}}\frac{\partial F}{\partial z_j}, \end{aligned}$$

$$(1.75)\qquad \begin{aligned} A_j^*F &= \frac{\sqrt{\pi}}{2\pi i}B\left(\frac{\partial}{\partial s_j} - i\frac{\partial}{\partial r_j}\right)\rho(r,s)B^{-1}F|_{r=s=0} \\ &= \frac{-1}{\sqrt{\pi}}\frac{\partial}{\partial \overline{w}_j}\beta(w)F|_{w=0} = \sqrt{\pi}z_jF. \end{aligned}$$

We leave it as an easy exercise for the reader to verify that the operators A_j and A_j^*, defined on the obvious domains

$$\mathrm{Dom}(A_j) = \{F \in \mathcal{F}_n : \partial F/\partial z_j \in \mathcal{F}_n\}, \qquad \mathrm{Dom}(A_j^*) = \{F \in \mathcal{F}_n : z_jF \in \mathcal{F}_n\},$$

are adjoints of each other. Moreover, they satisfy the commutation relations

$$(1.76)\qquad [A_j, A_k] = [A_j^*, A_k^*] = 0, \qquad [A_j, A_k^*] = \delta_{jk}I.$$

The operators A_j and A_j^* act very simply on the basis $\{\zeta_\alpha\}$. If we define 1_j to be the multi-index whose jth entry is 1 and whose other entries are 0, we clearly have

$$\frac{\partial z^\alpha}{\partial z_j} = \alpha_j z^{\alpha - 1_j}, \qquad z_j z^\alpha = z^{\alpha + 1_j},$$

and hence

$$(1.77) \qquad A_j\zeta_\alpha = \sqrt{\alpha_j}\,\zeta_{\alpha-1_j}, \qquad A_j^*\zeta_\alpha = \sqrt{\alpha_j+1}\,\zeta_{\alpha+1_j},$$

where $\zeta_{\alpha-1_j} = 0$ if $\alpha_j = 0$. In particular,

$$(1.78) \qquad \zeta_\alpha = \frac{1}{\sqrt{\alpha!}}(A_1^*)^{\alpha_1}\cdots(A_n^*)^{\alpha_n}\zeta_0.$$

Incidentally, by expanding $F \in \mathcal{F}_n$ in terms of the basis $\{\zeta_\alpha\}$ and using the Parseval equation, one can easily derive the formula

$$\|z_jF\|_{\mathcal{F}}^2 = \|F\|_{\mathcal{F}}^2 + \|\partial F/\partial z_j\|_{\mathcal{F}}^2,$$

from which it follows that

$$\mathrm{Dom}(A_j) = \mathrm{Dom}(A_j^*).$$

Let us now compute the inverse Bargmann transform. Since B is unitary, for $F \in \mathcal{F}_n$ and $g \in L^2(\mathbf{R}^n)$, we have

$$\langle B^{-1}F, g\rangle = \langle F, Bg\rangle_{\mathcal{F}} = 2^{n/4}\iint F(z)\overline{g(x)}e^{2\pi x\bar{z}-\pi x^2-(\pi/2)\bar{z}^2-\pi|z|^2}\,dx\,dz,$$

and hence

$$B^{-1}F(x) = 2^{n/4}\int F(z)e^{2\pi x\bar{z}-\pi x^2-(\pi/2)\bar{z}^2-\pi|z|^2}\,dz,$$

provided that the integrals are absolutely convergent. This will be the case if $|F(z)| \le Ce^{\delta|z|^2}$ for some $\delta < \pi/2$, and in particular if F is a polynomial. For a general $F \in \mathcal{F}_n$ the integral giving $B^{-1}F(x)$ may not converge, but we can compute $B^{-1}F$ by applying it to the partial sums of the Taylor series of F and taking the limit of the resulting functions in the L^2 norm. With this understanding, we can reformulate the Bargmann transform and its inverse as follows. We define the **Bargmann kernel** $B(z,x)$ by

$$(1.79) \qquad B(z,x) = 2^{n/4}e^{2\pi xz-\pi x^2-(\pi/2)z^2} \qquad (z \in \mathbf{C}^n, \quad x \in \mathbf{R}^n),$$

and we then have

$$(1.80) \quad Bf(z) = \int B(z,x)f(x)\,dx, \qquad B^{-1}F(x) = \int B(\bar{z},x)F(z)e^{-\pi|z|^2}\,dz.$$

For future reference, we exhibit the relationship between Hilbert-Schmidt operators on $L^2(\mathbf{R}^n)$ and on $\mathcal{F}_n$.

(1.81) Proposition. *Suppose $k \in L^2(\mathbf{R}^{2n})$, and let*

$$Tf(x) = \int k(x,y)f(y)\,dy, \qquad f \in L^2(\mathbf{R}^n).$$

Then

$$BTB^{-1}F(z) = \int K(z,\overline{w})F(w)e^{-\pi|w|^2}\,dw, \qquad F \in \mathcal{F}_n,$$

where K is the ($2n$-dimensional) Bargmann transform of k.

Proof: We first observe that the Bargmann kernels B_{2n} and B_n in dimensions $2n$ and n are related by

$$B_{2n}\big((z,w),\,(x,y)\big) = B_n(z,x)B_n(w,y),$$

so that, with $B = B_n$,

$$K(z,w) = \iint B(z,x)B(w,y)k(x,y)\,dx\,dy.$$

If F is a polynomial, we can apply formula (1.80) to write

$$BTB^{-1}F(z) = \iiint B(z,x)k(x,y)B(\overline{w},y)F(w)e^{-\pi|w|^2}\,dw\,dy\,dx.$$

One easily computes that $\|B(z,\cdot)\|_2 = e^{\pi|z|^2/2}$, so by the Schwarz inequality, the above integral is majorized by

$$\|B(z,\cdot)\|_2\|k\|_2 \int \|B(\overline{w},\cdot)\|_2|F(w)|e^{-\pi|w|^2}\,dw$$
$$= e^{\pi|z|^2/2}\|k\|_2 \int |F(w)|e^{-\pi|w|^2/2}dw,$$

which is finite since F is a polynomial. Thus we can integrate in x and y first to obtain

$$BTB^{-1}F(z) = \int K(z,\overline{w})F(w)e^{-\pi|w|^2}\,dw,$$

and this formula remains valid for arbitrary $F \in \mathcal{F}_n$ by continuity. ∎

Finally, we indicate how to modify the construction of the Fock–Bargmann representation for values of Planck's constant other than 1. If $h > 0$, we define the Fock space to be

$$\mathcal{F}_n^h = \{\, F : F \text{ is entire on } \mathbf{C}^n \text{ and } h^n\int |F(z)|^2 e^{-\pi h|z|^2}\,dz < \infty \,\}$$

and the Bargmann transform $B_h : L^2(\mathbf{R}^n) \to \mathcal{F}_n^h$ to be

$$B_h f(z) = e^{(\pi h/2)|z|^2} \langle \rho_h(p,q)f,\ \phi_h \rangle$$

where

$$z = p + iq \qquad \text{and} \qquad \phi_h(x) = \left(\frac{2}{h}\right)^{n/4} e^{-(\pi/h)x^2},$$

in other words,

$$B_h f(z) = \left(\frac{2}{h}\right)^{n/4} \int f(x) e^{2\pi xz - (\pi/h)x^2 - (\pi h/2)z^2}\, dx.$$

Then the representation

$$\beta_h(w) = B_h \rho_h(r,s) B_h^{-1} \qquad (w = r + is)$$

is given by

$$\beta_h(w)F(z) = e^{-(\pi h/2)|w|^2 - \pi h z\overline{w}} F(z + w).$$

On the other hand, if $h < 0$, the Fock space $\mathcal{F}_n^h$ consists of *antiholomorphic* functions:

$$\mathcal{F}_n^h = \{F \circ c : F \in \mathcal{F}_n^{|h|}\}, \qquad \text{where} \qquad c(z) = \overline{z}.$$

The Bargmann transform is

$$B_h f(\overline{z}) = e^{(\pi|h|/2)|z|^2} \langle \rho_h(p,q)f, \phi_{|h|} \rangle$$

where $\overline{z} = p - iq$ and $\phi_{|h|}$ is as above, in other words,

$$B_h f(\overline{z}) = \left(\frac{2}{|h|}\right)^{n/4} \int f(x) e^{-2\pi x\overline{z} + (\pi/h)x^2 + (\pi h/2)\overline{z}^2}\, dx.$$

The representation

$$\beta_h(w) = B_h \rho(r,s) B_h^{-1} \qquad (w = r + is)$$

is then given by

$$\beta_h(w)F(\overline{z}) = e^{(\pi h/2)|w|^2 + \pi h w\overline{z}} F(\overline{z} + \overline{w}).$$

Some Motivation and History. We begin by explicating the relationship between our space $\mathcal{F}_n$ and the Fock space of quantum mechanics. If $\mathcal{H}$ is any separable Hilbert space, the **Fock space over $\mathcal{H}$** is the complete tensor algebra over $\mathcal{H}$:

$$\mathcal{F}(\mathcal{H}) = \bigoplus_0^\infty (\otimes^k \mathcal{H}),$$

where $\otimes^k \mathcal{H}$ is the kth tensor power of $\mathcal{H}$ for $k \geq 1$ and $\otimes^0 \mathcal{H} = \mathbf{C}$. The Hilbert space structure on $\mathcal{F}(\mathcal{H})$ may be described as follows: if $\{e_j\}$ is an orthonormal basis for $\mathcal{H}$, then $\{e_{j_1} \otimes \cdots \otimes e_{j_k}\}$ is an orthonormal basis for $\otimes^k \mathcal{H}$, and the union of all these, together with the basis $\{1\}$ for $\otimes^0 \mathcal{H}$, is an orthonormal basis for $\mathcal{F}(\mathcal{H})$.

If $\mathcal{H}$ represents the state space for a quantum particle, $\otimes^k(\mathcal{H})$ can be considered as the state space for a system of k particles of the same type, and $\mathcal{F}(\mathcal{H})$ the state space for a system in which any number of particles can occur. In practice, however, particles are either bosons or fermions, which means that the k-particle states must be either symmetric or antisymmetric under interchancge of two particles. (The antisymmetry in the case of fermions is precisely the Pauli exclusion principle.) In these two cases, $\mathcal{F}(\mathcal{H})$ should be replaced by the **boson Fock space** $\mathcal{F}_s(\mathcal{H})$ consisting of all symmetric tensors or the **fermion Fock space** $\mathcal{F}_a(\mathcal{H})$ consisting of all antisymmetric tensors.

Our concern here is with the boson Fock space. The symmetrizer S which projects $\mathcal{F}(\mathcal{H})$ onto $\mathcal{F}_s(\mathcal{H})$ is given on $\otimes^k \mathcal{H}$ by

$$S(u_1 \otimes \cdots \otimes u_k) = \frac{1}{k!} \sum_\sigma u_{\sigma(1)} \otimes \cdots \otimes u_{\sigma(k)}$$

where σ ranges over the group of permutations of k letters. It is easy to verify that if $\{e_j\}$ is an orthonormal basis for $\mathcal{H}$, then

$$\left\{ E_\alpha = \sqrt{\frac{k!}{\alpha!}}\, S(e_1^{\alpha_1} \otimes e_2^{\alpha_2} \otimes \cdots) : \quad \sum \alpha_j = k, \quad k = 0, 1, 2, \ldots \right\}$$

is an orthonormal basis for $\mathcal{F}_s(\mathcal{H})$, where the superscripts α_j denote tensor powers.

Now let $\mathcal{H} = (\mathbf{C}^n)^*$, and let $\{e_j\}_1^n$ be the standard coordinate functions on $\mathbf{C}^n$. The elements of $\otimes^k \mathcal{H}$ are then k-linear functionals on $\mathbf{C}^n$, which are in one-to-one correspondence with homogeneous polynomials of degree k on $\mathbf{C}^n$. In order to make the normalizations come out right, we introduce a factor of $\sqrt{\pi^k/k!}$ into this correspondence:

$$S(e_1^{\alpha_1} \otimes \cdots \otimes e_n^{\alpha_n}) \longleftrightarrow \sqrt{\frac{\pi^k}{k!}}\, z_1^{\alpha_1} \cdots z_n^{\alpha_n} = \sqrt{\frac{\pi^k}{k!}}\, z^\alpha, \quad |\alpha| = k,$$

or in other words,

$$E_\alpha \longleftrightarrow \zeta_\alpha.$$

This correspondence clearly defines a unitary map from $\mathcal{F}_s((\mathbf{C}^n)^*)$ onto our space $\mathcal{F}_n$.

There are analogues of our operators A_j and A_j^* on an arbitrary $\mathcal{F}_s(\mathcal{H})$. Namely, given bases $\{e_j\}$ and $\{E_\alpha\}$ for $\mathcal{H}$ and $\mathcal{F}_s(\mathcal{H})$ as above, one defines A_j and A_j^* on the basis $\{E_\alpha\}$ by

$$A_j E_\alpha = \sqrt{\alpha_j}\, E_{\alpha-1_j}, \qquad A_j^* E_\alpha = \sqrt{\alpha_j+1}\, E_{\alpha+1_j}.$$

It is not hard to verify that

$$A_j^*\left(\sum_0^\infty v^{(k)}\right) = \sum_0^\infty \sqrt{k+1}\, S(e_j \otimes v^{(k)}), \qquad v^{(k)} \in \otimes^k \mathcal{H}.$$

More generally, for any $u \in \mathcal{H}$ one can define A_u^* by replacing e_j by u in this formula, and then define A_u to be the adjoint of A_u^*. A_u^* and A_u map the k-particle states into $(k+1)$-particle states and $(k-1)$-particle states, respectively; they are called **creation** and **annihilation** operators in quantum field theory.

The Fock spaces $\mathcal{F}(\mathcal{H})$, $\mathcal{F}_s(\mathcal{H})$, and $\mathcal{F}_a(\mathcal{H})$ were introduced by Fock [48]. It was also Fock [47] who first described (on the level of formal calculation) the use of $A = \partial/\partial z$, $A' = z$ to solve the commutator equation $[A, A'] = I$. The rigorous development of the representation of $\mathbf{H}_n$ on $\mathcal{F}_n$ and the intertwining operator B is due to Bargmann [11]; the same ideas also appear in work of Segal [126], [127], done independently at about the same time.

We pulled the Fock space and the Bargmann transform out of a hat by using the Fourier-Wigner transform. It is perhaps more enlightening to see the heuristic method by which Bargmann [11] derived them. To begin with, we observe that if P_j and Q_j are self-adjoint operators satisfying the canonical commutation relations (1.9) with $h = 1$, then the operators $A_j = \pi^{1/2}(Q_j + iP_j)$ and $A_j^* = \pi^{1/2}(Q_j - iP_j)$ satisfy the commutation relations (1.76), and conversely. The latter relations are also satisfied by the differential operators $\pi^{-1/2}\partial/\partial z_j$ and $\pi^{1/2} z_j$, so we are led to look for a Hilbert space of holomorphic functions on which these operators are adjoints of each other. As a candidate for such a space, we try the space $\mathcal{H}$ of entire functions in $L^2(\mathbf{C}^n, \omega(z,\overline{z})dz)$ where ω is a suitable positive weight function, and the condition we require is

$$\pi^{1/2}\int z_j F\overline{G}\omega\, dz = \pi^{-1/2}\int F\,\overline{\frac{\partial G}{\partial z_j}}\,\omega\, dz \qquad \text{for} \qquad F, G \in \mathcal{H}.$$

But if we integrate by parts, assuming that F, G, and ω are such that the boundary term vanishes, since F is holomorphic we obtain

$$\int F\,\overline{\frac{\partial G}{\partial z_j}}\,\omega\, dz = -\int \overline{G}\,\frac{\partial(F\omega)}{\partial \overline{z}_j}\, dz = -\int F\overline{G}\,\frac{\partial \omega}{\partial \overline{z}_j}\, dz,$$

so we must have $\partial\omega/\partial\bar{z}_j = -\pi z_j\omega$ for all j, or, with $z = x + iy$,

$$\frac{\partial\omega}{\partial x_j} + i\frac{\partial\omega}{\partial y_j} = -2\pi(x_j + iy_j)\omega.$$

Since ω is positive, this means that

$$\frac{\partial\omega}{\partial x_j} = -2\pi x_j\omega \quad\text{and}\quad \frac{\partial\omega}{\partial y_j} = -2\pi y_j\omega, \qquad\text{or}\quad \nabla_{x,y}(\log\omega) = (-2\pi x,\ -2\pi y),$$

so that $\log\omega = -\pi(x^2 + y^2) + C = -\pi|z|^2 + C$. We may choose $C = 0$; then $\omega = e^{-\pi|z|^2}$, so we obtain the space $\mathcal{F}_n$.

Next, we look for an operator

$$Bf(z) = \int f(x)B(z,x)\,dx$$

that maps $L^2(\mathbf{R}^n)$ onto $\mathcal{F}_n$ and intertwines the operators

$$\pi^{1/2}(X_j + iD_j) = \pi^{1/2}x_j + \frac{1}{2\pi^{1/2}}\frac{\partial}{\partial x_j} \text{ and } \pi^{1/2}(X_j - iD_j) = \pi^{1/2}x_j - \frac{1}{2\pi^{1/2}}\frac{\partial}{\partial x_j}$$

with $\pi^{1/2}z_j$ and $\pi^{-1/2}\partial/\partial z_j$. Again, a formal integration by parts yields

$$\int \big[(X_j \pm iD_j)f(x)\big]B(z,x)\,dx = \int f(x)\big[(X_j \mp iD_j)B(z,x)\big]\,dx,$$

so the intertwining conditions will be satisfied if

$$\left(\pi x_j + \frac{1}{2}\frac{\partial}{\partial x_j}\right)B(z,x) = \pi z_j B(z,x), \qquad \left(\pi x_j - \frac{1}{2}\frac{\partial}{\partial x_j}\right)B(z,x) = \frac{\partial B}{\partial z_j}(z,x),$$

or

$$\frac{\partial B}{\partial x_j} = 2\pi(z_j - x_j)B, \qquad \frac{\partial B}{\partial z_j} = \pi x_j B - \frac{1}{2}\frac{\partial B}{\partial x_j} = \pi(2x_j - z_j)B.$$

Hence

$$\frac{\partial(\log B)}{\partial x_j} = 2\pi(z_j - x_j), \qquad \frac{\partial(\log B)}{\partial z_j} = \pi(2x_j - z_j),$$

and integrating these equations gives

$$\log B = 2\pi xz - \pi x^2 - \tfrac{1}{2}\pi z^2 + C.$$

The constant C is chosen to be $\log 2^{n/4}$ to make the transform unitary, and the derivation is thereby complete.

7. Hermite Functions

The monomials $\zeta_\alpha(z) = \sqrt{\pi^{|\alpha|}/\alpha!}\,z^\alpha$ obviously play a distinguished role among all orthonormal bases for the Fock space $\mathcal{F}_n$, and one might therefore suspect that the corresponding functions $B^{-1}\zeta_\alpha$ should also be important in $L^2(\mathbf{R}^n)$. This is indeed the case. We call $B^{-1}\zeta_\alpha$ the αth (normalized, n-dimensional) **Hermite function** and denote it by h_α.

To compute h_α we utilize the operators

$$Z_j = (X_j + iD_j) = \pi^{-1/2}B^{-1}A_jB, \qquad Z_j^* = (X_j - iD_j) = \pi^{-1/2}B^{-1}A_j^*B$$

and their products

$$Z^\alpha = Z_1^{\alpha_1}\cdots Z_n^{\alpha_n}, \qquad Z^{*\alpha} = Z_1^{*\alpha_1}\cdots Z_n^{*\alpha_n}.$$

We observe that

$$Z_j^* f(x) = x_j f(x) - \frac{1}{2\pi}\frac{\partial f}{\partial x_j} = \frac{-1}{2\pi}e^{\pi x^2}\frac{\partial}{\partial x_j}(e^{-\pi x^2}f(x)),$$

so that

$$Z^{*\alpha}f(x) = \left(\frac{-1}{2\pi}\right)^{|\alpha|} e^{\pi x^2}\left(\frac{\partial}{\partial x}\right)^\alpha (e^{-\pi x^2}f(x)).$$

We have already noted in (1.72) that

$$h_0(x) = (B^{-1}\zeta_0)(x) = (B^{-1}E_0)(x) = 2^{n/4}e^{-\pi x^2}.$$

Therefore, by (1.78),

$$\begin{aligned} h_\alpha(x) &= \sqrt{\frac{1}{\alpha!}}\,(B^{-1}A^{*\alpha}\zeta_0)(x) = \sqrt{\frac{\pi^{|\alpha|}}{\alpha!}}\,(Z^{*\alpha}h_0)(x) \\ &= \frac{2^{n/4}}{\sqrt{\alpha!}}\left(\frac{-1}{2\sqrt{\pi}}\right)^{|\alpha|} e^{\pi x^2}\left(\frac{\partial}{\partial x}\right)^\alpha (e^{-2\pi x^2}). \end{aligned} \tag{1.81}$$

In particular, taking $n = 1$, we obtain the one-dimensional Hermite functions

$$h_j(x) = \frac{2^{1/4}}{\sqrt{j!}}\left(\frac{-1}{2\sqrt{\pi}}\right)^j e^{\pi x^2}\frac{d^j}{dx^j}(e^{-2\pi x^2}).$$

These are not quite the same as the Hermite functions usually found in the literature, because they are built from $e^{-\pi x^2}$ rather than $e^{-x^2/2}$ and are normalized differently. The **classical Hermite functions** on $\mathbf{R}$ are defined by

$$\widetilde{h}_j(x) = (-1)^j e^{x^2/2}\frac{d^j}{dx^j}e^{-x^2}.$$

It follows easily that

$$h_j(x) = \frac{2^{1/4}}{\sqrt{2^j j!}} \widetilde{h}_j(\sqrt{2\pi}\, x).$$

Returning now to the n-dimensional case, it is easy to read off from the above calculations a number of basic properties of the Hermite functions:

(i) The n-dimensional Hermite functions are products of one-dimensional Hermite functions; namely,

$$h_\alpha(x) = h_{\alpha_1}(x_1) \cdots h_{\alpha_n}(x_n).$$

(ii) The function $H_\alpha(x) = e^{\pi x^2} h_\alpha(x)$ is a polynomial of degree $|\alpha|$, called the αth **Hermite polynomial.** We have

$$H_\alpha(x) = 2^{(n/4)+|\alpha|} \sqrt{\frac{\pi^{|\alpha|}}{\alpha!}}\, x^\alpha + (\text{terms of degree } < |\alpha|).$$

(iii) Every polynomial of degree $\leq k$ on $\mathbf{R}^n$ is a linear combination of Hermite polynomials of degree $\leq k$. This follows from the preceding formula by induction on k.

(iv) Since $[Z_j, Z_k^*] = \pi^{-1}\delta_{jk} I$, one finds by induction that

$$[Z_j, Z^{*\alpha}] = \pi^{-1}\alpha_j Z^{*(\alpha - 1_j)}.$$

(v) We have

$$Z_j h_\alpha = \sqrt{\frac{\alpha_j}{\pi}}\, h_{\alpha - 1_j}, \qquad Z_j^* h_\alpha = \sqrt{\frac{\alpha_j + 1}{\pi}}\, h_{\alpha + 1_j}. \tag{1.82}$$

Of course this follows immediately from the corresponding property (1.78) of the ζ_α's, but it can also be verified directly by using (iv) and the formula $h_\alpha = \sqrt{\pi^{|\alpha|}/\alpha!}\, Z^{*\alpha} h_0$.

(vi) In dimension one we have

$$ZZ^* = (X + iD)(X - iD) = X^2 + D^2 + i[D, X] = D^2 + X^2 + (2\pi)^{-1} I.$$

The operator

$$2\pi(D^2 + X^2) = 2\pi x^2 - \frac{1}{2\pi}\frac{d^2}{dx^2}$$

is called the **Hermite operator** (adapted to our use of $e^{-\pi x^2}$ rather than $e^{-x^2/2}$; the usual Hermite operator is $x^2 - (d/dx)^2$). The one-dimensional Hermite functions are the eigenfunctions of this operator: by (1.82) we have

$$2\pi(D^2 + X^2) h_k = (2\pi ZZ^* - 1) h_k = 2\sqrt{\pi(k+1)}\, Z h_{k+1} - h_k = (2k+1) h_k.$$

In n dimensions, this equation together with (i) shows that h_α is an eigenfunction of the Hermite operators in each variable,

$$2\pi(D_j^2 + X_j^2)h_\alpha = (2\alpha_j + 1)h_\alpha, \tag{1.83a}$$

as well as of the n-dimensional Hermite operator $2\pi(D^2 + X^2)$, that is, $2\pi\sum_1^n(D_j^2 + X_j^2)$:

$$2\pi(D^2 + X^2)h_\alpha = (2|\alpha| + n)h_\alpha. \tag{1.83b}$$

(vii) $\{h_\alpha\}$ is an orthonormal basis for $L^2(\mathbf{R}^n)$. This follows from the corresponding property of the ζ_α's, but we can prove it directly from (1.81) as follows. We have

$$\langle h_\alpha,\, h_\beta\rangle = \sqrt{\frac{\pi^{|\alpha|+|\beta|}}{\alpha!\beta!}}\langle h_0,\, Z^\alpha Z^{*\beta}h_0\rangle.$$

Since $Z_j h_0 = 0$, if $\alpha_j > \beta_j$ for any j then repeated application of (iv) shows that $\langle h_\alpha,\, h_\beta\rangle = 0$, while if $\alpha = \beta$ it shows that $\|h_\alpha\|_2^2 = \|h_0\|_2^2$, and the latter number is 1. As for completeness, if $g \in L^2$ and $\langle g, h_\alpha\rangle = 0$ for all α, then by (iii), $\langle g,\, P(x)e^{-\pi x^2}\rangle = 0$ for all polynomials P. But then

$$\int g(x)e^{-\pi x^2}e^{2\pi i x\xi}\,dx = \sum_0^\infty \int g(x)e^{-\pi x^2}\frac{(2\pi i x\xi)^j}{j!}\,dx = 0,$$

so by Fourier uniqueness, $g(x)e^{-\pi x^2} = 0$ a.e., and hence $g = 0$.

(viii) The Hermite functions are eigenfunctions of the Fourier transform $\mathcal{F}$. Indeed, since $\mathcal{F}X_j = -D_j\mathcal{F}$ and $\mathcal{F}D_j = X_j\mathcal{F}$, we have

$$\mathcal{F}Z_j^* = \mathcal{F}(X_j - iD_j) = (-D_j - iX_j)\mathcal{F} = -iZ_j^*\mathcal{F}.$$

Moreover, $\mathcal{F}h_0 = h_0$, so

$$\mathcal{F}h_\alpha = \sqrt{\frac{\pi^{|\alpha|}}{\alpha!}}\,\mathcal{F}Z^{*\alpha}h_0 = (-i)^{|\alpha|}\sqrt{\frac{\pi^{|\alpha|}}{\alpha!}}\,Z^{*\alpha}h_0 = (-i)^{|\alpha|}h_\alpha. \tag{1.84}$$

We conclude this section by deriving two classical generating function identities for Hermite functions. The first one is more or less equivalent to the fact that the Bargmann transform maps the orthonormal basis $\{h_\alpha\}$ for $L^2(\mathbf{R}^n)$ to the orthonormal basis $\{\zeta_\alpha\}$ for $\mathcal{F}^n$; the second one is somewhat deeper. In order to obtain uniformity of convergence, we shall need the following lemma.

(1.85) Lemma. *There is a constant C, depending only on the dimension n, such that*

$$\|h_\alpha\|_\infty \le C(|\alpha|+1)^{n/2}.$$

Proof: In dimension 1 we have $\|h_j\|_2 = 1$ and, by (1.82),

$$\|Dh_j\|_2^2 = \frac{1}{2}\|(Z-Z^*)h_j\|_2^2 = \frac{1}{2\sqrt{\pi}}\|\sqrt{j}\,h_{j-1} - \sqrt{j+1}\,h_{j+1}\|_2^2 = \frac{1}{2}\sqrt{\frac{2j+1}{\pi}},$$

since $h_{j-1}\perp h_{j+1}$. Moreover, the Fourier inversion formula implies that $\|h_j\|_\infty \le \|\widehat{h}_j\|_1$, so by the Schwarz inequality and the Plancherel theorem,

$$\begin{aligned}\|h_j\|_\infty \le \int |\widehat{h}_j(\xi)|\,d\xi &\le \left(\int (1+\xi^2)|\widehat{h}_j(\xi)|^2\,d\xi\right)^{1/2}\left(\int (1+\xi^2)^{-1}d\xi\right)^{1/2} \\ &= \left(\|h_j\|_2^2 + \|Dh_j\|_2^2\right)^{1/2}\pi^{1/2} \\ &= \tfrac{1}{2}\sqrt{2j+1+4\pi}.\end{aligned}$$

The n-dimensional case now follows easily in view of (i) above. ∎

(1.86) Theorem. *We have*

$$\sum_{|\alpha|\ge 0} h_\alpha(x)\zeta_\alpha(z) = 2^{n/4}e^{2\pi xz - \pi x^2 - (\pi/2)z^2} = B(z,x),$$

where $B(z,x)$ is the Bargmann kernel (1.79). The series converges uniformly on $\mathbf{R}^n \times K$ for every compact $K \subset \mathbf{C}^n$, and also in $L^2(x)$ for each z.

Proof: The assertions about convergence follow from Lemma (1.85) and the fact that $\zeta_\alpha(z) = \sqrt{\pi^{|\alpha|}/\alpha!}\,z^\alpha$ tends to zero rapidly as $|\alpha|\to\infty$ for z in any compact set. To sum the series, observe that by (1.81),

$$h_\alpha(x) = \frac{2^{n/4}}{\sqrt{\alpha!}}\left(\frac{1}{2\sqrt{\pi}}\right)^{|\alpha|} e^{\pi x^2}\left(\frac{\partial}{\partial z}\right)^\alpha e^{-2\pi(x-z)^2}\Big|_{z=0}.$$

Hence, by Taylor's theorem,

$$e^{-2\pi(x-z)^2} = \sum \frac{(2\sqrt{\pi})^{|\alpha|}\sqrt{\alpha!}}{2^{n/4}e^{\pi x^2}}h_\alpha(x)\frac{z^\alpha}{\alpha!} = 2^{-n/4}e^{-\pi x^2}\sum h_\alpha(x)\zeta_\alpha(2z),$$

and the desired result follows immediately on replacing z by $z/2$. ∎

(1.87) Mehler's Formula. *For $x, y \in \mathbf{R}^n$ and $w \in \mathbf{C}$ with $|w| < 1$ we have*

$$\sum_{|\alpha|\geq 0} w^{|\alpha|} h_\alpha(x) h_\alpha(y) = \left(\frac{2}{1-w^2}\right)^{n/2} \exp\left[\frac{-\pi(1+w^2)(x^2+y^2)+4\pi wxy}{1-w^2}\right].$$

(Here $u = 2/(1-w^2)$ lies in the right half plane, and the square root in $u^{n/2}$ is the branch that is positive for $u > 0$.) The series converges absolutely and uniformly on compact sets of $\mathbf{R}^{2n} \times \{|w| < 1\}$, and also in $L^2(y)$ for each x and w.

Proof: The assertions about convergence follow easily from Lemma (1.85) and the orthonormality of $\{h_\alpha\}$. To sum the series, we replace z by wz in Theorem (1.86):

$$\sum w^{|\alpha|} h_\alpha(x) \zeta_\alpha(z) = 2^{n/4} e^{2\pi wxz - \pi x^2 - (\pi/2)w^2 z^2}.$$

We apply the inverse Bargmann transform to both sides. On the one hand, for fixed w and x, the left side converges in the Fock space norm by Lemma (1.85), and its inverse Bargmann transform is clearly

$$B^{-1}\left(\sum w^{|\alpha|} h_\alpha(x)\zeta_\alpha\right)(y) = \sum w^{|\alpha|} h_\alpha(x) h_\alpha(y).$$

On the other hand, the right side satisfies

$$\left|2^{n/4} e^{2\pi wxz - \pi x^2 - (\pi/2)z^2}\right| \leq Ce^{(\pi/2)(|w|^2+\epsilon)|z|^2} \leq Ce^{\delta|z|^2}$$

for some $\delta < \pi/2$, so we can apply (1.80) to see that its inverse Bargmann transform is

$$\begin{aligned} & 2^{n/2} \int e^{2\pi wxz - \pi x^2 - (\pi/2)w^2z^2 + 2\pi y\bar{z} - \pi y^2 - (\pi/2)\bar{z}^2} e^{-\pi|z|^2} dz \\ =& 2^{n/2} e^{-\pi(x^2+y^2)} \int e^{(\pi/2)(-w^2z^2 - \bar{z}^2 + 4wxz + 4y\bar{z})} e^{-\pi|z|^2} dz \end{aligned}$$

By Theorem 3 of Appendix A, this last integral equals

$$(1-w^2)^{-n/2} \exp\left[(-\pi)\frac{2w^2x^2 - 4wxy + 2w^2y^2}{1-w^2}\right],$$

and the result follows immediately. ∎

Several people have obtained extensions of Mehler's formula in which exponentials of more general quadratic functions are expanded in series of Hermite functions; see Louck [97] and the references given there.

8. The Wigner Transform

The **Wigner transform** of two functions f and g is the Fourier transform of their Fourier-Wigner transform:

$$W(f,g)(\xi,x) = \iint e^{-2\pi i(\xi p + xq)} V(f,g)(p,q)\,dp\,dq.$$

$W(f,g)$ was first introduced into the literature in the case $g = f$ by Wigner [156]; the general case seems to have been first studied by Moyal [108]. Since

$$V(f,g)(p,q) = \int e^{2\pi i qy} f(y+\tfrac{1}{2}p)\overline{g(y-\tfrac{1}{2}p)}\,dy, \tag{1.88}$$

we have

$$W(f,g)(\xi,x) = \iiint e^{-2\pi i(\xi p + xq - yq)} f(y+\tfrac{1}{2}p)\overline{g(y-\tfrac{1}{2}p)}\,dy\,dp\,dq,$$

so by the Fourier inversion theorem,

$$W(f,g)(\xi,x) = \int e^{-2\pi i \xi p} f(x+\tfrac{1}{2}p)\overline{g(y-\tfrac{1}{2}p)}\,dp \tag{1.89}$$

The expressions (1.88) and (1.89) are deceptively similar; it is hard to believe that they are always Fourier transforms of one another! In fact, a simple calculation shows that

$$W(f,g)(\xi,x) = 2^n V(f,\widetilde{g})(2x,-2\xi), \quad \text{where} \quad \widetilde{g}(x) = g(-x). \tag{1.90}$$

Like V, the sesquilinear transform W can be regarded as the restriction to functions of the form $f(x)\overline{g(y)}$ of the linear transform

$$\widetilde{W}F(\xi,x) = \int e^{-2\pi i \xi p} F(x+\tfrac{1}{2}p,\ x-\tfrac{1}{2}p)\,dp, \tag{1.91}$$

defined for functions F of $2n$ variables. $\widetilde{W}$ is the composition of the measure-preserving change of variables $(x,p) \to (x+\frac{1}{2}p,\ x-\frac{1}{2}p)$ with Fourier transformation in the second variable, so it preserves the classes $\mathcal{S}(\mathbf{R}^{2n})$ and $\mathcal{S}'(\mathbf{R}^{2n})$ and is unitary on $L^2(\mathbf{R}^{2n})$. Therefore:

(1.92) Proposition. *W maps $\mathcal{S}(\mathbf{R}^n) \times \mathcal{S}(\mathbf{R}^n)$ into $\mathcal{S}(\mathbf{R}^{2n})$ and extends to a map from $\mathcal{S}'(\mathbf{R}^n) \times \mathcal{S}'(\mathbf{R}^n)$ into $\mathcal{S}'(\mathbf{R}^{2n})$. Moreover, W maps $L^2(\mathbf{R}^n) \times L^2(\mathbf{R}^n)$ into $L^2(\mathbf{R}^{2n}) \cap C_0(\mathbf{R}^{2n})$ and satisfies*

$$\langle W(f_1,g_1),\ W(f_2,g_2)\rangle = \langle f_1,f_2\rangle\overline{\langle g_1,g_2\rangle} \tag{1.93}$$

and $\|W(f,g)\|_\infty \le \|f\|_2\|g\|_2$.

Proof: The assertions about $\mathcal{S}$ and $\mathcal{S}'$ and the unitarity relation (1.93) are consequences of the properties of $\widetilde{W}$. The estimate on $\|W(f,g)\|_\infty$ comes from the Schwarz inequality, and the fact that $W(f,g) \in C_0$ for $f,g \in L^2$ then follows since $\mathcal{S}$ is dense in L^2. ∎

(1.93) is often called **Moyal's identity.**

We summarize the basic transformation properties of $W(f,g)$ in the following proposition. The verifications of these formulas are all easy exercises which we leave to the reader.

(1.94) Proposition. *For $t \in \mathbf{R}\backslash\{0\}$, let $f^t(x) = |t|^{n/2} f(tx)$. Then:*

(a) $W(f^t,\, g^t)(\xi, x) = W(f,g)(t^{-1}\xi,\, tx)$.

(b) $W\big(\rho(a,b)f,\, \rho(c,d)g\big)(\xi,x)$
$= e^{\pi i(bc-ad)+2\pi i[(a-c)\xi+(b-d)x]} W(f,g)\big(\xi - \frac{1}{2}(b+d),\, x + \frac{1}{2}(a+c)\big)$.

(c) $W(\widehat{f},\, \widehat{g})(\xi,x) = W(f,g)(x,-\xi)$.

(d) $W(g,f) = \overline{W(f,g)}$.

As a special case of Proposition (1.94b), we have

$$W(\rho(a,b)f,\, \rho(a,b)g)(\xi,x) = W(f,g)(\xi - b,\, x+a), \tag{1.95a}$$

or, what is sometimes more convenient, with $\rho'(a,b) = \rho(-b,a)$ as in (1.27),

$$W(\rho'(a,b)f,\, \rho'(a,b)g)(\xi,x) = W(f,g)(\xi - a,\, x - b). \tag{1.95b}$$

The function $W(f,g)$ is of greatest intrinsic interest in the case $g = f$. In this case we shall write

$$W(f,f) = Wf$$

and call Wf the **Wigner distribution** of f. Wf was proposed by Wigner [156] as a substitute for the nonexistent joint probability distribution of momentum and position in the quantum state f. The motivation is as follows. Since the uncertainty principle imposes a limit on the precision with which momentum and position can be determined in the state f ($f \in L^2$, $\|f\|_2 = 1$), it does not make sense to speak of a joint probability distribution for these observables. However, if such a distribution existed, with density $\sigma(\xi,x)$, then the inverse Fourier transform of σ,

$$\iint e^{2\pi i(p\xi+qx)} \sigma(\xi,x)\, d\xi\, dx,$$

would be the expected value of the function $e^{2\pi i(p\xi+qx)}$ with respect to σ. The latter quantity has a natural and consistent interpretation in quantum mechanics, namely, the expected value of $e^{2\pi i(pD+qX)}$ in the state f:

$$\langle e^{2\pi i(pD+qX)} f,\, f\rangle = \langle \rho(p,q)f,\, f\rangle = V(f,f),$$

so we are led to take $\sigma = V(f,f)\widehat{\ } = Wf$.

Wf is usually not a genuine probability density, because it may assume negative values. (Indeed, if f is odd we have $Wf(0,0) = -\|f\|_2^2$.) It is, however, always real, by Proposition (1.94d), and in some sense it tries very hard to be a joint density for momentum and position. The following results provide supporting evidence for this heuristic assertion.

In the first place, Wf has the right marginal distributions: if we integrate out either position or momentum, we get the probability distribution for the other one.

(1.96) Proposition. *We have*

$$\int Wf(\xi,x)\,dx = |\widehat{f}(\xi)|^2 \qquad \text{and} \qquad \int Wf(\xi,x)\,d\xi = |f(x)|^2.$$

Proof: Letting $u = x + \frac{1}{2}p$ and $v = x - \frac{1}{2}p$, we have

$$\begin{aligned}\int Wf(\xi,x)\,dx &= \iint e^{-2\pi i\xi p} f(x+\tfrac{1}{2}p)\overline{f(x-\tfrac{1}{2}p)}\,dp\,dx \\ &= \iint e^{-2\pi i\xi u} f(u)\overline{e^{-2\pi i\xi v} f(v)}\,du\,dv \\ &= \widehat{f}(\xi)\overline{\widehat{f}(\xi)}.\end{aligned}$$

This proves the first assertion, and the second one follows from the Fourier inversion formula:

$$\begin{aligned}\int Wf(\xi,x)\,d\xi &= \iint e^{-2\pi i\xi p} f(x+\tfrac{1}{2}p)\overline{f(x-\tfrac{1}{2}p)}\,dp\,d\xi \\ &= \int \delta(p) f(x+\tfrac{1}{2}p)\overline{f(x-\tfrac{1}{2}p)}\,dp = |f(x)|^2. \quad \blacksquare\end{aligned}$$

Remark. We have been a bit sloppy here. An examination of these calculations shows that they are rigorously correct if $f \in L^1$ (which guarantees that $Wf(\xi,\cdot) \in L^1$) and $\widehat{f} \in L^1$ (which guarantees that $Wf(\cdot,x) \in L^1$). If f is merely in L^2, the integrals $\int Wf(\xi,x)dx$ and $\int Wf(\xi,x)d\xi$ need not be absolutely convergent, but the above formulas remain valid if they are suitably interpreted, a task which we leave to the reader.

As a corollary, under suitable hypotheses to ensure convergence, we obtain

$$\iint x_j Wf(\xi,x)\,d\xi\,dx = \int x_j|f(x)|^2 dx = \langle X_j f,\, f\rangle,$$
$$\iint \xi_j Wf(\xi,x)\,d\xi\,dx = \int \xi_j|\widehat{f}(\xi)|^2 d\xi = \langle D_j f,\, f\rangle.$$

This implies that the center of mass of Wf is $(\overline{\xi}, \overline{x})$ where $\overline{\xi}$ and $\overline{x}$ are the centers of mass of $|\widehat{f}|^2$ and $|f|^2$ respectively.

The next result shows that if the position or momentum spectrum of f is limited, there is a corresponding limitation on the support of Wf. In what follows, "supp(f)" means the smallest closed set outside of which $f = 0$ a.e., and we may assume that $f = 0$ everywhere outside supp(f).

(1.97) Proposition. *Let π_1 and π_2 be the projections from $\mathbf{R}^n \times \mathbf{R}^n$ onto the first and second factors, and for $E \subset \mathbf{R}^n$ let $H(E)$ denote the closed convex hull of E. Then*

$$\pi_1\big(\mathrm{supp}(Wf)\big) \subset H\big(\mathrm{supp}(\widehat{f})\big) \qquad \text{and} \qquad \pi_2\big(\mathrm{supp}(Wf)\big) \subset H\big(\mathrm{supp}(f)\big).$$

Proof: From (1.89), $Wf(\xi, x) = 0$ unless there is some p for which $x + \frac{1}{2}p$ and $x - \frac{1}{2}p$ are in supp(f); in this case x, being halfway between these points, is in $H\big(\mathrm{supp}(f)\big)$. This proves the second assertion, and the first one follows in the same way since $Wf(\xi, x) = W\widehat{f}(-x, \xi)$ by Proposition (1.94c). ∎

In view of these results, the Wigner distribution Wf can be viewed as a sort of portrait of the quantum state f in phase space. Another fact which supports this point of view is that

$$W(e^{2\pi i(aX - bD)}f)(\xi, x) = Wf(\xi - a,\, x - b),$$

so that momentum-position translations of f (cf. the discussion following (1.27)) correspond to ordinary translations of Wf. In this connection we should observe that Wf determines f up to a phase factor:

(1.98) Proposition. *$Wf = Wg$ if and only if $f = cg$ for some $c \in \mathbf{C}$ with $|c| = 1$.*

Proof: From formula (1.89) and the Fourier inversion theorem we see that $Wf = Wg$ if and only if

$$f(x + \tfrac{1}{2}p)\overline{f(x - \tfrac{1}{2}p)} = g(x + \tfrac{1}{2}p)\overline{g(x - \tfrac{1}{2}p)} \quad \text{for almost every } x, p,$$

in other words, $f(u)\overline{f(v)} = g(u)\overline{g(v)}$ for almost every u, v. The assertion is now obvious. ∎

A more classical interpretation is also available. Take $n = 1$, and let $f(t)$ represent the amplitude of a vibration—say, a sound wave—at time t. Then the Fourier representation $f(t) = \int e^{2\pi i\omega t}\widehat{f}(\omega)d\omega$ tells how f is synthesized from waves of definite frequencies, and Wf gives a picture of f in time-frequency space. This is rather like what is done in music. f might represent a musical composition, but composers almost never try to describe either f or $\widehat{f}$ directly;

rather, they make a "time-frequency plot" of f by writing notes on musical staves. For this reason, de Bruijn [37] has dubbed Wf the "musical score" of f.

For any f in L^2, by the Schwarz inequality we have $|Wf(\xi,x)| \leq \|f\|_2^2$, and hence

$$\left|\iint_E Wf(\xi,x)\,d\xi\,dx\right| \leq \|f\|_2^2 \cdot \operatorname{meas}(E).$$

On the other hand, by Proposition (1.96),

$$\iint_{\mathbf{R}^{2n}} Wf(\xi,x)\,d\xi\,dx = \|f\|_2^2.$$

Hence the mass of Wf cannot almost all be concentrated in a set E in phase space unless $\operatorname{meas}(E) \geq 1$. This is a form of the uncertainty principle; some quantitative versions of the uncertainty principle for Wf can be found in de Bruijn [37].

Let us return to the question of the positivity of Wf. Although Wf may assume negative values, it tends to be positive "on the average": it is easy to make Wf positive by convolving it with a suitable function G.

(1.99) Proposition. *Suppose $G \in (L^1 + L^2)(\mathbf{R}^{2n})$ satisfies $\int (Wf)G \geq 0$ for all $f \in L^2(\mathbf{R}^n)$. Then $Wf * G \geq 0$ pointwise for all $f \in L^2(\mathbf{R}^n)$. In particular, this is the case if $G = Wg$ for some $g \in L^2(\mathbf{R}^n)$; in fact,*

$$Wf * Wg(\xi,x) = |V(\widetilde{f},\,g)(-x,\xi)|^2$$

where $\widetilde{f}(x) = f(-x)$.

Proof: We observe that by (1.95) and Proposition (1.94a) (with $t = -1$),

$$Wf(\xi-\eta,\,x-y) = W\widetilde{f}(\eta-\xi,\,y-x) = W\big(\rho(-x,\xi)\widetilde{f}\big)(\eta,y)$$

and hence

$$Wf * G(\xi,x) = \iint W\big(\rho(-x,\xi)\widetilde{f}\big)(\eta,y)G(\eta,y)\,d\eta\,dy.$$

This proves the first assertion. If $G = Wg$ then G is real, so by Moyal's identity (1.93) the above equation gives

$$\begin{aligned} Wf * Wg(\xi,x) &= \langle W\big(\rho(-x,\xi)\widetilde{f}\big),\, Wg\rangle \\ &= |\langle \rho(-x,\xi)\widetilde{f},\, g\rangle|^2 = |V(\widetilde{f},g)(-x,\xi)|^2. \quad \blacksquare \end{aligned}$$

Other examples of G's satisfying the condition $\int (Wf)G \geq 0$ for all f may be found in Janssen [85].

To understand the meaning of Proposition (1.99), let us consider a specific example. For $a > 0$, let

$$\phi_a^0(x) = (2a)^{n/4} e^{-\pi a x^2} \qquad \text{and} \qquad \Phi_a = W\phi_a^0.$$

A simple calculation shows that

$$\Phi_a(\xi, x) = 2^n e^{-2\pi(ax^2 + a^{-1}\xi^2)} = \phi_a(x)\phi_{1/a}(\xi) \quad \text{where} \quad \phi_a(x) = (2a)^{n/2} e^{-2\pi x^2},$$

so Proposition (1.99) shows that $Wf * \Phi_a \geq 0$ for all $f \in L^2$ and $a > 0$. Now, ϕ_a is a Gaussian of total mass 1 whose central peak has width roughly $\|X_j\phi_a\|_2 = 1/\sqrt{4\pi a}$, so convolving Wf by Φ_a more or less amounts to averaging Wf over balls of radius $1/\sqrt{4\pi a}$ in x and over balls of radius $\sqrt{a/4\pi}$ in ξ. The uncertainty principle does not allow position and momentum to be measured simultaneously with complete precision, but it allows position to be measured with an error ϵ and momentum to be measured with an error δ provided that $\epsilon\delta \geq (4\pi)^{-1}$. So it should be possible for the *averages* of position and momentum over balls of radius $\epsilon = 1/\sqrt{4\pi a}$ and $\delta = \sqrt{a/4\pi}$ to have a joint distribution—and this is $Wf * \Phi_a$.

More generally, if we average over larger balls (that is, convolve with more spread-out Gaussians) we get something strictly positive, while averaging over smaller balls doesn't work, as the following result of de Bruijn [37] shows:

(1.100) Proposition. *Let*

$$\Phi_{a,b}(\xi, x) = 2^n (ab)^{-n/2} \exp(-2\pi)\left(\frac{\xi^2}{a} + \frac{x^2}{b}\right), \qquad a, b > 0,$$

and suppose $f \in L^2(\mathbf{R}^n)$. If $ab = 1$ then $Wf * \Phi_{a,b} \geq 0$. If $ab > 1$ then $Wf * \Phi_{a,b} > 0$. If $ab < 1$ then $Wf * \Phi_{a,b}$ may be negative.

Proof: We have proved the first assertion above. If $ab > 1$, pick $c < a$ and $d < b$ with $cd = 1$. By the semigroup property of Gaussians (easily verified by taking Fourier transforms),

$$Wf * \Phi_{a,b} = (Wf * \Phi_{c,d}) * \Phi_{a-c,b-d}.$$

But $Wf * \Phi_{c,d} \geq 0$ and $\Phi_{a-c,b-d} > 0$, so the second assertion follows. As for the last one, we leave it as an exercise for the reader to verify that if $f(x) = xe^{-\pi x^2}$ on $\mathbf{R}$, or more generally $f(x) = x_1 e^{-\pi x^2}$ on $\mathbf{R}^n$, then $Wf * \Phi_{a,b}(0,0) < 0$ if $ab < 1$. ∎

There remains the question of when Wf is itself nonnegative. We have seen that this is the case when $f(x) = e^{-\pi a x^2}$. More generally, if

$$
f(x) = e^{-xAx+bx+c} \quad \text{where} \quad \begin{array}{l} A \in GL(n,\mathbf{C}),\ b \in \mathbf{C}^n,\ c \in \mathbf{C}, \\ \text{and } \operatorname{Re} A \text{ is positive definite,} \end{array} \tag{1.101}
$$

then $Wf > 0$. It is not hard to prove this by a brute force calculation of Wf, and the reader is welcome to do so. We shall return to this point in Section 4.5, when we shall have the machinery to make it utterly transparent. The remarkable thing, however, is that the functions (1.101) are the *only* ones with nonnegative Wigner distributions.

(1.102) Theorem. (Hudson [77]) *If $f \neq 0 \in L^2(\mathbf{R}^n)$ and $Wf \geq 0$ then f is of the form (1.101).*

Proof: For $z \in \mathbf{C}^n$ let $\psi_z(x) = e^{-\pi x^2 - 2\pi i z x}$. It is easily checked that

$$
W\psi_z(\xi, x) = 2^{n/2} \exp\left(-2\pi x^2 - 2\pi(\xi + \operatorname{Re} z)^2 + 4\pi(\operatorname{Im} z)x\right) > 0,
$$

so that if $Wf \geq 0$, (1.93) yields

$$
|\langle \psi_z, f\rangle|^2 = \int (W\psi_z)(Wf) > 0 \quad \text{for all } z.
$$

Therefore, if we set

$$
G(z) = \langle \psi_z, f\rangle = \int \overline{f(x)} e^{-\pi x^2 - 2\pi i z x}\, dx,
$$

G is a nonvanishing entire function of z. Moreover,

$$
\|\psi_z\|_2^2 = \int e^{-2\pi x^2 + 4\pi(\operatorname{Im} z)x}\, dx = 2^{n/2} e^{2\pi(\operatorname{Im} z)^2},
$$

so

$$
|G(z)| \leq \|f\|_2 \|\phi_z\|_2 \leq C e^{\pi |z|^2}.
$$

We claim that $G(z)$ is therefore of the form $e^{zAz+bz+c}$. Granted this, we must have $\operatorname{Re}(A)$ negative definite since $G|\mathbf{R}^n = (\overline{f} e^{-\pi x^2})\widehat{\ } \in L^2$. But then $\overline{f}(x) e^{-\pi x^2} = \mathcal{F}^{-1}(G|\mathbf{R}^n)(x)$ is of the form (1.101), and hence so is f.

As for the claim: this is a reasonably well known fact, at least in dimension 1, but here is a direct proof. Since G is entire and nonvanishing, the function

$$
H(z) = \log G(0) + \int_{[0,z]} \frac{G'(t)}{G(t)}\, dt
$$

(where $[0, z]$ is the line segment from 0 to z) is entire, and $G = e^H$. The bound on G means that

$$\operatorname{Re} H(z) \leq C' + \pi|z|^2. \tag{1.103}$$

Let $H(z) = \sum a_\alpha z^\alpha$. Given $r > 0$, we set

$$z(\theta_1, \ldots, \theta_n) = z(\theta) = (re^{i\theta_1}, \ldots, re^{i\theta_n}), \qquad 0 \leq \theta_j \leq 2\pi,$$

so that

$$\operatorname{Re} H(z(\theta)) = \tfrac{1}{2} \sum (a_\alpha r^{|\alpha|} e^{i\alpha\theta} + \overline{a}_\alpha r^{|\alpha|} e^{-i\alpha\theta}).$$

The right side is a Fourier series, so

$$\begin{aligned} a_\alpha r^{|\alpha|} &= \frac{2}{(2\pi)^n} \int_{[0,2\pi]^n} [\operatorname{Re} H(z(\theta))] e^{-i\alpha\theta} d\theta \quad \text{for} \quad \alpha \neq 0, \\ \operatorname{Re} a_0 &= \frac{1}{(2\pi)^n} \int_{[0,2\pi]^n} [\operatorname{Re} H(z(\theta))] d\theta, \end{aligned}$$

and hence

$$\begin{aligned} |a_\alpha| r^{|\alpha|} + 2\operatorname{Re} a_0 &\leq \frac{2}{(2\pi)^n} \int_{[0,2\pi]^n} \left[|\operatorname{Re} H(z(\theta))| + \operatorname{Re} H(z(\theta)) \right] d\theta \\ &= \frac{4}{(2\pi)^n} \int_{[0,2\pi]^n} \max(\operatorname{Re} H(z(\theta)),\, 0) d\theta \\ &\leq 4(C' + \pi r^2), \end{aligned}$$

by (1.103). Letting $r \to \infty$ we conclude that a_α must vanish unless $|\alpha| \leq 2$, so H is a polynomial of degree ≤ 2. ∎

A generalization of Hudson's theorem, pertaining to Wigner distributions of more general (non-L^2) functions, can be found in Janssen [87].

9. The Laguerre Connection

In this section we calculate the Fourier-Wigner and Wigner transforms of the Hermite functions. The answers turn out to involve the Laguerre polynomials $L_k^{(j)}$, defined for nonnegative integers j and k by

$$L_k^{(j)}(x) = \sum_{m=0}^{k} \frac{(k+j)!}{(k-m)!\,(j+m)!} \frac{(-x)^m}{m!}.$$

(For each j, the polynomials $L_k^{(j)}$ are orthogonal on $(0, \infty)$ with respect to the measure $x^j e^{-x}\,dx$.)

It suffices to consider the one-dimensional case, since n-dimensional Hermite functions are products of one-dimensional Hermite functions, and the Fourier-Wigner and Wigner transforms preserve the product structure; that is,

$$V(h_\alpha,\, h_\beta)(p, q) = \prod_{j=1}^{n} V(h_{\alpha_j},\, h_{\beta_j})(p_j, q_j),$$

and similarly for W.

(1.104) Theorem. *Suppose $p, q \in \mathbf{R}$ and $w = p + iq$. Then*

$$V(h_j, h_k)(p, q) = \begin{cases} \sqrt{\frac{k!}{j!}}\, e^{-(\pi/2)|w|^2} (\sqrt{\pi}\, w)^{j-k} L_k^{(j-k)}(\pi|w|^2) & \text{for } j \geq k, \\ \sqrt{\frac{k!}{j!}}\, e^{-(\pi/2)|w|^2} (-\sqrt{\pi}\, \overline{w})^{k-j} L_j^{(k-j)}(\pi|w|^2) & \text{for } j \leq k. \end{cases}$$

In particular,

$$V(h_j, h_j)(p, q) = e^{-(\pi/2)|w|^2} L_j^{(0)}(\pi|w|^2).$$

Proof: The simplest method is to perform the calculations in Fock space. We have

$$\begin{aligned} V(h_j, h_k)(p, q) &= \langle \rho(p, q) h_j,\, h_k \rangle = \langle \beta(w)\zeta_j,\, \zeta_k \rangle_{\mathcal{F}} \\ &= \sqrt{\frac{\pi^{j+k}}{j!k!}} \int e^{-\pi z \overline{w} - (\pi/2)|w|^2} (z + w)^j \overline{z}^k e^{-\pi |z|^2}\, dz \\ &= \sqrt{\frac{\pi^{j+k}}{j!k!}}\, e^{-(\pi/2)|w|^2} \sum_{m=0}^{j} \frac{j!\, w^{j-m}}{m!\,(j-m)!} \int z^m \overline{z}^k e^{-\pi z \overline{w} - \pi |z|^2} dz \end{aligned}$$

$$= \sqrt{\frac{\pi^{j+k}}{j!k!}}\, e^{-(\pi/2)|w|^2} \sum_{m=0}^{j} \frac{j!\, w^{j-m}}{m!\,(j-m)!} (-\pi)^{-m} \left(\frac{\partial}{\partial \overline{w}}\right)^m \int \overline{z}^k e^{-\pi z \overline{w} - \pi |z|^2} dz.$$

But

$$\int \overline{z}^k e^{-\pi z \overline{w} - \pi |z|^2} dz = \langle E_{-w},\, z^k \rangle_{\mathcal{F}} = \overline{\langle z^k,\, E_{-w} \rangle_{\mathcal{F}}} = (-\overline{w})^k.$$

Hence the sum in the last formula for $V(h_j, h_k)$ becomes

$$\sum_{m=0}^{\min(j,k)} \frac{j!\, w^{j-m}}{m!\,(j-m)!} (-1)^{k-m} \pi^{-m} \frac{k!\, \overline{w}^{k-m}}{(k-m)!}.$$

If $j \geq k$ we make the substitution $m \to k - m$, obtaining

$$V(h_j, h_k)(p,q) = \sqrt{\frac{k!}{j!}}\, e^{-(\pi/2)|w|^2} \sum_{m=0}^{k} \frac{j!\,(\sqrt{\pi}\,w)^{j-k}(-\pi|w|^2)^m}{(k-m)!\,(j+m)!\,m!}$$
$$= \sqrt{\frac{k!}{j!}}\, e^{-(\pi/2)|w|^2} (\sqrt{\pi}\,w)^{j-k} L_j^{(j-k)}(\pi|w|^2),$$

while if $j \leq k$ we make the substitution $m \to j - m$ and obtain similarly

$$V(h_j, h_k)(p,q) = \sqrt{\frac{j!}{k!}}\, e^{-(\pi/2)|w|^2} (-\sqrt{\pi}\,w)^{k-j} L_j^{(k-j)}(\pi|w|^2). \quad ▮$$

This result sems to have been first pointed out in the case $j = k$ by Klauder [90]. In the general case it was proved independently by Itzykson [81], Miller [106], and Vilenkin [145], and it has since been rederived by Peetre [116], Howe [74], and possibly others. Let us examine some consequences of it for the Laguerre functions

$$l_j(t) = e^{-t/2} L_j^{(0)}(t).$$

In the first place, the orthonormality of the l_j's in $L^2(0,\infty)$ follows immediately from the unitarity of V:

$$\int_0^\infty l_j(t) l_k(t)\,dt = \int_0^\infty \int_0^{2\pi} l_j(\pi r^2) l_j(\pi r^2) r\,d\theta\,dr$$
$$= \int_{\mathbf{C}} l_j(\pi|w|^2) l_k(\pi|w|^2)\,dw = \langle V(h_j,h_j),\, V(h_k,h_k)\rangle = \delta_{jk}.$$

Secondly, if we set

$$\mathcal{L}_j(p,q) = l_j\big(\pi(p^2+q^2)\big),$$

then, by Proposition (1.46), the operator $\rho(\mathcal{L}_j)$ is the orthogonal projection onto h_j. Hence, by (1.83), the spectral resolution of the Hermite operator $\pi(D^2 + X^2)$ is given by

$$\pi(D^2 + X^2) = \sum_0^\infty (j + \tfrac{1}{2})\rho(\mathcal{L}_j).$$

More generally, suppose $\phi(t)$ is any measurable function on $(0,\infty)$ that is $O\big((1+t)^N\big)$ for some N. Then we can expand ϕ in a series of Laguerre functions:

$$\phi = \sum_0^\infty c_j l_j, \qquad c_j = \int_0^\infty \phi(t)\lambda_j(t)\,dt.$$

(The series converges in L^2 if $\phi \in L^2$ or in a suitable weak sense for more general ϕ.) If

$$\Phi(p,q) = \phi\big(\pi(p^2+q^2)\big),$$

we then have $\Phi = \sum c_j \mathcal{L}^j$ (the sum converging at least in $\mathcal{S}'(\mathbf{R}^2)$), and $\rho(\Phi) = \sum c_j \rho(\mathcal{L}_j)$. We can therefore summarize the situation as follows:

$$\rho\Big[\phi\big(\pi(p^2+q^2)\big)\Big] = \widetilde{\phi}\big(\pi(D^2+X^2)\big),$$

where the right side denotes the ordinary functional calculus of self-adjoint operators and $\widetilde{\phi}$ is the function on the spectrum of $\pi(D^2+X^2)$ defined by

$$\widetilde{\phi}(j+\tfrac{1}{2}) = \int_0^\infty \phi(t) l_j(t)\, dt.$$

This result is due to Peetre [116], who used it to give Fourier-analytic derivations of various properties of Laguerre functions. (See also Itzykson [81].)

Closely related to these results is a formula due to Geller [55] that expresses the group Fourier transform of radial functions on $\mathbf{H}_n$ (i.e., functions $f(p,q,t)$ that depend only on p^2+q^2 and t) in terms of Laguerre transforms. The connection between Laguerre functions and Fourier analysis on $\mathbf{H}_n$ has been exploited in the study of various translation-invariant operators on $\mathbf{H}_n$ by de Michele and Mauceri [39], Jerison [88], Nachman [110], and Beals, Greiner, and Vauthier [20].

Next we compute $W(h_j,h_k)$. This can be done easily by applying equation (1.90) to our formula for $V(h_j,h_k)$, since the h_j's are all either odd or even. However, for the sake of variety, we shall present an independent derivation, following Janssen [85], that utilizes the Bargmann transform in a somewhat different way.

(1.105) Theorem. *Suppose $\xi, x \in \mathbf{R}$ and $z = x + i\xi$. Then*

$$W(h_j,h_k)(\xi,x) = \begin{cases} 2(-1)^k\sqrt{\frac{k!}{j!}}\, e^{-2\pi|z|^2}(2\sqrt{\pi}\, z)^{j-k} L_k^{(j-k)}(4\pi|z|^2) & \text{for } j \geq k, \\ 2(-1)^j\sqrt{\frac{j!}{k!}}\, e^{-2\pi|z|^2}(2\sqrt{\pi}\, z)^{k-j} L_j^{(k-j)}(4\pi|z|^2) & \text{for } j \leq k. \end{cases}$$

In particular,

$$Wh_j(\xi,x) = (-1)^j e^{-2\pi|z|^2} L_j^{(0)}(4\pi|z|^2).$$

Proof: For $u \in \mathbf{C}$, let

$$B_u(x) = B(u,x) = 2^{1/4} e^{2\pi ux - \pi x^2 - (\pi/2)u^2}.$$

Then by formula (1.89), $W(B_u, B_{\overline{v}})(\xi, x)$ equals

$$\begin{aligned}
&2^{1/2} e^{2\pi x(u+v) - (\pi/2)(u^2+v^2) - 2\pi x^2} \int e^{-2\pi i p[\xi + i(u-v)/2]} e^{(-\pi/2)p^2}\, dp \\
&= 2e^{2\pi x(u+v) - (\pi/2)(u^2+v^2) - 2\pi x^2 - 2\pi[\xi + i(u-v)/2]^2} \\
&= 2e^{-2\pi|z|^2} e^{2\pi(\overline{z}u + zv) - \pi uv} \\
&= 2e^{-2\pi|z|^2} \sum_{l,m,n=0}^{\infty} \frac{(-\pi)^l (2\pi\overline{z})^m (2\pi z)^n}{l!\, m!\, n!} u^{m+l} v^{n+l} \\
&= 2e^{-2\pi|z|^2} \sum_{j,k=0}^{\infty} u^j v^k \sum_{l=0}^{\min(j,k)} \frac{(-\pi)^l (2\pi\overline{z})^{j-l} (2\pi z)^{k-l}}{l!\, (j-l)!\, (k-l)!}.
\end{aligned}$$

On the other hand, by Theorem (1.86),

$$W(B_u, B_{\overline{v}}) = \sum_{j,k=0}^{\infty} \sqrt{\frac{\pi^{j+k}}{j!k!}}\, u^j v^k W(h_j, h_k).$$

Hence

$$W(h_j, h_k)(\xi, x) = 2e^{-2\pi|z|^2} \sqrt{\frac{j!k!}{\pi^{j+k}}} \sum_{l=0}^{\min(j,k)} \frac{(-\pi)^l (2\pi\overline{z})^{j-l} (2\pi z)^{k-l}}{l!\, (j-l)!\, (k-l)!},$$

and the same manipulations as in the proof of Theorem (1.104) can be used to express this last quantity in terms of Laguerre polynomials. ∎

Remark. If we compare the formulas for $V(h_j, h_j)$ and $Wh_j = W(h_j, h_j)$ in Theorems (1.104) and (1.105), we obtain the following result. If

$$F_j(w) = e^{-(\pi/2)|w|^2} L_j^{(0)}(\pi|w|^2), \qquad w \in \mathbf{C} \cong \mathbf{R}^2,$$

then

$$\widehat{F}_j(z) = 2(-1)^j F_j(2z).$$

If this equation is written out in polar coordinates, it reduces to the formula

$$\int_0^\infty e^{-r^2/2} L_j^{(0)}(r^2) J_0(rs) r\, dr = (-1)^j e^{-s^2/2} L_j^{(0)}(s^2)$$

(where J_0 is the Bessel function of order zero), which can be found in tables of Hankel transforms.

10. The Nilmanifold Representation

In this section we discuss yet another interesting way of realizing the irreducible representations of $\mathbf{H}_n$. We restrict attention to the representation ρ and leave the generalization to ρ_h, $h \neq 1$, to the reader.

Here it will be convenient to use the polarized form $\mathbf{H}_n^{\mathrm{pol}}$ of the Heisenberg group. We recall that this is $\mathbf{R}^{2n+1}$ with the group law

$$(p,q,t)(p',q',t') = (p+p',\, q+q',\, t+t'+pq'),$$

and that the map $\alpha : \mathbf{H}_n^{\mathrm{pol}} \to \mathbf{H}_n$ defined by

$$\alpha(p,q,t) = (p,\, q,\, t - \tfrac{1}{2}pq)$$

is an isomorphism. The representation ρ of $\mathbf{H}_n$ corresponds to the representation $\rho^{\mathrm{pol}} = \rho \circ \alpha$ of $\mathbf{H}_n^{\mathrm{pol}}$, given by

$$\rho^{\mathrm{pol}}(p,q,t)f(x) = e^{2\pi i(t+qx)}f(x+p) = e^{2\pi i t}e^{2\pi i qX}e^{2\pi i pD}f(x). \tag{1.106}$$

Let Γ denote the subset of $\mathbf{H}_n^{\mathrm{pol}}$ consisting of points whose coordinates are all integers:

$$\Gamma = \{(p,q,t) \in \mathbf{H}_n^{\mathrm{pol}} : \quad p,q \in \mathbf{Z}^n \text{ and } t \in \mathbf{Z}\}.$$

Then Γ is a discrete subgroup of $\mathbf{H}_n^{\mathrm{pol}}$, and the right coset space

$$M = \Gamma\backslash\mathbf{H}_n^{\mathrm{pol}}$$

is a compact nilmanifold. It is easily verified that the half-open unit cube

$$Q^{2n+1} = [0,1)^{2n+1} \subset \mathbf{H}_n^{\mathrm{pol}}$$

is a fundamental domain for Γ, that is, each right coset of Γ contains precisely one point of Q^{2n+1}. Hence M can be considered topologically as the closed unit cube $\overline{Q}^{2n+1}$ with certain pieces of its boundary identified with each other, a sort of "twisted torus." (In fact, M is a nontrivial circle bundle over the $2n$-torus.) Moreover, Haar measure on $\mathbf{H}_n^{\mathrm{pol}}$—namely, Lebesgue measure—induces an invariant measure on M, so measure-theoretically we can think of M as the cube Q^{2n+1} with Lebesgue measure. We shall identify functions on M with Γ-invariant functions on $\mathbf{H}_n^{\mathrm{pol}}$ or with functions on Q^{2n+1}, as the occasion warrants.

The action of $\mathbf{H}_n^{\mathrm{pol}}$ on M by right translation determines the regular representation R of $\mathbf{H}_n^{\mathrm{pol}}$ on $L^2(M)$:

$$R(X)f(\Gamma Y) = f(\Gamma Y X).$$

$L^2(M)$ breaks up into a sum of R-invariant subspaces $\mathcal{H}_j$ according to the action of the center of $\mathbf{H}_n^{\mathrm{pol}}$:

$$\mathcal{H}_j = \{f \in L^2(M) : R(0,0,t)f = e^{2\pi ijt} f\}, \qquad j \in \mathbf{Z}.$$

In other words,

$$f \in \mathcal{H}_j \iff f(p,q,t) = e^{2\pi ijt} f(p,q,0). \tag{1.107}$$

If we think of $f \in L^2(M)$ as a function on Q^{2n+1}, the expansion $f = \sum_{-\infty}^{\infty} f_j$, $f_j \in \mathcal{H}_j$, is just the Fourier series of f in the variable t:

$$f_j(p,q,t) = e^{2\pi ijt} \int_{Q^{2n+1}} f(p,q,\tau) e^{-2\pi ij\tau}\, d\tau.$$

By the Stone–von Neumann theorem, the restriction of the representation R to $\mathcal{H}_j$ is equivalent to a direct sum of copies of ρ_j. Our interest here is in the case $j = 1$, where, as we shall see, the restriction of R is irreducible.

Let us define a map T from functions on $\mathbf{R}^n$ to functions on $\mathbf{H}_n^{\mathrm{pol}}$ as follows:

$$Tf(p,q,t) = \sum_{K \in \mathbf{Z}^n} f(p+K) e^{2\pi iKq} e^{2\pi it}.$$

If $f \in \mathcal{S}(\mathbf{R}^n)$, say, this series clearly converges nicely to a C^∞ function on $\mathbf{H}_n^{\mathrm{pol}}$. Moreover, if $(a,b,j) \in \Gamma$,

$$\begin{aligned} Tf\big((a,b,j)(p,q,t)\big) &= Tf(p+a,\, q+b,\, t+j+aq) \\ &= \sum f(p+a+K) e^{2\pi iK(q+b)} e^{2\pi i(t+j+aq)} \\ &= \sum f(p+a+K) e^{2\pi i(a+K)q} e^{2\pi it}. \end{aligned}$$

But on relabeling the index K of summation as $K-a$, we see that this last sum is nothing but $Tf(p,q,t)$. Thus Tf is Γ-invariant, and we may (and do) regard it as a function on M. As such, Tf is in $\mathcal{H}_1$ by (1.107), and we have

$$\begin{aligned} \|Tf\|_{L^2(M)}^2 &= \int_{Q^{2n+1}} |Tf(p,q,t)|^2\, dp\, dq\, dt \\ &= \int_{[0,1)^{2n}} \Big|\sum f(p+K) e^{2\pi iKq}\Big|^2\, dp\, dq \\ &= \sum \int_{[0,1)^n} |f(p+K)|^2\, dp \qquad \text{(by Parseval)} \\ &= \|f\|_2^2, \end{aligned}$$

so T is an isometry from $L^2(\mathbf{R}^n)$ into $\mathcal{H}_1$. T is actually surjective onto $\mathcal{H}_1$, for if $g \in \mathcal{H}_1$ we can expand g in a Fourier series on Q^{2n+1}:

$$g(p,q,t) = e^{2\pi it} \sum_{J,K\in\mathbf{Z}^n} a_{JK} e^{2\pi i(Jp+Kq)}, \qquad (p,q,t) \in Q^{2n+1}.$$

We then have $g = Tf$ where f is the function defined piecemeal on $\mathbf{R}^n$ by

$$f(x+K) = \sum_J a_{JK} e^{2\pi iJx} \quad \text{for} \quad x \in [0,1)^n, \quad K \in \mathbf{Z}^n.$$

Finally, observe that

$$\begin{aligned} T\left[\rho^{\text{pol}}(p,q,t)f\right](p',q',t') &= \sum e^{2\pi i[t+q(p'+K)]} f(p'+p+K) e^{2\pi iKq'} e^{2\pi it'} \\ &= \sum f(p'+p+K) e^{2\pi iK(q'+q)} e^{2\pi i(t'+t+p'q)} \\ &= \left[R(p,q,t)Tf\right](p',q',t'). \end{aligned}$$

We have therefore proved:

(1.109) Theorem. *The transform T defined by (1.108) is a unitary map from $L^2(\mathbf{R}^n)$ to $\mathcal{H}_1$ which intertwines ρ^{pol} and $R|\mathcal{H}_1$. In particular, $R|\mathcal{H}_1$ is an irreducible unitary representation of $\mathbf{H}_n^{\text{pol}}$ that is equivalent to ρ^{pol}.*

Remark. For $|j| \geq 2$, the representation $R|\mathcal{H}_j$ is not irreducible, but it is of finite multiplicity. See Auslander [7] or Brezin [26] for a detailed analysis of its structure. For $j = 0$, functions on $\mathcal{H}_0$ can be identified with functions on the $2n$-torus $\mathbf{R}^{2n}/\mathbf{Z}^{2n}$, and the irreducible subspaces are the one-dimensional spans of the functions $e^{2\pi i(Jp+Kq)}$, $J, K \in \mathbf{Z}^n$.

The central variable t enters the above calculations, as usual, in a rather trivial way, so we can provide an alternative description of the space $\mathcal{H}_1$ and the transform T that does not mention it. Namely, if $f \in \mathcal{H}_1$, f is determined according to (1.107) by the function

$$f_0(p,q) = f(p,q,0)$$

on $\mathbf{R}^{2n}$. (Here we regard f as a function on $\mathbf{H}_n^{\text{pol}}$.) The Γ-invariance of f translates into the following quasi-periodicity property of f_0:

$$f_0(p+a,\ q+b) = e^{-2\pi iaq} f_0(p,q) \quad \text{for} \quad a, b \in \mathbf{Z}^n. \tag{1.110}$$

A function satisfying (1.110) is competely determined by its values on the unit cube Q^{2n} in $\mathbf{R}^{2n}$, and its absolute value is actually periodic in all variables.

Moreover, the norm of f in $L^2(M)$ equals the norm of f_0 in $L^2(Q^{2n})$. Hence if we define

$$\mathcal{H}^1 = \left\{ f : f \text{ satisfies (1.110) and } \int_{Q^{2n}} |f(x)|^2 dx < \infty \right\},$$

the map $f \to f_0$ is unitary from $\mathcal{H}_1$ to $\mathcal{H}^1$. The corresponding map $T_0 : L^2(\mathbf{R}^n) \to \mathcal{H}^1$ is given by

$$T_0 f(p,q) = (Tf)_0(p,q) = \sum_{K \in \mathbf{Z}^n} f(p+K) e^{2\pi i K q}, \tag{1.111}$$

and the corresponding representation R_0 of $\mathbf{H}_n^{\mathrm{pol}}$ on $\mathcal{H}^1$ is given by

$$R_0(r,s,t) f(p,q) = f(p+r,\, q+s) e^{2\pi i (t+sp)}. \tag{1.112}$$

The transform T and its close relative T_0 are referred to in the literature as the **Weil-Brezin transform** (the name we shall adopt) or the **Zak transform.** T_0 was described by Weil [151, pp. 164-5], although it is implicitly present in much earlier works; and T was introduced later by Brezin [26]. Meanwhile, T_0 and the representation R_0 were discovered independently by Zak [158], [159], who found them useful for solving problems in solid state physics involving motion in a periodic potential and for studying other quantum phenomena in which periodic variables occur. Zak calls the representation R_0 (or rather its infinitesimal version) the **kq representation**, k and q being the standard names of the quasi-position and quasi-momentum variables in solid state physics.

We now discuss some interesting properties of the Weil-Brezin transform T_0 and the representation R_0.

(i) The infinitesimal representation dR_0 of $\mathfrak{h}_n$ on $\mathcal{H}^1$ is given by

$$dR_0(r,s,t) f = \sum_1^n \left(r_j \frac{\partial f}{\partial p_j} + s_j \frac{\partial f}{\partial q_j} + 2\pi i s_j p_j f \right) + 2\pi i t f.$$

This follows easily from (1.112). A similar formula holds for the infinitesimal representation dR on $\mathcal{H}_1$, provided that one considers elements of $\mathcal{H}_1$ as Γ-invariant functions on $\mathbf{H}_n^{\mathrm{pol}}$ rather than functions on M. (The coordinates p_j are not well defined on M. The reader may check that if $f \in \mathcal{H}_1$, the function $\partial f / \partial q_j + 2\pi i p_j f$ is Γ-invariant although $\partial f / \partial q_j$ and $2\pi i p_j f$ individually are not.)

(ii) The operators $\rho(a,0) = e^{2\pi i a D}$ and $\rho(0,b) = e^{2\pi i b X}$ on $L^2(\mathbf{R}^n)$ commute when a and b are both in $\mathbf{Z}^n$, and the Weil-Brezin transform provides the spectral resolution for this family of commuting normal operators. Indeed, if $a \in \mathbf{Z}^n$,

$$T_0(e^{2\pi i a D} f)(p,q) = \sum f(p+a+K)e^{2\pi i K q}$$
$$= \sum f(p+K)e^{2\pi i (K-a)q} = e^{-2\pi i a q} T_0 f(p,q),$$

and if $b \in \mathbf{Z}^n$,

$$T_0(e^{2\pi i b X} f)(p,q) = \sum f(p+K)e^{2\pi i b(p+K)}e^{2\pi i K q}$$
$$= e^{2\pi i b p} T_0 f(p,q).$$

Combining these results with the fact that $\rho(a,0)\rho(0,b) = \epsilon^{\pi i a b}\rho(a,b)$, we have

$$T_0\big(\rho(a,b)f\big)(p,q) = (-1)^{ab} e^{2\pi i(bp-aq)} T_0 f(p,q) \quad \text{for} \quad a,b \in \mathbf{Z}^n. \tag{1.113}$$

(iii) T_0 is not only an isometry from $L^2(\mathbf{R}^n)$ to $L^2(Q^{2n})$ but also a contraction from $L^p(\mathbf{R}^n)$ to $L^p(Q^{2n})$ for $1 \le p \le 2$. Indeed, for $p = 1$ we have

$$\|T_0 f\|_1 = \int_{Q^{2n}} \left|\sum f(p+K)e^{2\pi i K q}\right| dp\, dq$$
$$\le \int_{Q^{2n}} \sum |f(p+K)|\, dp = \|f\|_1,$$

and the case $1 < p < 2$ follows by interpolation.

(iv) The following slightly bizarre property of the space $\mathcal{H}^1$ was observed by Zak and Janssen [86].

(1.114) Proposition. *Every continuous function in $\mathcal{H}^1$ has zeros.*

Proof: Suppose $f \in \mathcal{H}^1$ is continuous. Since $\mathbf{R}^{2n}$ is simply connected, if f were nonvanishing we could write $f(p,q) = e^{2\pi i \psi(p,q)}$ for some continuous function ψ. The quasi-periodicity (1.110) implies that for every $a,b \in \mathbf{Z}^n$ there exists $K_{a,b} \in \mathbf{Z}$ such that

$$\psi(p+a,\, q+b) = \psi(p,q) - aq + K_{a,b}.$$

But this is self-contradictory: on the one hand,

$$\psi(a,b) = \psi(0,b) - ab + K_{a,0} = \psi(0,0) - ab + K_{a,0} + K_{0,b},$$

and on the other,

$$\psi(a,b) = \psi(a,0) + K_{0,b} = \psi(0,0) + K_{a,0} + K_{0,b}. \qquad \blacksquare$$

(v) We now assume $n = 1$. Given $\tau \in \mathbf{C}$ with $\operatorname{Im}\tau > 0$, let

$$\phi_\tau(x) = e^{\pi i \tau x^2}.$$

Then $\phi_\tau \in L^2(\mathbf{R})$, and we have

$$T_0\phi_\tau(u,v) = \sum_{k=-\infty}^{\infty} e^{2\pi i k v + \pi i \tau (u+k)^2} = e^{\pi i \tau u^2}\vartheta_3(z,q),$$

where

$$z = \pi(v + \tau u), \qquad q = e^{\pi i \tau}, \qquad \vartheta_3(z,q) = \sum_{-\infty}^{\infty} q^{k^2} e^{2ikz}.$$

ϑ_3 is one of the basic Jacobi theta functions, in the notation of Whittaker and Watson [154], the others being

$$\begin{aligned}
\vartheta_1(z,q) &= -ie^{iz+i\pi\tau/4}\vartheta_3(z + \tfrac{1}{2}\pi(\tau - 1),\, q),\\
\vartheta_2(z,q) &= \vartheta_1(z + \tfrac{1}{2}\pi,\, q) = e^{iz+i\pi\tau/4}\vartheta_3(z + \tfrac{1}{2}\pi\tau,\, q),\\
\vartheta_4(z,q) &= \vartheta_3(z - \tfrac{1}{2}\pi,\, q),
\end{aligned}$$

where $q = e^{\pi i \tau}$ throughout. Since

$$T_0\big(\rho^{\mathrm{pol}}(r,s)\phi_\tau\big)(u,v) = R_0(r,s)T_0\phi_\tau(u,v) = e^{2\pi i s u}T_0\phi_\tau(u+r,\, v+s),$$

these other theta functions can be obtained (up to factors involving only elementary exponential functions) as Weil-Brezin transforms of the functions $\rho^{\mathrm{pol}}(r,s)\phi_\tau$ where $s + \tau r$ is $\frac{1}{2}(\tau - 1)$, $\frac{1}{2}\tau$, or $\frac{1}{2}$.

These relations suggest that the Heisenberg group and the nilmanifold M should be of use in the study of theta functions. That is indeed the case, and this connection has been much exploited in recent years. See Auslander [7], Auslander–Tolimieri [8], Igusa [80], and Mumford [109].

11. Postscripts

The Heisenberg group plays a role in many other parts of analysis besides the subjects discussed in this monograph. In this section we provide a brief description of some of these areas and a few selected references, mostly expository works from whose bibliographies the reader can obtain a more complete guide to the literature. At the outset, let us mention the article of Howe [75], which surveys several aspects of the Heisenberg group that are considered here from a somewhat different point of view.

Several Complex Variables. As is well known, the unit disc in $\mathbf{C}$ can be mapped onto the upper half plane by a fractional linear transformation, and the boundary of the upper half plane (namely the real axis) can be identified with the group of horizontal translations of the plane. There is a similar situation in higher dimensions. We work in $\mathbf{C}^{n+1}$ and denote points in $\mathbf{C}^{n+1}$ by (ζ, τ) where $\zeta \in \mathbf{C}^n$ and $\tau \in \mathbf{C}$. The analogue of the unit disc is the unit ball,

$$B_{n+1} = \{(\zeta, \tau) \in \mathbf{C}^{n+1} : |\zeta|^2 + |\tau|^2 < 1\},$$

and the analogue of the upper half plane is the Siegel domain

$$D_{n+1} = \{(\zeta, \tau) \in \mathbf{C}^{n+1} : \operatorname{Im} \tau > |\zeta|^2\}.$$

It is easily checked that the fractional linear transformation

$$\phi(\zeta, \tau) = \left(\frac{\zeta}{i(\tau - 1)}, \frac{\tau + 1}{i(\tau - 1)}\right)$$

maps B_{n+1} onto D_{n+1}. Finally, the analogue of the horizontal translation group is the Heisenberg group $\mathbf{H}_n$, which acts on $\mathbf{C}^{n+1}$ by holomorphic affine transformations:

$$L_{(z,t)}(\zeta, \tau) = (\zeta + z, \tau - 4t + i|z|^2 - 2i\bar{z}\zeta).$$

Here we are using complex cooridnates $z = p + iq$ on $\mathbf{H}_n$ as we did in discussing the Bargmann transform. The reader may verify that L is indeed a left action,

$$L_{(z,t)}L_{(z',t')} = L_{(z,t)(z',t')},$$

and that the transformations $L_{(z,t)}$ map the domain D_{n+1} and its boundary ∂D_{n+1} onto themselves. The action on ∂D_{n+1} is simply transitive, so $\mathbf{H}_n$ can be identified with ∂D_{n+1} by the correspondence

$$(z, t) \longleftrightarrow L_{(z,t)}(0, 0) = (z, i|z|^2 - 4t). \tag{1.115}$$

At this point it should be said that in the complex analysis literature it is customary to use a different parametrization of $\mathbf{H}_n$. Namely, one replaces the coordinate t by $-t/4$ so that $(z, t) \in \mathbf{H}_n$ becomes identified with $(z, t + i|z|^2) \in \partial D_{n+1}$.

Since the action L of $\mathbf{H}_n$ is holomorphic, the Cauchy-Riemann operators on D_{n+1} are invariant under it, and the induced complex of operators on ∂D_{n+1}, the so-called $\bar{\partial}_b$ complex, can actually be considered as a complex of left-invariant operators on $\mathbf{H}_n$ via the identification (1.115). One can then apply Fourier-analytic techniques on $\mathbf{H}_n$ to study these operators in detail.

The unit ball and the Siegel domain D_{n+1} are the simplest members of the class of strongly pseudoconvex domains, which is of fundamental importance in the theory of several complex variables. One can show that if $\Omega \subset \mathbf{C}^{n+1}$ is strongly pseudoconvex, for any $P \in \partial\Omega$ there is a holomorphic coordinate system with origin at P in which $\partial\Omega$ closely approximates ∂D_{n+1} near the origin. Using this fact, the analysis on the Heisenberg group can be transferred to $\partial\Omega$ to yield refined information about the $\overline{\partial}$ and $\overline{\partial}_b$ complexes on general strongly pseudoconvex domains. This program was initiated in Folland–Stein [51]; see also the survey articles of Folland [49], Stanton [130], and Beals-Fefferman-Grossman [15], and the references given there.

The identification of $\mathbf{H}_n$ with ∂D_{n+1} also leads to another derivation of the Fock-Bargmann representation, as was pointed out to the author by F. Ricci. Transfer Lebesgue measure on $\mathbf{H}_n$ to ∂D_{n+1}, and consider the subspace H^2 of $L^2(\partial D_{n+1})$ consisting of functions that are nontangential limits of holomorphic functions on D_{n+1}. If $F \in H^2$ and $h \in \mathbf{R}$, let

$$F_h(\zeta) = \int_{\operatorname{Im}\tau=|\zeta|^2} F(-\zeta,\tau)e^{-\pi ih\tau/2}d\tau = \int_{\mathbf{R}} F(-\zeta,\, s+i|\zeta|^2)e^{-\pi ih(s+i|\zeta|^2)/2}ds.$$

(We use $-\zeta$ instead of ζ just to make the formulas below turn out more neatly.) The first equation shows that $F_h(\zeta)$ is an entire function of ζ for each h, as the contour of integration can be deformed to be locally independent of ζ. The second one shows that $e^{-\pi h|\zeta|^2/2}F_h(\zeta)$ is, for each ζ, the Fourier transform of the function $s \to F(-\zeta,\, s+i|\zeta|^2)$, evaluated at $h/4$. Since the latter function extends holomorphically to the half plane $\operatorname{Im} s > 0$, one sees by Cauchy's theorem that $F_h(\zeta) = 0$ for $h < 0$. Moreover, by the Plancherel theorem,

$$\|F\|_2^2 = \tfrac{1}{4}\int_0^\infty\int_{\mathbf{C}^n} |F_h(\zeta)|^2 e^{-\pi h|\zeta|^2}d\zeta\, dh.$$

This formula exhibits H^2 as the direct integral of the Fock spaces $\mathcal{F}_n^h$ $(h > 0)$ defined in Section 1.6.

The Heisenberg group acts unitarily on H^2 by left translation:

$$U_{(z,t)}F(\zeta,\tau) = F\big(L_{(z,t)}^{-1}(\zeta,\tau)\big) = F(\zeta - z,\, \tau + 4t + i|z|^2 - 2i\overline{z}\zeta).$$

Applying the transform $F \to F_h$, we obtain

$$\begin{aligned}
\big(U_{(z,t)}F\big)_h(\zeta) &= \int F(-\zeta - z,\, s + i|\zeta|^2 + 4t + i|z|^2 + 2i\overline{z}\zeta)e^{-\pi ih(s+i|\zeta|^2)/2}ds \\
&= \int F(-\zeta - z,\, s + 4t - 2\operatorname{Im}\overline{z}\zeta + i|\zeta + z|^2)e^{-\pi ih(s+i|\zeta|^2)/2}ds \\
&= \int F(-\zeta - z,\, s + i|\zeta + z|^2)e^{-\pi ih(s-4t+2\operatorname{Im}\overline{z}\zeta+i|\zeta|^2)/2}ds \\
&= e^{-\pi ih(-4t-2i\overline{z}\zeta-i|z|^2)/2}\int F(-\zeta - z,\, s + i|\zeta + z|^2)e^{-\pi ih(s+i|\zeta+z|^2)/2}ds \\
&= e^{2\pi iht-\pi h\overline{z}\zeta-(\pi h/2)|z|^2}F_h(\zeta + z),
\end{aligned}$$

which is the Fock-Bargmann representation on $\mathbf{H}_n$ on $\mathcal{F}_n^h$. (For $h < 0$, one plays the same game with $\overline{H}^2$, the space of boundary values of antiholomorphic functions on D_{n+1}.)

Representation Theory. The Stone–von Neumann theorem provided inspiration for, and is a paradigmatic special case of, two of the fundamental results of modern representation theory: the Mackey imprimitivity theorem and the Kirillov classification of irreducible unitary representations of nilpotent Lie groups. See Mackey [99] for a lucid account of the path that leads from the Stone–von Neumann theorem to the imprimitivity theorem, and Mackey [100] for more information on the imprimitivity theorem and its applications. For the Kirillov theory and some of its extensions, see Moore [107] and Wallach [150].

Partial Differential Equations and Harmonic Analysis. In the past two decades a considerable amount of study has been devoted to partial differential operators constructed from non-commuting vector fields, in which the non-commutativity plays an essential role in determining the regularity properties of the operators. (One of the most important instances of this situation is the $\overline{\partial}_b$ complex on the boundary of a domain on $\mathbf{C}^{n+1}$, as discussed above.) The operators of this sort that can be most readily analyzed are left-invariant operators on graded nilpotent Lie groups (such as the Heisenberg group) that are homogeneous with respect to the natural dilations on these groups. In this setting one can develop non-Abelian, non-isotropic analogues of many of the tools of Euclidean harmonic analysis—singular integrals, Green's functions, various function spaces (Sobolev, Lipschitz, Hardy, etc.). These techniques, together with the representation theory of the groups, yield precise results for invariant differential operators, which can then be transferred to more general operators. Among the foundational papers in this subject are Folland–Stein [51], Rothschild–Stein [124], and Rockland [123]; see also Folland [49], Folland–Stein [52], Helffer–Nourrigat [68], Taylor [136], and Taylor [137].

In 1957 Hans Lewy shocked the world of analysis by producing the first example of a differential equation that is not locally solvable. Lewy's unsolvable operator is nothing but $(X + iY)$ on $\mathbf{H}_1$, to which he was led because of its connection with complex analysis. (It is essentially $\overline{\partial}_b$ on $\mathbf{H}_1$.) More recently, Greiner, Kohn, and Stein [60] have used the techniques mentioned above to give a complete characterization of the functions g for which $(X + iY)f = g$ is solvable.

Abstract Heisenberg Groups. There is an analogue of the Heisenberg group (or, more precisely, of the reduced, polarized Heisenberg group) associated to an arbitrary locally compact Abelian group G. If G is such a group, let T denote the group of complex numbers of modulus one and $\widehat{G}$ the group of continuous homomorphisms from G to T. ($\widehat{G}$ is a locally compact group with the compact-open topology.) The **Heisenberg group** of G is the locally compact group

$\mathbf{H}(G)$ whose underlying space is $G \times \widehat{G} \times T$ and whose group law is

$$(g, \chi, z)(g', \chi', z') = \big(gg',\ \chi\chi',\ zz'\chi'(g)\big).$$

$\mathbf{H}(G)$ has a family of Schrödinger representations ρ_j, indexed by a nonzero integer j, that act on $L^2(G)$ by

$$\rho_j(g, \chi, z)f(g') = z^j \chi(g')^j f(gg').$$

The analogue of the Stone–von Neumann theorem (again a special case of the Mackey imprimitivity theorem) is the fact that up to unitary equivalence, the ρ_j's exhaust all irreducible unitary representations of $\mathbf{H}(G)$ that are nontrivial on the center. (The representations that are trivial on the center are just the characters of $G \times \widehat{G}$, lifted to $\mathbf{H}(G)$.) The groups $\mathbf{H}(G)$, in which G is an adèle group or the additive group of a vector space over a local field, have been applied to problems in number theory by Weil [151].