Essential Self-Adjointness of the Symplectic Dirac Operators

by

A. Nita

B.A., University of California, Irvine, 2007

M.A., University of Colorado, Boulder, 2010

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Mathematics

2016

This thesis entitled:
Essential Self-Adjointness of the Symplectic Dirac Operators
written by A. Nita
has been approved for the Department of Mathematics

Prof. Alexander Gorokhovsky

Prof. Carla Farsi

Date ____________________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Nita, A. (Ph.D., Mathematics)

Essential Self-Adjointness of the Symplectic Dirac Operators

Thesis directed by Prof. Alexander Gorokhovsky

The main problem we consider in this thesis is the essential self-adjointness of the symplectic Dirac operators D and $\tilde{D}$ constructed by Katharina Habermann in the mid 1990s. Her constructions run parallel to those of the well-known Riemannian Dirac operators, and show that in the symplectic setting many of the same properties hold. For example, the symplectic Dirac operators are also unbounded and symmetric, as in the Riemannian case, with one important difference: the bundle of symplectic spinors is now infinite-dimensional, and in fact a Hilbert bundle. This infinite dimensionality makes the classical proofs of essential self-adjointness fail at a crucial step, namely in local coordinates the coefficients are now seen to be unbounded operators on $L^2(\mathbb{R}^n)$. A new approach is needed, and that is the content of these notes. We use the decomposition of the spinor bundle into countably many finite-dimensional subbundles, the eigenbundles of the harmonic oscillator, along with the simple behavior of D and $\tilde{D}$ with respect to this decomposition, to construct an inductive argument for their essential self-adjointness. This requires the use of ancillary operators, constructed out of the symplectic Dirac operators, whose behavior with respect to the decomposition is transparent. By an analysis of their kernels we manage to deduce the main result one eigensection at a time.

Dedication

I would like to dedicate this thesis to my parents, Alexander and Maria Nita, who were angels throughout the entire project, giving me the support and encouragement without which I could not have made it. Their selfless dedication to me throughout this difficult intellectual endeavor will never be forgotten.

Acknowledgements

I would like to thank first and foremost my advisor Alexander Gorokhovsky, whose patience and guidance saw this project through to its conclusion through the many hurdles I faced along the way. Sebastian Casailina-Martin also provided much encouragement and opportunities to flesh out much of the material on analysis on manifolds, through talks I gave in his complex geometry seminar. I would also like to thank Carla Farsi for giving me the opportunity to talk about K-theory and algebraic topology, and Jeffrey Fox for several important discussions on the relationship between physics and index theory. There are many others I should thank, particularly Marty Walter and Pedro Berrizbeitia, who took valuable time away from their own work to come to my aid in the final stages of the thesis.

Contents

Chapter

Chapter 1

Introduction

We shall give here a proof of the essential self-adjointness of the symplectic Dirac operators D and $\tilde{D}$ acting on sections the symplectic spinor bundle $\mathbf{Q}$ over a compact symplectic manifold (M, ω). This is a preliminary result in the larger quest for a symplectic analog of the well-known Atiyah-Singer index theorem on a Riemannian manifold. To understand the importance of Dirac operators and their self-adjointness in this quest, it is necessary to briefly review the historical development of the classical index theorem, its generalizations, and various proofs.

The classical index theorem originated in the famous papers of Atiyah, Singer and Segal of the 1960s and early 1970s, [5], [6], [7], [8], [9], and [10], and gave an equivalence between the analytic index of an elliptic operator P acting on sections of a complex vector bundle E over a Riemannian manifold (M, g) and that manifold's topological index, via the equation

$$\operatorname{index}(P) = \{\operatorname{ch}(P) \cdot \mathcal{T}(M)\}[M]$$

The term on the left is the analytic index, defined as $\operatorname{index}(P) = \dim \ker P - \dim \operatorname{coker}(P)$, while the term on the right consists of topological invariants: the Chern character, a ring homomorphism $\operatorname{ch} : K(M) \to H^{\bullet}(M; \mathbb{Q})$, taking the symbol of P, realized as an element of the K-theory ring $K(M)$, into the rational cohomology ring $H^{\bullet}(M; \mathbb{Q})$ of M, the Todd class $\mathcal{T}(M)$ of M, and the fundamental class $[M]$ of M. The motivation for this theorem was Gel'fand's 1960 observation [36] that the analytical index $\operatorname{index}(P)$ of an elliptic operator is always homotopy invariant, and so should describe a topological invariant. It had also been proven by Borel and Hirzebruch in

1958, [18], that the $\hat{A}$-genus of a spin manifold was always an integer, even though this topological quantity was defined in terms of an infinite series. Atiyah and Singer's paper [5] explained this unusual result as the index of the manifold's Dirac operator.

The first announcement of the general theorem came in that same 1963 paper [5], though its proof arrived somewhat later, in Palais' 1965 book [86]. According to Atiyah and Singer themselves, [6], this early proof used the computationally taxing and ill suited methods of cobordism and rational cohomology groups, which failed to reveal the basic underlying mechanics of the theorem's simple statement, while also making difficult the theorem's generalizations. The three papers of 1968, [6], [7], [8], replaced cobordism and homology with just K-theory, which was more in line with Grothendieck's outlook on such matters.

The years 1968-1973 saw several advances, the most important of which, for us, was the introduction heat equation methods to the proof method of the index theorem. These brought with them some revealing simplifications in certain cases, and reduced the dependency on topological machinery with the introduction of more analytic tools. The first inroads were made by McKean and Singer, [80], in 1967. They looked at the Dirac operator $d + d^* : \Omega^{ev}(M) \to \Omega^{odd}(M)$ on a Riemannian manifold (M, g) to prove the index theorem equating the index of $d + d^*$ to the Euler characteristic of M. Observing that the integrand α, given by the heat equation methods, and the normalized Gauss curvature K had the same total integral, they asked whether $\alpha = K$ locally, by some remarkable cancellations of higher order terms in local coordinates. Patodi [87] in 1971 answered affirmatively. As recounted in Atiyah, Bott and Patodi [11], the drawbacks to Patodi's methods were the complicated algebraic process of cancellation. Gilkey's 1973 approach [40] showed that the higher derivatives could be eliminated on a priori grounds, so that the integrand α should never involve higher derivatives. The drawback here, however, was with the lengh and difficulty of the proofs. In Atiyah et al's estimation, this was due to the reliance on general differential operator theory, rather than the Riemannian geometry at hand. With a change in tactics, toward the Riemannian geometric tools, Atiyah, Bott and Padodi succeeded in simplifying the analytical proof of Gilkey enough to make it manageable.

The last step in this direction, using the heat equation methods and cancellation of higher order terms in the local expansion, came with Getzler's 1985 paper [38], which introduced his eponymous symbol calculus involving only the top order terms in the asymptotic expansion of the heat kernel of $\not{D}^2$. This paper gave a clear and simple proof of the local index theorem, and now serves as the model in the smooth category (see for example Melrose [81]). The Dirac operator case is a specific case of the general index theorem [5], of course, and it seems that it was a bit of mathematical folklore that the Dirac case could imply the general topological case. The present author has just seen two new preprints by Paul Baum and Erik van Erp on just this topic, [13], [14]. These papers give a proof of the implication of the general index theorem from the special Dirac case, using K-homology and Bott periodicity. So it seems that the general elliptic case and the special Dirac operator case are in fact equivalent.

The heat kernel approach to the classical Atiyah-Singer index theorem by means of Dirac operators is indeed the model of our investigations. Of course, the symplectic category is vastly different from the Riemannian category, first and foremost because symplectic manifolds do not have local invariants, as a consequence of Darboux's theorem. The topological side, therefore, is quite different and employs ideas deriving more from algebraic geometry than from differential geometry. Our goal here, however, is to understand only the analytical side of the hypothetical 'index theorem.' We aim merely to lay the analytical groundwork for heat kernel methods.

The first step in this direction is ensuring that our Dirac operators are self-adjoint, or at least essentially self-adjoint. Generally speaking, self-adjointness of unbounded operators on a Hilbert space are important because their spectrum is real, and thus the functional calculus may be applied to them to get continuous or even Borel functions of these operators. For example, we may exponentiate them, or take square roots, if the spectrum is nonnegative. Merely symmetric operators may not have real spectrum, and there are examples of such cases, for instance $i\frac{d}{dx}$ on $L^2([0,1])$ with domain $\{f \in AC([0,1]) \mid f(0) = f(1) = 0\}$. Essentially self-adjoint operators are symmetric operators whose *closures* are self-adjoint, and these enjoy the same benefits as self-adjoint operators, only without our needing to specify their domains ahead of time. In the case of our symmetric differential operators D and $\tilde{D}$, the application of the functional calculus we are most interested is in exponention. Then we can apply the heat kernel methods to try to make some connections with the underlying *symplectic* topology of our manifold.

The notion of spinors came to light in Brauer and Weyl's 1935 paper [20], following earlier investigations of Cartan and originating in Dirac's work on the spin of the electron. Spinors may be upgraded to spinor fields on a Riemannian manifold, and here begins the notion of a spin manifold, whose topology was studied by Borel and Hirzebruch [18], and thence by Atiyah and Singer [5] by means of Dirac operators. The symplectic analog of a spinor was introduced by Kostant [65] in 1974, and the symplectic Dirac operators were introduced in 1995 by Habermann [50]. Subsequent investigations of the symplectic Dirac operators were conducted by Habermann in the papers [51], [52] and [53]. The construction of these operators was completely analogous to the classical Dirac operators on a spin manifold, as operators acting on sections of the spin bundle.

The difference in the symplectic case is that we replace the Clifford algebra bundle, inside which sits the spin group and its Lie algebra, with the Weyl algebra bundle, whose sections are the symplectic spinor fields. The Weyl algebra bundle uses the symplectic form ω instead of the Riemannian metric g to give the relations $u \cdot v - v \cdot u = -\omega(u, v)1$ (this is the symplectic analog to $u \cdot v + v \cdot u = -g(u, v)1$ in the Riemannian setting). The skew-symmetry of ω has important consequences for the Weyl algebra bundle, the most eye-catching of which is that it makes the Weyl algebra infinite-dimensional. Another difference is that the representations of the metaplectic group, the double cover of the symplectic group, are not matrix representations. They are unitary representations on the Hilbert space $L^2(\mathbb{R}^n)$, and are unique up to isomorphism. This is in contrast to the representation of the spin group, the double cover of the special orthogonal group, which is a matrix representation on $\mathbb{C}^{2^n}$ (and variants thereof on $\mathbb{R}^k$ or $\mathbb{H}^\ell$, where k and ℓ depend on $n = \dim M$, see Lawson and Michelsohn [71]). These facts point to the need for infinite-dimensional Hilbert bundles of spinors over M rather than finite-dimensional. The implications for the symplectic Dirac operators are that, unlike the Riemannian case, the Dirac operators have unbounded coefficient operators in local coordinates, making the traditional proofs of their essential self adjointness (Wolf [110], Chernoff [25]) difficult to emulate.[1]

Another difficulty arising in the symplectic case is that, though we show the symplectic Dirac operators to be essentially self-adjoint, they are also known to not be elliptic, unlike the Riemannian case. Their commutator $P = i[\tilde{D}, D]$ is, however, elliptic, and this fact will be useful for a variety of purposes. These technical differences, particularly the jump to infinite dimensions, make the symplectic Dirac operator case both more unusual and also more interesting given the intricate topological constructions available in the symplectic category (see McDuff and Salamon [76], [77]).

Let us now briefly describe the structure of this thesis. Since the setting for our Dirac operators is the smooth symplectic category, following some general observations on bilinear vector spaces, we begin Chapter 2 with the *linear* symplectic category, consisting of symplectic vector spaces and linear symplectic morphisms. This is necessary because the fibers of the tangent bundle of a symplectic manifold are symplectic vector spaces. The model for the linear symplectic category is the standard symplectic vector space $(\mathbb{R}^{2n}, \omega_0)$, where $\omega_0 = \sum_{j=1}^{n} dp^j \wedge dq^j$ is given in terms of the dual basis vectors $dp^j = \mathbf{e}_{n+j}^*$ and $dq^j = \mathbf{e}_j^*$ of the standard orthonormal basis vectors. The notation here is meant to recall the physicists' convention of viewing $\mathbb{R}^{2n}$ as the phase space of a free particle. The configuration space of such a particle is $\mathbb{R}^n$, and coordinates $q_j = \mathbf{e}_j$ are position coordinates of the particle. The momentum coordinates p_j should technically lie in the cotangent bundle $T^*\mathbb{R}^n = \mathbb{R}^n \times (\mathbb{R}^n)^*$ of $\mathbb{R}^n$, whose coordinates are $(q_1, \ldots, q_n, dp^1, \ldots, dp^n)$. However, by use of the nondegenerate form ω_0 we may identify $\mathbb{R}^n$ with its dual, and by this means we identify $T^*\mathbb{R}^n$ with $\mathbb{R}^{2n}$.

As we will see, all symplectic vector spaces (V, ω) are linearly symplectomorphic to the standard symplectic space $(\mathbb{R}^{2n}, \omega_0)$, so the nature of the whole linear symplectic category boils down to that of this space. Now, $\mathbb{R}^{2n}$ is also a Euclidean vector space with standard metric g_0 the

[1] Habermann, in her original paper [50], cited Wolf's proof as applicable to the symplectic Dirac operators, but the present author couldn't understand how that proof could be modified to work in the symplectic case.

dot product. By means of the standard complex structure, an endomorphism

$$J_0 = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$$

of $\mathbb{R}^{2n}$ squaring to $-I_{2n}$, we may also endow $\mathbb{R}^{2n}$ with a complex vector space structure of complex dimension n. Complex-scalar multiplication is given by

$$(a + ib)\mathbf{v} := a\mathbf{v} + bJ(\mathbf{v})$$

making $\mathbb{R}^{2n}_{J_0}$ complex-isomorphic to $\mathbb{C}^n$. The general linear group $\mathrm{GL}(2n, \mathbb{R})$ acts transitively by conjugation on the space $\mathcal{J}(\mathbb{R}^{2n})$ of all complex structures on $\mathbb{R}^{2n}$, with stabilizer isomorphic to $\mathrm{GL}(n, \mathbb{C})$, so we have the diffeomorphism

$$\mathcal{J}(\mathbb{R}^{2n}) \approx \mathrm{GL}(2n, \mathbb{R}) / \mathrm{GL}(n, \mathbb{C})$$

Each $J \in \mathcal{J}(\mathbb{R}^{2n})$ similarly endows $\mathbb{R}^{2n}$ with a complex structure, by exactly the same procedure as above, giving an isomorphisms $\mathbb{R}^{2n}_J \cong \mathbb{C}^n$. Another $\mathrm{GL}(2n, \mathbb{R})$-action, this time on the space of symplectic forms $\Omega(\mathbb{R}^{2n})$, shows the diffeomorphism

$$\Omega(\mathbb{R}^{2n}) \approx \mathrm{GL}(2n, \mathbb{R}) / \mathrm{Sp}(n, \mathbb{R})$$

where $\mathrm{Sp}(n, \mathbb{R})$ may be identified with the stabilizer of ω_0 under this action. We will also demonstrate that the space of metrics $\mathcal{M}(\mathbb{R}^{2n})$ is diffeomorphic to the positive definite symmetric matrices $\mathrm{Sym}_{2n}(\mathbb{R})^+$,

$$\mathcal{M}(\mathbb{R}^{2n}) \approx \mathrm{Sym}_{2n}(\mathbb{R})^+$$

So we see that there is no shortage of symplectic forms, complex structures, and metrics on $\mathbb{R}^{2n}$. The natural question, then, is, 'What sort of interactions there between these structures?'

The most important interaction is ω-compatibility. Given a symplectic form ω on $\mathbb{R}^{2n}$, a complex structure J is said to be compatible with ω if $g(u, v) := \omega(u, J(v))$ defines a metric. It will be shown in this chapter that the space of such ω-compatible complex structures is also large, and even connected and contractible. Moreover, we also have a converse condition: given a complex structure J on $\mathbb{R}^{2n}$, there always exists a symplectic form ω such that J is ω-compatible, and the space of such symplectic forms is convex and contractible. We may conclude, therefore, that the linear symplectic category is well-equipped with compatibile Euclidean, symplectic and complex structures, in particular ω-compatibility is a property not only of J for given ω, but of ω given J.

Moving to the smooth symplectic category such compatibility conditions do not always occur. One additional feature of symplectic manifolds is the closedness of the symplectic form $\omega \in \Omega^2(M)$. Fiberwise, ω may be identified with the standard linear symplectic form ω_0, but globally we require more. From this condition alone we see why compatibility becomes more difficult in the smooth setting. The class of manifolds with compatible Riemannian, symplectic and complex structures is rather restrictive, being the Käler class. It follows from the fiberwise case that manifolds equipped with symplectic structures are always equipped with compatible almost complex structures, and conversely. The term 'almost' in the smooth setting refers to the potential lack of an integrability condition on J, meaning simply that J may not come from a complex structure on the manifold M. The obstructions to compatibility are by nature topological, and though it is not our focus in these notes, we will pause and state some of the known results in this area which are relevant to understanding the various subclasses of symplectic manifolds.

Following this, we describe the basic objects associated to symplectic manifolds: Hamiltonian vector fields, the Poisson bracket, symplectomorphisms, and symplectic connections. We will also sketch a proof of the fundamental Darboux theorem, which asserts that locally all symplectic manifolds are symplectomorphic to an open subset of the standard symplectic space $(\mathbb{R}^{2n}, \omega_0)$. This is the key feature distinguishing the symplectic category from the Riemannian.

After the geometric setting has been adequately described, we return to analysis. We dedicate Chapter 3 to the construction of the relevant functions spaces needed to handle unbounded differential operators: Sobolev spaces, locally convex spaces and distributions. We do this first in the Euclidean setting, then describe the method of transferring these spaces to manifolds, even in the noncompact case. Then we use these spaces to enlarge the class of differentiable functions to weakly differentiable functions. The embedding theorems and the Fourier transform are recalled, before describing the general theory of differential operators on a manifold. Invariant definitions are given, and filtered algebras are constructed, before we give the definition of the symbol of a differential operator. The symbol's importance lies in the fact that elliptic operators are defined to be those whose symbol is always invertible fiberwise off the zero vector. This naturally leads into elliptic theory. We describe elliptic regularity, which is of fundamental importance for us. In its simplest version, it states that if P is elliptic and a (not necessarily smooth) section s satisfies the weak equality $Ps = 0$, then s is in fact smooth. Thus, the kernel of an elliptic operator P consists of *smooth* sections. If P is also self-adjoint, then its spectrum is particularly simple, being real and discrete, and we will describe these points in detail.

Before venturing into elliptic theory, however, we will spend some time on the general theory of self-adjoint unbounded operators on a Hilbert space $\mathcal{H}$. Section 3.2 of Chapter 3 is devoted to complete proofs of von Neumann's famous theorems characterizing self-adjoint unbounded operators. Importantly for us, we will see why an unbounded symmetric operator T on a Hilbert space $\mathcal{H}$ is self-adjoint if and only if $\ker(T^* \pm iI) = \{0\}$. This characterization is the one we will employ in our proof of the essential self-adjointness of the symplectic Dirac operators D and $\tilde{D}$.

This completed, we then embark in Chapter 4 on the technicalities of the construction of the symplectic spinor bundle $\mathbf{Q}$ and, in Chapter 5, on the construction of the symplectic Dirac operators acting on its sections. The Weyl algebra is defined and its properties studied, before shifting gears and considering the metaplectic representation. The metaplectic group $\mathrm{Mp}(n, \mathbb{R})$ is the double cover of the symplectic group $\mathrm{Sp}(n, \mathbb{R})$, the structure group of any symplectic vector bundle (in particular the tangent bundle of a symplectic manifold), and the fact that it has no matrix representation, but only one (up to isomorphism) unitary representation on $L^2(\mathbb{R}^n)$, leads to the unusual properties of symplectic spinors. Firstly, the symplectic spinor bundle $\mathbf{Q}$ is the associated $L^2(\mathbb{R}^n)$ Hilbert vector bundle

$$\mathbf{Q} = P_{\mathrm{Mp}(n,\mathbb{R})} \times_{\mathfrak{m}} L^2(\mathbb{R}^n)$$

to the metaplectic prinicpal bundle via the metaplectic representation $\mathfrak{m}$. This is an infinite-dimensional Hilbert bundle, and symplectic spinors are sections of this bundle. The smooth sections turn out to be precisely those sections with image in $P_{\mathrm{Mp}(n,\mathbb{R})} \times_{\mathfrak{m}} \mathcal{S}(\mathbb{R}^n)$, the Schwartz space-valued symplectic spinors. The Lie algebras $\mathfrak{mp}(n, \mathbb{R})$ and $\mathfrak{sp}(n, \mathbb{R})$ of the metaplectic and symplectic groups, respectively, may be embedded in the Weyl algebra, and the tangent map $d\rho$ of the double covering map $\rho : \mathrm{Mp}(n, \mathbb{R}) \to \mathrm{Sp}(n, \mathbb{R})$ may be understood in terms of the Weyl algebra product. Once the position and momentum coordinates of the base manifold (M, ω) are quantized, $q_j \mapsto ix_j$ (a multiplication operator on $L^2(\mathbb{R}^n)$), $p_j \mapsto \partial/\partial x_j$, the quantization map is shown to satisfy the defining Weyl algebra relation, allowing us to extend it to the entire Weyl algebra. This allows us to define symplectic Clifford multiplication, $X \cdot s$, multiplying a symplectic spinor field s by a vector field X. We can understand this as a linear combination of position and momentum operators acting on s fiberwise. These are unbounded operators, and here is where the symplectic case begins to seriously diverge from the Riemannian case.

Given a symplectic connection ∇ on (M, ω), we may lift it to the symplectic spinor bundle $\mathbf{Q}$. The symplectic Dirac operators are then defined to be the compositions

$$\Gamma(\mathbf{Q}) \xrightarrow{\nabla} \Gamma(T^*M \otimes \mathbf{Q}) \xrightarrow[g^\sharp]{\omega^\sharp} \Gamma(TM \otimes \mathbf{Q}) \xrightarrow{\mu} \gamma(\mathbf{Q})$$

where $\omega^\sharp$ and $g^\sharp$ are isomorphisms between T^*M and TM deriving from the symplectic and Riemannian structure of M, and μ is the symplectic Clifford product described above. These are the operators studied by Katharina Habermann in the papers [50], [51], [52] and [53]. Under appropriate conditions on M, D and $\tilde{D}$ are symmetric, and it is the subject matter of this thesis to show that they are essentially self-adjoint.

Traditional functional analytic methods tend to fail because of the unboundedness of symplectic Clifford multiplication. In local coordinates, this means that D and $\tilde{D}$ have coefficients which are unbounded operators on the fibers $L^2(\mathbb{R}^n)$. There is, however, some extra information we can work with here. Our main tool is the harmonic oscillator. In $\mathbb{R}^n$, the harmonic oscillator H_0 is known to be essentially self-adjoint and to have the Hermite functions as eigenfunctions. These form a complete orthonormal system for $L^2(\mathbb{R}^n)$, and therefore decompose $L^2(\mathbb{R}^n)$ into finite-dimensional eigenspaces of H_0. The metaplectic representation, restricted to the double cover of the unitary group, the maximal compact subgroup of the symplectic group, commutes with H_0, thus giving a decomposition of the symplectic spinor bundle $\mathbf{Q}$ into finite-dimensional subbundles $\mathbf{Q}_\ell$. Our symplectic Dirac operators D and $\tilde{D}$ combine to construct ancillary Dirac-type operators $Z = D + i\tilde{D}$ and $Z^* = D - i\tilde{D}$, which have very simple behavior with respect to the decomposition of $\mathbf{Q}$. Z moves sections of $\mathbf{Q}_\ell$ down by one to sections of $\mathbf{Q}_{\ell-1}$, and Z^* moves sections up by one. This simple behavior is exploited, along with information about the kernels of the operators D, $\tilde{D}$, Z, Z^* and $P = i[\tilde{D}, D]$, to obtain an inductive proof that any symplectic spinor in $\ker(D^* \pm iI)$ must be 0. Von Neumann's theorem then applies to say that D and $\tilde{D}$ are essentially self-adjoint. Along the way we prove some related results, and look at the local behavior of our symplectic Dirac operators. Elliptic regularity is employed in statements about the kernel of P, which is known to be elliptic.

Chapter 2

Symplectic Manifolds

2.1 Preliminaries on Bilinear Spaces

2.1.1 Matrix Representations of Bilinear Forms

Let V be a vector space over a field F and let B be a bilinear form on V (we will also use the notation $\langle\cdot,\cdot\rangle$ for a bilinear form, and we denote the set of all bilinear forms on V by $\mathrm{Hom}^2_F(V;F)$). When V is finite-dimensional with ordered basis $\beta = (v_1, \ldots, v_n)$, any bilinear form B on V possesses a **matrix representation** $[B]_\beta \in M_n(F)$ with respect to β, namely

$$[B]_\beta := \begin{pmatrix} B(v_1,v_1) & B(v_1,v_2) & \ldots & B(v_1,v_n) \\ B(v_2,v_1) & B(v_2,v_2) & \ldots & B(v_2,v_n) \\ \vdots & \vdots & \ddots & \vdots \\ B(v_n,v_1) & B(v_n,v_2) & \ldots & B(v_n,v_n) \end{pmatrix} \tag{2.1}$$

Moreover, the action of B in matrix terms is given by the following proposition. It's proof is straightforward so we omit it.

Proposition 1 *If V is a finite-dimensional vectors space over a field F and β is an ordered basis for V, then for any bilinear form $B \in \mathrm{Hom}^2_F(V;F)$ an any vectors $x, y \in V$ we have*

$$B(x,y) = [x]^T_\beta [B]_\beta [y]_\beta \qquad \blacksquare$$

Theorem 1 *The function $\varphi_\beta : \mathrm{Hom}^2_F(V;F) \to M_n(F)$, $\varphi_\beta(B) := [B]_\beta$, is an isomorphism. Therefore, $\mathrm{Hom}^2_F(V;F) \cong M_n(F)$ and $\dim(\mathrm{Hom}^2_F(V;F)) = n^2$.* $\blacksquare$

We now ask what happens when we switch bases from β to γ on V. The answer is $[B]_\beta$ and $[B]_\gamma$ are congruent matrices, and conversely congruent matrices X and Y represent the same bilinear form B with respect to different bases β and γ. This follows directly from the change of basis theorem from linear algebra and the previous theorem.

Theorem 2 *Let V be an n-dimensional vector space over a field F, and let β and γ be two ordered bases for V. If $M_{\beta,\gamma}$ is the change of coordinate matrix changing β-coordinates to γ-coordinates, then a bilinear form $B \in \mathrm{Hom}_F^2(V;F)$ gives congruent matrix representations with respect to the two bases:*

$$[B]_\beta = M_{\beta,\gamma}^T [B]_\gamma M_{\beta,\gamma} \tag{2.2}$$

Conversely, if two matrices $X, Y \in M_n(F)$ are congruent, that is

$$X = P^T Y P \tag{2.3}$$

for some $P \in \mathrm{GL}(n,F)$, and if $Y = [B]_\gamma$ represents a bilinear form B on V with respect to a basis γ, then $X = [B]_\beta$ represents B with respect to some other basis β for V, and $P = M_{\beta,\gamma}$ is the change of coordnates matrix.

Proof: By Change of Coordinates Theorem in linear algebra, for all $v \in V$ we have $[v]_\gamma = M_{\beta,\gamma}[v]_\beta$, which means

$$B(x,y) = [x]_\gamma^T [B]_\gamma [y]_\gamma = (M_{\beta,\gamma}[v]_\beta)^T [B]_\gamma M_{\beta,\gamma}[v]_\beta = [v]_\beta^T \big(M_{\beta,\gamma}^T [B]_\gamma M_{\beta,\gamma}\big)[v]_\beta$$

Conversely, if $X = P^T Y P$ and $Y = [B]_\gamma$, then we must have $P = M_{\beta,\gamma}$. ∎

2.1.2 Orthogonality and Isotropy

In what follows we assume (V, B) is a bilinear space. We establish here some terminology.

Two vectors u and v in V are called **orthogonal** if $B(u,v) = 0$, and this fact is denoted by $u \perp v$. We remark here that $B(u,v) = 0$ does not necessarily imply $B(v,u) = 0$ (see the section on reflexivity below). We also write $u \perp S$ when we mean to say $u \perp s$ for all $s \in S$ for any subset S

of V, and similarly we write $S \perp v$ and $S \perp T$. The direct sum $U \oplus S$ for subspaces U and S of V is written $U \bigoplus S$ when we also have $U \perp S$.

Define two associated maps $B_L \in \mathrm{Hom}_F(V, V^*)$ and $B_R \in \mathrm{Hom}_F(V, V^*)$ by

$$B_L(v) := B(v, \cdot)$$

$$B_R(w) := B(\cdot, w)$$

The kernel of B_L is also called the **left radical** of the space (V, B), and the kernel of B_R is called the **right radical** of (V, B), and the notation

$$\mathrm{rad}_L(V) \equiv^{\perp} V := \ker B_L \qquad \text{and} \qquad \mathrm{rad}_R(V) \equiv V^{\perp} := \ker B_R$$

is used. If W is a subspace of V, we write

$$^{\perp}W := \{v \in V \mid v \perp w \text{ for all } w \in W\} = \ker(B_L : V \to W^*)$$

and

$$W^{\perp} := \{v \in V \mid w \perp v \text{ for all } w \in W\} = \ker(B_R : V \to W^*)$$

and call these the **left orthogonal complement** and the **right orthogonal complement** of W, respectively. We do not necessarily have $^{\perp}W = W^{\perp}$ (see the reflexivity section below).

The **norm** of a vector $v \in V$ is the field element $B(v, v)$.

A vector $v \in V$ is called **isotropic** if $v \neq 0$ and $B(v, v) = 0$, i.e. $v \perp v$, while if $B(v, v) \neq 0$ it is called **anisotropic**. An **isotropic subspace** U of V is one which contains an isotropic vector, and an **anisotropic subspace** contains no isotropic vectors. A **totally isotropic subspace** U of V consists *entirely* of isotropic vectors. When applied to the entirety of V, this terminology about subpaces being isotropic or anisotropic is sometimes carried over to the bilinear form B, which is thus isotropic or anisotropic according as V is one or the other. This is the case especially when B is a symmetric form (see below).

Remark 1 When B is symplectic (that is nondegenerate and skew-symmetric, see below), then the term 'isotropic' has a different meaning. Let us explain why. Since B is alternating whenever it is skew-symmetric, by Theorem 3 in the next section, assuming of course $\mathrm{ch}(F) \neq 2$, we already have that $B(v, v) = 0$ for all $v \in V$, so that V is trivially totally isotropic in the symplectic case in the sense of isotropy introduced above. In the symplectic case, then, an **isotropic subspace** U of V is one which is contained in its orthogonal complement, $U \subseteq U^{\perp}$, that is $u \perp v$ for all $u, v \in U$. ■

2.1.3 Symmetry and Skew-Symmetry

We recall that there are three important classes of bilinear forms: symmetric, skew-symmetric and alternating.

(1) A **symmetric** bilinear form satisfies $B(x, y) = B(y, x)$ for all $x, y \in V$.

(2) A **skew-symmetric** bilinear form satisfies $B(x, y) = -B(y, x)$ for all $x, y \in V$.

(3) An **alternating** bilinear form satisfies $B(x, x) = 0$ for all $x \in V$.

A bilinear space (V, B) is called a **symmetric space**, a **skew-symmetric space** or an **alternating space** according to it's bilinear form (see Szymiczek [97], Greub [42], Roman [92]).

Theorem 3 *If* $\mathrm{ch}(F) \neq 2$, *then any bilinear form* $B \in \mathrm{Hom}_F^2(V, F)$ *is alternating iff it is skew-symmetric, and the zero form is the only bilinear form that is both symmetric and skew-symmetric.*

Proof: If B is alternating, then for all $x, y \in V$ we have $B(x + y, x + y) = 0$. Expanding this and noting that $B(x, x) = B(y, y) = 0$ gives the result, $B(x, y) + B(y, x) = 0$. For the converse B is skew-symmetric, then for all $x \in V$ we have $B(x, x) = -B(x, x)$, or $2B(x, x) = 0$, and therefore, since $2 \neq 0$ in F, $B(x, x) = 0$. For the last statement, suppose that for all $x, y \in V$ we have $B(x, y) = -B(y, x) = B(y, x)$, where the first equality is by skew-symmetry and the second by symmetry, then $2B(x, y) = 0$, so $B(x, y) = 0$. ■

Next, we observe that in a sense symmetric and skew-symmetric bilinear forms are the two types of forms out of which all bilinear forms are made.

Theorem 4 *Let* V *be a vector space over a field* F *with* $\mathrm{ch}(F) \neq 2$, *and let*

$$\mathrm{Hom}_{F,\mathrm{sym}}^2(V; F)$$

denote the subspace of $\mathrm{Hom}_F^2(V;F)$ *of all symmetric biliniear forms on* V*, and let*

$$\mathrm{Hom}_{F,\mathrm{skew}}^2(V;F) \quad \textit{or} \quad \mathrm{Hom}_{F,\mathrm{alt}}^2(V;F)$$

denote the subspace of all skew-symmetric (equiv. alternating) bilinear forms on V*. Then,*

$$\mathrm{Hom}_F^2(V;F) = \mathrm{Hom}_{F,\mathrm{sym}}^2(V;F) \oplus \mathrm{Hom}_{F,\mathrm{skew}}^2(V;F) \qquad \blacksquare$$

Corollary 1 *If* $\mathrm{ch}(F) \neq 2$*, then, letting* $\mathrm{Sym}_n(F)$ *and* $\mathrm{Skew}_n(F)$ *denote, respectively, the spaces of symmetric and skew-symmetric* $n \times n$ *matrices with entries in* F*, we have*

$$M_n(F) = \mathrm{Sym}_n(F) \oplus \mathrm{Skew}_n(F) \qquad \blacksquare$$

2.1.4 Reflexivity and Degeneracy

Another reason for focusing on the symmetric and skew-symmetric forms is that these are the only reflexive forms.

We assume as usual that $\mathrm{ch}(F) \neq 2$. Two vectors $x, y \in V$ are said to be **orthogonal** if $B(x,y) = 0$ and this fact is denoted $x \perp y$. Orthogonality defines a binary relation $\perp$ on V. In the case that $\perp$ is reflexive, that is $x \perp y \implies y \perp x$, or in other words $B(x,y) = 0 \implies B(y,x) = 0$, the bilinear form B is called **reflexive**. Reflexivity is needed in order to be able define the kernel, or radical, of B and determine to what extent B is degenerate (that is, to what extent it collapses subspaces of V to 0). A priori these are not well defined ideas, since B acts potentially differently in each coordinate. In fact, not all bilinear forms are reflexive; an example is the bilinear form B on the set of 2×2 real matrices given by $B(X,Y) = \sum_{i,j=1}^2 (XY)_{ij}$. Obviously symmetric and skew-symmetric forms are symmetric, and as we will show below, these are the only reflexivity forms on V.

Theorem 5 *Let* V *be a vector space over a field* F*. A bilinear form* $B \in \mathrm{Hom}_F^2(V,F)$ *is reflexive iff it is either symmetric or skew-symmetric.*

Proof: Suppose first that B is reflexive. Note that for all $x, y, z \in V$ we have

$$B(x, B(x,y)z - B(x,z)y) = B(x,y)B(x,z) - B(x,z)B(x,y) = 0 \tag{2.4}$$

Reflexivity then implies

$$B(x,y)B(z,x) - B(x,z)B(y,x) = B(B(x,y)z - B(x,z)y, x) = 0 \tag{2.5}$$

Choosing $z = x$ shows that

$$B(x,x)(B(x,y) - B(y,x)) = 0 \tag{2.6}$$

for all $x, y \in V$. From this it is apparent that either $B(x,x) = 0$ or $B(x,y) - B(y,x) = 0$ for all $x, y \in V$. If $B(x,y) - B(y,x) = 0$ for all $x, y \in V$, then B is symmetric. Otherwise, if we suppose $B(x,y) - B(y,x) \neq 0$ for some $x, y \in V$, we will show that $B(z,z) = 0$ for all $z \in V$, and this implies that B is skew-symmetric, because $0 = B(x+y, x+y) = B(x,x) + B(x,y) + B(y,x) + B(y,y)$ implies $B(x,y) = -B(y,x)$.

Suppose, therefore, that for some $x, y \in V$ we have $B(x,y) - B(y,x) \neq 0$. In that case, $B(x,x) = 0$. Take any $z \in V$. We must show that $B(z,z) = 0$. If $B(z,x) \neq B(x,z)$, then by (2.6) we have $B(z,z) = 0$. So suppose $B(z,x) = B(x,z)$, and likewise suppose $B(z,y) = B(y,z)$. Then, from (2.5) we get that

$$\begin{aligned} 0 &= (x,y)B(z,x) - B(x,z)B(y,x) \\ &= B(z,x)B(x,y) - B(z,x)B(y,x) \\ &= B(z,x)(B(x,y) - B(y,x)) \end{aligned}$$

Since we assumed $B(x,y) - B(y,x) \neq 0$, we must have $B(z,x) = 0$. Similarly we get that $B(z,y) = 0$. By the assumed reflexivity we also have $B(x,z) = B(y,z) = 0$, and consequently

$$B(x+z, y) = B(x,y) + B(z,y) = B(x,y) \neq B(y,x) = B(y,x) + B(y,z) = B(y, x+z) \tag{2.7}$$

Using (2.6) with $x' = x + z$ and $y' = y$ we get

$$0 = B(x', x')(B(x', y') - B(y', x')) = B(x+z, x+z)\overbrace{(B(x+z, y) - B(y, x+z))}^{\neq 0 \text{ by } (2.7)}$$

so that we must have $B(x+z, x+z) = 0$. Expanding this and noting that $B(x,x) = B(z,x) = 0$ we get

$$0 = B(x+z, x+z) = \overbrace{B(x,x) + B(x,z) + B(z,x)}^{=0} + B(z,z) = B(z,z)$$

The converse is immediate, for if B is symmetric or alternating, then $B(x,y) = 0$ implies $B(y,x) = \pm B(x,y) = 0$. ■

If B is reflexive, then the right and left radicals are the same, and we may speak of *the* **kernel** of B and *the* **radical** of (V, B). In this case,

$$\ker B = \mathrm{rad}(V) = V^{\perp}$$

and all of these are equal to $\ker B_L = \ker B_R$. We say that a reflexive form B, or sometimes the space (V, B), is **nondegenerate** or **nonsingular** if

$$\ker B = \mathrm{rad}(V) = \{0\}$$

We may equally well speak of the $\mathrm{rank}(B_L)$ and $\mathrm{rank}(B_R)$, or in the case that B is reflexive, of $\mathrm{rank}\, B$, by the same reasoning.

Proposition 2 *If (V, B) is a finite-dimensional reflexive bilinear space, then B is nondegenerate iff it's matrix $[B]_\beta$ with respect to any basis β for V is nonsingular.*

Proof: B is nondegenerate iff $\ker(B) = \ker(B_L) = \ker(B_R) = \{0\}$ iff $\ker([B]_\beta) = \{0\}$, where $[B]_\beta$ can be viewed as acting by multiplication on the left, i.e. as a matrix representation of B_R. ■

Remark 2 In the case that B is reflexive and nondegenerate, and V is finite-dimensional, we get that B_L and B_R are isomorphisms between V and V^*. When B is an inner product, B_L is usually denoted

$$B^\flat := B_L \qquad \text{or} \qquad B_L(v) = v^\flat$$

and it's inverse B_L^{-1} is denoted by

$$B^\sharp := B_L^{-1} \qquad \text{or} \qquad B_L^{-1}(\alpha) = \alpha^\sharp$$

These are called the **flat** and **sharp** operators, respectively. Similar notation is used when B is a symplectic form. ■

Remark 3 (Nondegenerate Subspaces and Riesz Representation) When V is a finite-dimensional space, and B is a reflexive bilinear form on V, then, even if B fails to be nondegenerate on all of V, it may nevertheless be so on a subspace W of V. In this case, we can restrict B to $W \times W$, and get the standard isomorphisms $B^\flat = B_L$ and $B^\sharp = B_L^{-1}$ between W and W^*. Moreover, if we restrict only the second component of B to W, that is we allow B_L to have domain all of V, then we get a linear map

$$B_L|_{V\times W} : V \to W^*, \qquad B_L(v) := B(v, \cdot)|_W$$

The kernel of this map is clearly $W^\perp$,

$$\ker B_L|_{V\times W} = W^\perp$$

and moreover the finite-dimensional Riesz representation theorem applies to give the existence of a unique $v \in W$ such that $B_L|_{V\times W}(w) = B(v, w)$ for all $w \in W$. Even if W is nondegenerate we can still find a representing v, this time in any of V, though it may not be unique. (See Theorem 11.6 in Roman [92]). ■

Proposition 3 *If B is a reflexive bilinear map on a finite-dimensional F-vector space V, then for any subspace we have that*

$$\mathrm{rad}(W) = W \cap W^\perp$$

and thus $B|_{W\times W}$ is nondegenerate iff $\mathrm{rad}(W) = \{0\}$.

Proof: Since $\mathrm{rad}(W) = \ker(B|_{W\times W}) = \{w \in W \mid B(w, v) = 0,\ \forall v \in W\} = W \cap W^\perp$ and $B_L|_W \in \mathrm{GL}(W)$ iff $\ker(B_L|_W) = \{0\}$ iff $B_L|_W$ and therefore $B|_{W\times W}$ is nondegenerate. ■

Theorem 6 *If (V, B) is a finite-dimensional reflexive bilinear space, then for any subspace W of V we have*

$$\dim W + \dim W^{\perp} = \dim V + \dim(W \cap V^{\perp})$$

Consequently, if W is nondegenerate, then, $V = W \bigoplus W^{\perp}$. Indeed, the following are equivalent:

(1) $V = W + W^{\perp}$.

(2) W *is nondegenerate.*

(3) $V = W \bigoplus W^{\perp}$.

Moreover, we always have

$$V = \mathrm{rad}(V) \bigoplus S = V^{\perp} \bigoplus S$$

for some nondegenerate subspace S of V. ■

2.1.5 Symmetric Spaces

Let V be a vector space over a field F and let $B \in \mathrm{Hom}_F^2(V; F)$ be a reflexive bilinear form on V. If V is finite-dimensional with basis $\beta = (v_1, \ldots, v_n)$, then we say that β is an **orthogonal basis** if all the basis vectors are mutually orthogonal, i.e. $B(v_i, v_j) = 0$ whenever $i \neq j$. We call β **orthonormal** if $B(v_i, v_j) = \delta_{ij}$. We say that B is **diagonalizable** if there is a basis β for V such that $[B]_\beta = \mathrm{diag}(B(v_1, v_1), \ldots, B(v_n, v_n))$. Since such a matrix is symmetric, B is a fortiori symmetric (if γ is any other basis for V, then $[B]_\gamma = M^T_{\beta,\gamma}[B]_\beta M_{\beta,\gamma}$ is also symmetric). Thus diagonalizability implies symmetry. The converse also holds. To show this we will need the following proposition.

Proposition 4 *Let (V, B) be a reflexive finite-dimensional bilinear space over a field F and use Theorem 6 to obtain a nondegenerate subspace S of V such that $V = S \bigoplus V^{\perp}$. Then B is diagonalizable iff $B|_{S \times S}$ is diagonalizable.*

Proof: The forward direction is clear, while the other direction follows by extending any orthogonal basis $(b_1, \ldots, b_k)$ for S to an orthogonal basis for V, e.g. by adjoining any basis $(b_{k+1}, \ldots b_n)$ for $V^{\perp}$, since for $i, j \geq k + 1$ we have $B(b_i, b_j) = 0$. ■

Proposition 5 *Let (V, B) be a finite-dimensional bilinear space over a field F. Then B is diagonalizable iff it is symmetric.*

Proof: We have already remarked in the opening paragraph of this section that the diagonalizability of B implies its symmetry. Now suppose B is symmetric. By the previous proposition it suffices to prove the result for the nondegenerate subspace S complementing $V^\perp$, $V = S \bigoplus V^\perp$. Toward this end let $B' := B|_{S\times S}$, and use induction on $\dim S$. If $k = \dim S = 1$, then there exists a $b_1 \in S\backslash\{0\}$ such that $B'(b_1, b_1) \neq 0$, and for the basis $\beta = (b_1)$ we have that $[B']_\beta = [b_1]$, whcih is symmetric. Suppose, then, that the result holds for $k-1 = \dim S$ where $k-1 \geq 1$ and consider the case k. Since B' is nondegenerate, there exists a $b_1 \in S\backslash\{0\}$ such that $B'(b_1, b_1) \neq 0$. Then $\mathrm{span}(b_1)$ is a nondegenerate subspace of S, and by Theorem 6 we have $S = \mathrm{span}(b_1) \bigoplus \mathrm{span}(b_1)^\perp$. Since $\dim \mathrm{span}(b_1)^\perp = k-1$, the induction hypothesis applies to give the existence of an orthogonal basis $(b_2, \ldots, b_k)$ for $\mathrm{span}(b_1)^\perp$. Since b_1 is by definition in the orthogonal complement of $\mathrm{span}(b_1)^\perp$, we have $b_1 \perp b_j$ for all $j = 2, \ldots, k$, and so $(b_1, \ldots, b_k)$ forms an orthogonal basis for S. The linear independence of $b_1 \ldots, b_k$ is clear, since if $a_1b_1 + \cdots a_kb_k = 0$ for $a_i \in F$, applying B to both sides in the first coordinate and placing v_i in the second shows that $a_iB(b_i, b_i) = 0$, and since $B(b_i, b_i) \neq 0$, we must have $a_i = 0$. ■

Let V be a vector space over a field F, with the usual assumption that $\mathrm{ch}(F) \neq 2$. A **quadratic form** on V is a map $q : V \to F$ together with an **associated symmetric bilinear map** $B_q \in \mathrm{Hom}^2_{F,\mathrm{sym}}(V; F)$ satisfying

(1) $q(av) = a^2 q(v)$ for all $a \in F$ and all $v \in V$.

(2) $B_q(u, v) = \frac{1}{2}\big[q(u + v) - q(u) - q(v)\big]$ for all $u, v \in V$.

The pair (V, q) is called a quadratic space.

Proposition 6 *There is a one-to-one correspondence between quadratic forms on V and bilinear forms on V. Given q, we have a unique bilinear form B_q as defined in (2) above, while conversely given $B \in \mathrm{Hom}^2_{F,\mathrm{sym}}(V; F)$ we define q by $q(v) := B(q, q)$, and in this case $B = B_q$.*

Proof: Given q, we are automatically give a symmetric form B_q by (2) in the definition. So it only remains to check that given $B \in \mathrm{Hom}^2_{F,\mathrm{sym}}(V;F)$ we get a quadratic form q and $B_q = B$. Toward this end, note that if we define q by $q(v) := B(v,v)$, then, defining B_q as in (2), we have

$$B_q(v,v) = \frac{1}{2}\big[q(v+v) - 2q(v)\big] = \frac{1}{2}\big[4q(v) - 2q(v)\big] = q(v) = B(v,v)$$

for all $v \in V$, and

$$\begin{aligned} B_q(u,v) &= \frac{1}{2}\big[q(u+v) - q(u) - q(v)\big] \\ &= \frac{1}{2}\big[B(u+v,u+v) - B(u,u) - B(v,v)\big] \\ &= \frac{1}{2}\big[B(u,u) + 2B(u,v) + B(v,v) - B(u,u) - B(v,v)\big] \\ &= B(u,v) \end{aligned}$$

for all $u, v \in V$, so in fact $B = B_q$. ■

Thus, quadratic spaces (V, q) and symmetric spaces (V, B) are basically the same thing. Henceforth, we shall identify quadratic and symmetric spaces. As a side remark, we note that the identification is so common in the literature that B_q and q are often not even distinguished notationally, so that $q(u, v)$ is used instead of $B_q(u, v)$.

The terms **isotropic**, **anisotropic** and **totally isotropic**, which apply to V and B, therefore also apply to q. And q is **nondegenerate** or **nonsingular** iff B is. This terminology shows up in the build-up to Witt's extension and cancellation theorems, which we do not include here, and which are used to prove the main theorem of this section, namely Sylvester's Law of Inertia. See Szymiczek [97] for details.

Theorem 7 (Sylvester's Law of Inertia) *Let V be a finite-dimensional space over a field F of characteristic different from 2, and suppose F is an ordered field (e.g. $F = \mathbb{R}$). If B is a nondegenerate symmetric bilinear form and $\beta = (b_1, \ldots, b_n)$ and $\gamma = (c_1, \ldots, c_n)$ are two orthogonal bases for V, then the number of positive diagonal entries $b_{ii} := B(b_i, b_i)$ equals the number of positive diagonal entries $c_{ii} := B(c_i, c_i)$ in the matrix representations $[B]_\beta$ and $[B]_\gamma$, respectively,* *and similarly with the number of negative diagonal entries. These two numbers, p and q, are therefore invariants of the space (V, B), and the pair (p, q) is called the* **signature** *of the space or of the form B.* ■

Remark 4 We could enlarge the definition of signature to include degenerate symmetric forms, but then we would have a triple (p, q, r), where r would be the dimension of the space $V^\perp$ in the decomposition $V = S \bigoplus V^\perp$ of V, and where S is nondegenerate. This is sometimes taken as an alternative definition of signature. ■

Remark 5 There is a canonical diagonalization of a nondegenerate symmetric form B, namely

$$[B]_\beta = I_p \oplus -I_q = \begin{pmatrix} I_p & O \\ O & -I_q \end{pmatrix}$$

by applying an appropriate congruence to any given diagonalization of B. ■

Suppose B is nondegenerate and $\beta = (b_1, \ldots, b_p, p_{p+1}, \ldots, b_{n=p+q})$ is an orthogonal basis, and let $P = \mathrm{span}(b_1, \ldots, b_p)$ and $N = \mathrm{span}(b_{p+1}, \ldots, b_{p+q})$ be the **positive** and the **negative parts** of V, respectively, where $q(v_i) > 0$ for $i = 1, \ldots, p$ and $q(v_i) < 0$ for $i = p+1, \ldots, p+q$. A symmetric form B, and also its associated quadratic form q, is called **positive definite** if $q(v) = B(v, v) \geq 0$ for all v and $q(v) = B(v, v) = 0$ iff $v = 0$. From the decomposition of $V = P \bigoplus N$ we see immediately that B is positive definite iff $q = 0$, for if $v = \sum_{i=1}^n v_i b_i$, then

$$B(v, v) = [v]_\beta^T [B]_\beta [v]_\beta = \sum_{i=1}^n v_i^2$$

which will be negavie if we choose $v \in N$. Moreover, we clearly have $B(v, v) = 0$ iff all $v_i = 0$ iff $v = 0$ in this case.

A similar procedure applies to the case when B is **negative definite**, i.e. $q(v) = B(v, v) \leq 0$ for all $v \in V$ and $q(v) = B(v, v) = 0$ iff $v = 0$.

Notation 1 *Let us denote the space of positive definite symmetric bilinear forms*

$$\mathrm{Hom}^2_{F,\mathrm{Sym}}(V; F)^+$$ ■

2.1.6 Skew-Symmetric Spaces

Let V be a vector space over a field F and let $B \in \mathcal{L}^2(V;F)$ be a reflexive bilinear form on V. A pair of vectors $u, v \in V$ is called a **hyperbolic pair** if u and v are isotropic and $B(u,v) = 1$,

$$B(u,u) = B(v,v) = 0 \qquad \text{and} \qquad B(u,v) = 1$$

The span, $H := \operatorname{span}(u,v)$, of a hyperbolic pair is called a **hyperbolic plane**. A **hyperbolic space** is any subspace $\mathcal{H}$ of V which can be decomposed into an orthogonal direct sum of hyperbolic planes,

$$\mathcal{H} = H_1 \bigoplus H_2 \bigoplus \cdots \bigoplus H_k$$

If V itself is a hyperbolic space, and each H_i is spanned by the hyperbolic pair (u_i, v_i), then the ordererd set $(u_1, v_1, u_2, v_2, \ldots, u_k, v_k)$ is called a **hyperbolic basis** for V, or a **symplectic basis** in the case that B is skew-symmetric.

Remark 6 Any hyperbolic pair (u,v) is linearly independent, so u is not a scalar multiple of v. For if $v = au$ for some $a \in F \backslash \{0\}$, then $1 = B(u,v) = aB(u,u) = 0$, an impossibility. ■

Remark 7 Let $\beta = (u,v)$ be an ordered hyperbolic pair and let $H = \operatorname{span}(u,v)$. Then

$$[B|_{H\times H}]_\beta = \begin{pmatrix} B(u,u) & B(u,v) \\ B(v,u) & B(v,v) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \pm 1 & 0 \end{pmatrix} = \begin{cases} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, & \text{if } B \text{ is symmetric} \\ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, & \text{if } B \text{ is skew-symmetric} \end{cases} \qquad ■$$

Henceforth we suppose that (V,B) is a nondegenerate reflexive bilinear space over a field F of characteristic other than 2, so that a fortiori B is symmetric or skew symmetric (Theorem 5). If U is a subspace of V and W is any subspace of V containin U, i.e. $U \subseteq W \subseteq V$, then W is called an **extension** of U. A **nondegenerate completion** of U is a minimal nondegenerate extension of U.

Theorem 8 *Let (V,B) be a finite-dimensional nondegenerate reflexive bilinear space over a field F of characteristic other than* 2*.*

(1) *If $v \in V$ is isotropic, and $\operatorname{span}(v) \bigoplus U$ exists as a subspace of V for some subspace U of V, then $\operatorname{span}(v)$ can be extended to a hyperbolic plane $H = \operatorname{span}(v, z)$ such that $H \bigoplus U$ exists as a subspace of V.*

(2) *Let W be a nondegenerate subspace of V and suppose there are linearly independent vectors $v_1, \ldots, v_k \in V$ such that $U = \operatorname{span}(v_1, \ldots, v_k) \bigoplus W$ exists as a subspace. If $v_1, \ldots, v_k \in \operatorname{rad}(U) = U \cap U^{\perp}$, so that the v_i are isotropic, then $\operatorname{span}(v_1, \ldots, v_k)$ can be extended to a hyperbolic space $\mathcal{H}_k := H_1 \bigoplus \cdots \bigoplus H_k$ for which $(v_1, z_1, \ldots, v_k, z_k)$ is a hyperbolic basis, and moreover U has a proper nondegenerate extension*

$$\overline{U} := \mathcal{H}_k \bigoplus W = H_1 \bigoplus \cdots \bigoplus H_k \bigoplus W$$

Finally, if $(v_1, \ldots, v_k)$ for a basis for $\operatorname{rad}(U)$, then $\dim \overline{U} = \dim U + \dim \operatorname{rad}(U)$. We call $\overline{U}$ a **hyperbolic extension** *of U. If U is nondegenerate (i.e. $\operatorname{rad}(U) = \{0\}$, then we say that U is a* **hyperbolic extension of itself***.* ■

The difficult task is proving (1). Then (2) follows by an induction argument. See Roman [92], Theorem 11.10, for a full proof.

Theorem 9 *Let (V,B) be a finite-dimensional nondegenerate reflexive bilinear space over a field F of characteristic other than* 2*. Then for any subspace U the following are equivalent:*

(1) *$T = \mathcal{H} \bigoplus W$, for some hyperbolic space $\mathcal{H}$ of V, is a hyperbolic extension of U.*

(2) *T is a minimal nonsdegenerate completion of U.*

(3) *T is a nondegenerate extension of U such that $\dim \overline{U} = \dim U + \dim \operatorname{rad}(U)$.*

Proof: If $U \subseteq X \subset V$ and X is nondegenerate, then by the previous theorem there is a hyperbolic extension $K \bigoplus W$ of U such that $U \leq K \bigoplus W \leq X$, so every nondegenerate extension of U contains a hyperbolic extension of U. Since all hyperbolic extension $\overline{U}$ of U have the same dimension, $\dim \overline{U} = \dim U + \dim \operatorname{rad}(U)$, none is propertly contained in any other. ■

Remark 8 If (V, B) is a skew-symmetric space over F, with $\mathrm{ch}(F) \neq 2$, then B is not diagonalizable unless it is totally degenerate, that is $\mathrm{rad}(V) = V^\perp = V$, since we already know that diagonalizability is equivalent to symmetry, and the only symmetric and skew-symmetric form is the zero form. Thus, since we cannot hope to diagonalize a nondegenerate skew-symmetric form, the best we can hope to do is find a a symplectic basis and decompose V into hyperbolic planes. This works even if B is degenerate, but not totally degenerate. Indeed, if B is not totally degenerate, then $V^\perp \subsetneq V$, so there is a $v \in V \backslash V^\perp$, which means there is also a $w \in W$, necessarily different from v since B is alternating, such that $B(v, w) \neq 0$. By Theorem 8, there is a nontrivial nondegenerate subspace S of V such that $V = V^\perp \textcircled{\perp} S$. The claim now is that we can decompose S into hyperplanes, that is S is a hyperbolic space, and therefore $V = V^\perp \textcircled{\perp} H_1 \textcircled{\perp} \cdots \textcircled{\perp} H_k$. We prove this below. ■

Theorem 10 *Let (V, B) be a skew-symmetric space over a field F of characteristic different from 2, or else an alternating space over any F. Then*

$$\begin{aligned} V &= \mathcal{H} \textcircled{\perp} V^\perp \\ &= H_1 \textcircled{\perp} \cdots \textcircled{\perp} H_k \textcircled{\perp} V^\perp \end{aligned}$$

where the H_i are hyperplanes and $\mathcal{H} = H_1 \textcircled{\perp} \cdots \textcircled{\perp} H_k$ is a hyperbolic space. Consequently, if $\operatorname{rank} B = \dim \mathcal{H}$, *and if $\beta = (v_1, z_1, \ldots, v_k, z_k)$ is the corresponding hyperbolic basis for $\mathcal{H}$, then, extending to a basis $\tilde{\beta} = (v_1, z_1, \ldots, v_k, z_k, w_1, \ldots, w_{n-2k})$ for V we get the matrix represetntation*

$$[B]_{\tilde{\beta}} = \left(\bigoplus_{i=1}^{k} J \right) \oplus O_{n-2k} \begin{pmatrix} J & O & \cdots & O & O \\ O & J & \cdots & O & O \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ O & O & \cdots & J & O \\ O & O & \cdots & O & O_{n-2k} \end{pmatrix}$$

where $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. *If V is nondegenerate, then we just get $[B]_{\tilde{\beta}} = \bigoplus_{i=1}^{k} J$. Moreover, in the* *nondegenerate case, if we re-order the basis as* $\gamma = (v_1, \dots, v_k, z_1, \dots, z_k)$*, then*

$$[B]_\gamma = \begin{pmatrix} O_k & I_k \\ -I_k & O_k \end{pmatrix}$$

or alternatively if we order it as $\delta = (z_1, \dots, z_k, v_1, \dots, v_k)$*, then*

$$[B]_\delta = J_0 := \begin{pmatrix} O_k & -I_k \\ I_k & O_k \end{pmatrix}$$

In particular, we see that a nondegenerate skew-symmetric space V *has to be even-dimensional.*

Proof: If $V^\perp = V$, then $B = 0$ and any basis is orthogonal, whence $[B]_\beta = O_n$. If $V^\perp \subsetneq V$, then choose $v_1 \in V \backslash V^\perp$ and use the fact that $v_1 \neq V^\perp$ to find $w \in V$ such that $B(v_1, w) \neq 0$ (here $w \neq v$, because B is alternating), and let

$$z_1 := \frac{1}{B(v_1, w)}$$

Then note that $B(v_1, v_1) = B(z_1, z_1) = 0$ and $B(v_1, z_1) = 1$, so (v_1, z_1) forms a hyperbolic pair and $H_1 = \text{span}(v_1, z_n)$ a hyperbolic plane. Note that $H_1 \cap H_1^\perp = \{0\}$, since $B(v_1, z_1) = 1$ and (v_1, z_1) forms a basis for H_1, so by Proposition 3 H_1 is a nondegenerate subspace. Theorem 6 thus applies to give that $V = H_1 \bigoplus H_1^\perp$.

Now use induction on dimension: if the statement holds true for all spaces of dimension $\leq n$, then consider V with $\dim V = n + 1$. By the above argument, $V = H_1 \bigoplus H_1^\perp$, and $\dim H_1^\perp = (n + 1) - 2 = n - 1$, so if $\text{rad}\, H_1^\perp = \{0\}$, otherwise we can apply the induction hypothesis to get $H_1^\perp = H_2 \bigoplus \cdots \bigoplus H_k \bigoplus \text{rad}(H_1^\perp)$. But

$$\text{rad}(H_1^\perp) = H_1^\perp \cap H_1^{\perp\perp} = (H_1 \bigoplus H_1^\perp)^\perp = V^\perp = \text{rad}(V)$$

so in this case we have $V = H_1 \bigoplus H_2 \bigoplus \cdots \bigoplus H_k \bigoplus \text{rad}(V)$. ■

If a skew-symmetric bilinear form B is nondegenerate, we call it a **symplectic form**. Symplectic forms are conventionally denoted ω. The skew-symmetric space (V, ω) is then called a

symplectic vector space. We have just seen that a symplectic space always possesses a symplectic basis, and so is a hyperbolic space, decomposable into hyperbolic planes. By necessity V is an even-dimensional space.

Remark 9 We recall the vector space isomorphism $\mathcal{L}^2_{\text{skew}}(V;F) \cong \left(\bigwedge^2 V\right)^*$, via the universal property defining $\bigwedge^2 V$. When V is finite-dimensional, we also have

$$\text{Hom}^2_{F,\text{skew}}(V;F) \cong \bigwedge^2 V$$

If $(v_1, \ldots, v_n)$ is a basis for V, then a basis for $\bigwedge^2 V$ is $\{v_i \wedge v_j \mid 1 \leq i < j \leq n\}$, so a typical skew-symmetric 2-form is a sum of basic 2-forms,

$$B = \sum_{i<j} a_{ij} v_i \wedge v_j$$

where $a_{ij} \in F$. For example, in $\mathbb{R}^3$, in terms of the standard basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, a typical skew-symmetric 2-form is given by

$$B = a_{12}\mathbf{e}_1 \wedge \mathbf{e}_2 + a_{13}\mathbf{e}_1 \wedge \mathbf{e}_3 + a_{23}\mathbf{e}_2 \wedge \mathbf{e}_3$$

where $a_{ij} \in \mathbb{R}$. The effect of, say, $\mathbf{e}_1 \wedge \mathbf{e}_2$ on a pair of vectors $(\mathbf{u}, \mathbf{v})$ is

$$\begin{aligned}
\mathbf{e}_1 \wedge \mathbf{e}_2(\mathbf{u}, \mathbf{v}) &= \mathbf{e}^1(\mathbf{u})\mathbf{e}^2(\mathbf{v}) - \mathbf{e}^1(\mathbf{v})\mathbf{e}^2(\mathbf{u}) \\
&= u_1 v_2 - u_2 v_1 \\
&= \det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix} \\
&= \|(u_1, u_2, 0) \times (v_1, v_2, 0)\|
\end{aligned}$$

which is the area of the parallelogram spanned by the projections of $\mathbf{u}$ and $\mathbf{v}$ onto the xy-plane. Similarly,

$$\begin{aligned}
\mathbf{e}_1 \wedge \mathbf{e}_3(\mathbf{u}, \mathbf{v}) &= u_1 v_3 - u_3 v_1 = -\|((u_1, 0, u_3) \times (v_1, 0, v_3)\| \\
\mathbf{e}_1 \wedge \mathbf{e}_2(\mathbf{u}, \mathbf{v}) &= u_2 v_3 - u_3 v_2 = \|((0, u_2, u_3) \times (0, v_2, v_3)\|
\end{aligned}$$

■

2.1.7 Morphisms of Bilinear Spaces

Let (V, B_V) and (W, B_W) be bilinear vector spaces over the same field F and let $T \in \mathrm{Hom}_F(V, W)$ be a linear map. The **pullback** of the form B_W by T to V is the bilinear form

$$T^*B_W := B_W \circ (T \times T), \qquad \text{i.e.} \qquad T^*B_W(u, v) := B_W(T(u), T(v)),\ \forall u, v \in V$$

We say T **preserves the form** B_V if we have $T^*B_W = B_V$, that is if for all $u, v \in V$ we have $B_W(T(u), T(v)) = B_V(u, v)$, and we call such a form-preserving map an **morphism of bilinear spaces**. It is the structure-preserving map needed to make the collection of bilinear spaces into a *category*, $\mathcal{B}$. It's objects are bilinear spaces and its morphisms are form-preserving F-linear maps,

$$\mathrm{Ob}(\mathcal{B}) = \{(V, B) \mid V \in \mathrm{Ob}(F\text{-Vect}),\ B \in \mathrm{Hom}_F^2(V; F)\}$$
$$\mathrm{Hom}_{\mathcal{B}}\big((V, B_V),\ (W, B_W)\big) = \{T \in \mathrm{Hom}_F(V, W) \mid T^*B_W = B_V\}$$

An **isometry** of bilinear spaces (V, B_V) and (W, B_W) is an *invertible* morphism $T \in \mathrm{Hom}_{\mathcal{B}}\big((V, B_V), (W, B_W)\big)$. We denote the set of isometries

$$\mathrm{GL}\big((V, B_V),\ (W, B_W)\big)$$

The isometries of a single bilinear space (V, B) form a group under composition, called the **group of isometries**, and denoted

$$\mathrm{Isom}(V, B) \quad \text{or} \quad \mathrm{GL}(V, B) := \{T \in \mathrm{GL}(V) \mid T^*B = B\}$$

There are two broad cases we want to consider, the orthogonal group and the symplectic group. We consider only *finite-dimensional* vector spaces here.

(1) If B is symmetric and nondegenerate, with associated quadratic form q, and V is finite-dimensional, then the space $(V, B) = (V, q)$ is called a **quadratic vector space**. If B (and so q) is also positive definite, then $(V, B) = (V, q)$ is called an **inner product space** or a **Euclidean space**. The group of isometries of a quadratic space is called the **orthogonal group** and is more commonly denoted

$$O(V, B) := \mathrm{GL}(V, B) = \{T \in \mathrm{GL}(V) \mid T^*B = B\} \quad \text{or} \quad O(V, q) = \{T \in \mathrm{GL}(V) \mid T^*q = q\}$$

depending on whether we prefer to use B or q. Of course, $T^*q = q$ iff $T^*B = B$ (see the next theorem), so the two groups are the same. Its elements are called **orthogonal operators**.

(2) An isometry of symplectic spaces (V, ω_V) and (W, ω_W) is called a **linear symplectomorphism**, and an isometry $\omega \in \mathrm{Isom}(V, \omega)$ is also sometimes called a **symplectic isometry**. The group of symplectomorphisms of (V, ω) is called the **symplectic group**, and is denoted

$$\mathrm{Sp}(V, \omega) \text{ or } \mathrm{Sp}(V) := \{T \in \mathrm{GL}(V) \mid T^*\omega = \omega\}$$

and its elements are called **symplectic operators** or symplectic matrices when considering their matrix representatives.

Theorem 11 *Let (V, B_V) and (W, B_W) be reflexive bilinear vector spaces over the same field F, with $\mathrm{ch}(F) \neq 2$, and let $T \in \mathrm{Hom}_F(V, W)$.*

(1) *Given a basis $\beta = (b_1, \ldots, b_n)$ for V, T is an isometry iff T is bijective and preserves the form on β, i.e. $B_W(T(b_i), T(b_j)) = B_V(b_i, b_j)$ for all $b_i, b_j \in \beta$.*

(2) *If (V, B_V) and (W, B_W) are symmetric, then T is an isometry iff both T is bijective and $q_W(T(v)) = q_V(v)$ for all $v \in V$.*

(3) *If T is an isometry and $V = E \bigoplus E^\perp$ and $W = F \bigoplus F^\perp$, then $T(E) = F$ and $T(E^\perp) = F^\perp$.*

Proof: (1) If T is an isometry, then T is invertible, and so bijective. Moreover $B_W(T(u), T(v)) = B_V(u, v)$ for all $u, v \in V$, so in particular this is true for basis vectors. Conversely, if $B_W(T(b_i), T(b_j)) = B_V(b_i, b_j)$ for all $b_i, b_j \in \beta$, then for any $u = \sum_{i=1}^n s_i b_i$ and $v = \sum_{i=1}^n t_i b_i$ in V we have $B_W(T(u), T(v)) = \sum_{ij} s_i t_j B_W(b_i, b_j) = \sum_{ij} s_i t_j B_V(b_i, b_j) = B_V(u, v)$. Since T is bijective, it is invertible with linear inverse (from basic linear algebra), and so T is an isometry.

(2) The forward direction is clear because $q_W(T(v)) = B_W(T(v), T(v)) = B_V(v, v) = q_V(v)$, so suppose T is bijective and $q_W(T(v)) = q_V(v)$ for all $v \in V$. Then expanding and using the

symmetry of B_W and B_V we have for all $u, v \in V$ that

$$\begin{aligned} q_W(T(u)) &+ 2B_W(T(u), T(v)) + q_W(T(v)) \\ &= B_W(T(u), T(u)) + 2B_W(T(u), T(v)) + B_W(T(v), T(v)) \\ &= B_W(T(u+v), T(u+v)) \\ &= B_V((u+v), (u+v)) \\ &= B_V(u,u) + 2B_V(u,v) + B_V(v,v) \\ &= q_V(u) + 2B_V(u,v) + q_V(v) \end{aligned}$$

Since $q_W(T(u)) = q_V(u)$ and $q_W(T(v)) = q_V(v)$, we can cacle those and get $B_W(T(u), T(v)) = B_V(u, v)$.

(3) Suppose $V = E \bigoplus E^{\perp}$ and T is an isometry. If $T(E) = F$ and $W = F \bigoplus F^{\perp}$, then for all $u \in E$ and $v \in E^{\perp}$ we have $B_W(T(u), T(v)) = B_V(u, v) = 0$ and conversely $B_W(T(u), T(v)) = B_V(u, v) = 0$ implies $v \in V^{\perp}$ for all $u \in V$, so that $T(E^{\perp}) = T(E)^{\perp} = F^{\perp}$. ■

2.2 Grassmannians

Let V be an n-dimensional vector space over a field F, which we will generally consider to be $\mathbb{R}$ or $\mathbb{C}$ in these notes.

Definition 1 Nowl let k be an integer, $0 \leq k \leq n = \dim V$, and denote by

$$\mathrm{G}(k, V) \quad \text{and} \quad \mathrm{G}(k, n) := \mathrm{G}(k, F^n)$$

the set of all k-dimensional subspaces of V. We call $\mathrm{G}(k, V)$ the **Grassmannian** or **Grassmann manifold** of k-dimensional subspaces of V. When $k = 1$, this is the projective space $P(V) = \mathrm{G}(1, V)$. ■

2.2.1 Topological Manifold Structure of Grassmanians

We show in this section that for any real or complex n-dimensional vector space V and any $0 \leq k \leq n$ the Grassmannian $\mathrm{G}(k, V)$ is a topological manifold of dimension $n(n-k)$. In fact, $\mathrm{G}(k, V)$ is a smooth manifold, which we will demonstrate below, but the constructions in this section have their own virtues. They allow us to see clearly the relationship between k-frames, k-dimensional subspaces, group actions, and the geometric algebra underlying $\mathrm{G}(k, V)$. We begin by introducing the concept of a Stiefel manifold $V_k(V)$ of k-frames for V. The topology of Stiefel manifolds is easy to understand and can be used to construct the topology for $\mathrm{G}(k, V)$, as a quotient topology.

Definition 2 Let V be an n-dimensional vector space over a field F (which we take to be $\mathbb{R}$ or $\mathbb{C}$ in these notes). Define the **Stiefel manifold** of k-frames for V,

$$V_k(V) := \{(v_1, \ldots, v_k) \in \prod_{j=1}^{k} V \mid v_1, \ldots, v_n \ \text{ are linearly independent}\} \tag{2.8}$$

If V is equipped with an inner product g, we can define the subset of *orthonormal* k-frames,

$$V_k^0(V) := \{(v_1, \ldots, v_k) \in V_k(V) \mid v_1, \ldots, v_n \ \text{ are orthonormal}\} \tag{2.9}$$

which is also called a Stiefel manifold. Our notation follows that of Milnor and Stasheff [79]. ■

When working with $V = F^n$, we can characterize elements $(v_1, \dots, v_k) \in V_k(F^n)$ in terms of $n \times k$ matrices, namely those whose columns are precisely the vectors v_j: define the map

$$\begin{aligned} &V_k(F^n) \to M_{n,k}(F) \\ &(v_1, v_2, \dots, v_k) \mapsto A := \begin{pmatrix} v_1 & v_2 & \cdots & v_k \end{pmatrix} \end{aligned}$$

and note that a k-tuple of n-vectors $(v_1, \dots, v_k) \in \prod_{j=1}^k F^n$ lies in $V_k(F^n)$ iff the corresponding matrix satisfies $A^*A \in \mathrm{GL}(k, F)$, so that we have the bijection

$$V_k(F^n) \approx \{A \in M_{n,k}(F) \mid A^*A \in \mathrm{GL}(k, F)\} \tag{2.10}$$

In the real case, this is easily seen to be a type of Gram matrix, spanning a k-dimensional parallelpiped in $\mathbb{R}^n$, and $A^TA \in \mathrm{Sym}_k(\mathbb{R})$ in fact, not merely $\mathrm{GL}(k, \mathbb{R})$.

Example 1 For example, if $(\mathbf{u}, \mathbf{v})$ is a 2-frame in $\mathbb{R}^3$, then it's corresponding matrix A gives

$$A^TA = \begin{pmatrix} \mathbf{u}^T \\ \mathbf{v}^T \end{pmatrix} \begin{pmatrix} \mathbf{u} & \mathbf{v} \end{pmatrix} = \begin{pmatrix} \mathbf{u} \cdot \mathbf{u} & \mathbf{u} \cdot \mathbf{v} \\ \mathbf{v} \cdot \mathbf{u} & \mathbf{v} \cdot \mathbf{v} \end{pmatrix}$$

The square root of the absolute value of the determinant of this matrix is the area of the parallelogram spanned by $\mathbf{u}$ and $\mathbf{v}$:

$$\mathrm{Area}(P(\mathbf{u}, \mathbf{v})) = \sqrt{|\det(A^TA)|}$$

and its area is nonzero precisely when $\mathbf{u}$ and $\mathbf{v}$ are linearly independent. ■

A corollary of this result is:

$$V_k^0(\mathbb{R}^n) \approx \{A \in M_{n,k}(\mathbb{R}) \mid A^TA = I_k\} \tag{2.11}$$

and

$$V_k^0(\mathbb{C}^n) \approx \{A \in M_{n,k}(\mathbb{C}) \mid A^*A = I_k\} \tag{2.12}$$

Note that $(v_1, \dots, v_k) \in V_k(F^n)$ iff $\operatorname{rank} A = k$. Put the matrix A into standard form,

$$M_{n,k}(F) \ni A \mapsto \begin{pmatrix} B \\ C \end{pmatrix}$$

where $B \in \mathrm{GL}(k,F)$ and $C \in M_{n-k}(F)$, which involves finitely many elementary row-interchanging operations, and which can be subsumed into a single left-multiplication by $P \in \mathrm{GL}(n,F)$, then apply the determinant function to B, that is apply det to the projection of PA onto the first $k \times k$ entries. Since each of the operations, the $\mathrm{GL}(n,F)$-action, the projection (which is linear), and the determinant, is smooth, their composition is smooth, and hence continuous. Denote this composition by

$$\varphi : M_{n,k}(F) \to F$$

$$\varphi(A) := \det(\pi_{k\times k} PA)$$

Then we have that

$$V_k(F^n) = \varphi^{-1}(F^*)$$

and since F^* is open, so is $V_k(F^n)$ as a subset of $F^{kn} \cong M_{n,k}(F)$. In particular, $V_k(F^n)$ is a kn-dimensional smooth manifold.

Moreover, since $(v_1,\dots,v_k) \in V_k^0(F^n)$ iff $A^*A = I_k$, in which case $\det(A^*A) = 1$, consider the map

$$\psi : M_{n,k}(F) \to F$$

$$\psi(A) = \det(A^*A)$$

Since $A \mapsto A^*A$ and the determinant function are smooth, so is ψ, and

$$V_k^0(F^n) = \psi^{-1}(1)$$

is therefore a closed set. In fact, it is compact, since if we extend $(v_1,\dots,v_k)$ to an orthonormal basis for F^n, we have that the matrix $\tilde{A} = (A|B)$ with columns v_i is unitary/orthogonal, and so we can view A as a submatrix of $\tilde{A}$ which satisfies, for all $v \in F^k \cong F^k \times \{0_{n-k}\} \subseteq F^n$ with $\|v\| = 1$, $\|Av\| \leq 1$ (because $\|\tilde{A}v\| \leq 1$). Thus, $V_k^0(F^n)$ is also bounded (in the operator norm here, but in fact in any norm, since all norms are equivalent in finite dimensions). By the Heine-Borel theorem $V_k^0(F^n)$ is compact. It's dimension is smaller than that of $V_k(V)$. To see this, consider first the case where $V = \mathbb{R}^n$ is real. Define the smooth map

$$\psi : M_{n,k}(\mathbb{R}) \to \mathrm{Sym}_k(\mathbb{R})$$

$$\psi(A) := A^T A - I_k$$

We claim that ψ is a submersion with kernel $V_k^0(\mathbb{R}^n) = \psi^{-1}(0)$. Since $\mathrm{Sym}_k(\mathbb{R})$ is a vector space, so trivially a manifold, we can avoid using charts in computing the derivative of ψ at A, which is merely the Leibniz rule,

$$D\psi(A) \in \mathrm{Hom}_{\mathbb{R}}\big(M_{n,k}(\mathbb{R}), \mathrm{Sym}_k(\mathbb{R})\big)$$

$$D\psi(A) = A^T(\cdot) + (\cdot)^T A, \quad \text{i.e.} \quad D\psi(A)B = A^T B + B^T A$$

To see the surjectivity of $D\psi(A)$ at any $A \in V_k^0(\mathbb{R})$, let $B \in \mathrm{Sym}_k(\mathbb{R})$ and let $\frac{1}{2}AB \in M_{n,k}(\mathbb{R})$. Then, the fact that $A^T A = I_k$ implies

$$D\psi(A)\left(\frac{1}{2}AB\right) = \frac{1}{2}(A^T(AB) + B^T A^T A) = B$$

so ψ has full rank and is therefore a submersion (cf Lee [73, Proposition 4.1]), with $V_k^0(\mathbb{R}^n) = \psi^{-1}(0)$ an embedded submanifold of $M_{n,k}(\mathbb{R})$. Moreover, $\operatorname{codim} V_k^0(\mathbb{R}^n) = \dim \mathrm{Sym}_k(\mathbb{R}) = k(k+1)/2$, and therefore $\dim V_k^0(\mathbb{R}^k) = \dim M_{n,k}(\mathbb{R}) - \dim \mathrm{Sym}_k(\mathbb{R}) = nk - k(k+1)/2$ (Lee [73, Corollary 5.14]). Analogous arguments prove that for V complex, $\dim_{\mathbb{R}} V_k(\mathbb{C}^n) = 2nk$ and $\dim_{\mathbb{C}} V_k^0(\mathbb{C}^n) = 2nk - k^2$, since the image of ψ in this case is $U(k)$, which has real dimension k^2.

Under the isomorphism $V \cong F^n$ we have therefore proved the following result:

Proposition 7 *If V is a real vector space of dimension n, then the Stiefel manifold $V_k(V)$ is a smooth kn-dimensional manifold. The Stiefel submanifold $V_k^0(V)$ of orthonormal k-frames is a compact submanifold of $V_k(V)$ of dimension $nk - \frac{1}{2}k(k+1)$. If V is a complex space, then $V_k(V)$ and $V_k^0(V)$ are real manifolds of dimensions $\dim V_k(V) = 2nk$ and $\dim V_k^0(V) = 2nk - k^2$.* ■

We delay calculating the dimensions of $V_k(V)$ and $V_k^0(V)$ until the Section 2.2.3 below, where we consider these spaces as homogeneous spaces under the left-multiplication action of $\mathrm{GL}(V)$. For then we can use the Lie group orbit-stabilizer theorem to get an explicit diffeomorphism onto a quotient Lie group whose dimensions are known.

Let us, following Milnor and Stasheff [79], define the map

$$q : V_k(V) \to \mathrm{G}(k,V) \tag{2.13}$$

$$q(v_1,\dots,v_k) := \mathrm{span}_F(v_1,\dots,v_k) \tag{2.14}$$

which is clearly surjective, and let $q_0 = q|_{V_k^0(V)}$. Let us then use q to put the quotient topology on $\mathrm{G}(k,V)$, so that $U \subseteq \mathrm{G}(k,V)$ is open iff $q^{-1}(U) \subseteq V_k(F^n)$ is open. Notice that the Gram-Schmidt process gives us the following commutative diagram,

$$\begin{tikzcd}
V_k^0(F^n) \arrow[r, hook, "i"] \arrow[dr, "q_0"'] & V_k(F^n) \arrow[r, "G.S."] \arrow[d, "q"'] & V_k^0(F^n) \arrow[dl, "q_0"] \\
& \mathrm{G}(k,V) &
\end{tikzcd}$$

so that we may also view $\mathrm{G}(k,V)$ as the quotient under the map $q_0 : V_k^0(F^n) \to \mathrm{G}(k,V)$.

Proposition 8 $\mathrm{G}(k,V)$ *is a compact completely Hausdorff space.*

Proof: Compactness follows from the compactness of $V_k^0(V)$ and the continuity of the quotient map q_0. To show that $\mathrm{G}(k,V)$ is completely Hausdorff, meaning that any two distinct points U and U' in $\mathrm{G}(k,V)$ can be separated by a continuous function, we need to find a continuous function $f : \mathrm{G}(k,V) \to F$ such that $f(U) = 0$ and $f(U') = 1$. Toward this end, let U and U' be distince points in $\mathrm{G}(k,V)$. Endow V with a metric g (hermitian in the case $F = \mathbb{C}$), then decompose V as $V = U \bigoplus U^\perp$, in which case any v may be uniquely written as $v = u + u^\perp$, where $u \in U$ and $u^\perp \in U^\perp$. Then, the distance from U to v is given by

$$d(v,U) = \|u^\perp\| = \|v-u\| = \big(g(v,v) - \mathrm{proj}_U\, v\big)^{1/2} = \Big(g(v,v) - \sum_{j=1}^k g(v,b_j)\Big)^{1/2}$$

for any orthonormal basis $(b_1, \ldots, b_k)$ for U. To see that this function is continuous in U for any fixed $v \in V$, define

$$f_v : \mathrm{G}(k, V) \to F$$

$$f_v := d(v, U)$$

and note that the composition $f_v \circ q_0 : V_k(F^n) \to F$ is continuous, as evidenced by the formula

$$(f_v \circ q_0)(b_1, \ldots, b_k) := \Big(g(v, v) - \sum_{j=1}^{k} g(v, b_j)\Big)^{1/2}$$

which is the square root of a linear function in the b_j. Consequently, since the quotient topology induced by q_0 is characterized by the continuity condition, $f_v \circ q_0$ is continuous iff f_v is, we have that f_v is continuous.

Consider now two distinct subspaces U and U' in $\mathrm{G}(k, V)$ and choose $v \in U \backslash U'$. Then clearly $f_v(U) = 0$ and $f_v(U') > 0$. ■

Proposition 9 $\mathrm{G}(k, V)$ *is locally Euclidean, and therefore a topological manifold of dimension* $k(n-k)$.

Proof: Fix $W_0 \in \mathrm{G}(k, V)$, and consider the orthogonal projection $p : V = W_0 \bigoplus W_0^\perp \to W_0$. Then define the subset $U_{W_0} \subseteq \mathrm{G}(k, V)$ by

$$U_{W_0} := \{W \in \mathrm{G}(k, V) \mid p(W) = X_0\} = \{W \in \mathrm{G}(k, V) \mid W \cap W_0^\perp = \{0\}\}$$

We demonstrate in the next section (where these sets are treated as chart domains for the smooth structure on $\mathrm{G}(k, V)$) that this set is in bijective correspondence with the graphs of linear maps, and therefore with the maps themselves, between W_0 and $W_0^\perp$, that is $U_{W_0} \approx \mathrm{Hom}_F(W_0, W_0^\perp) \cong F^{k(n-k)}$, and that this bijection is moeover a homeomorphism (see the proof of Proposition 11 below). ■

Proposition 10 *The map* $\perp\colon \mathrm{G}(k,V) \to \mathrm{G}(n-k,V)$*,* $U \mapsto U^\perp$*, with respect to a given metric* g *on* V*, is a homeomorphism.*

Proof: Fix $W_0 \in \mathrm{G}(k,V)$, so that $V = W_0 \bigoplus W_0^\perp$, and fix a basis $(u_1,\ldots,u_{n-k}) \in V_{n-k}(V)$ for $W_0^\perp$. Then define the function $f : q^{-1}(U) \to V_k(V)$ as follows: for each $(v_1,\ldots,v_k) \in q^{-1}(U)$ apply the Gram-Schmidt process to $(v_1,\ldots,v_k,u_1,\ldots,u_{n-k})$ to obtain the orthonormal n-frame $(v_1',\ldots,v_n') \in V_n(V)$ for V. Then define f to be

$$f(v_1,\ldots,v_k) := (v_{k+1}',\ldots,v_n')$$

Now, the diagram

$$\begin{tikzcd} q^{-1}(U) \arrow[r, "f"] \arrow[d, "q"'] & V_k(V) \arrow[d, "q"] \\ U \arrow[r, "\perp"'] & G(k,V) \end{tikzcd}$$

is commutative. Since f is continuous (being the composition of the Gram-Schmidt process with the projection onto $W_0^\perp$), $q \circ f$ is continuous, which means $\perp \circ q$ is continuous, and therefore $\perp$ is continuous. ■

2.2.2 Smooth Manifold Structure of Grassmanians

We show in this section that for any real n-dimensional vector space V and any $0 \leq k \leq n$ the Grassmannian $\mathrm{G}(k,V)$ is a smooth manifold of dimension $n(n-k)$.

We begin by constructing charts for $\mathrm{G}(k,V)$. Toward this end, we introduce the concept of transversality, which will be needed here, and which will also be useful later for manifolds.

Definition 3 We say that two subspaces U_0 and U_1 of V are **transversal** if their sum equals all of V:

$$U_0 + U_1 = V \tag{2.15}$$

We do not require that the sum be direct, i.e. that $U_0 \cap U_1 = \{0\}$. However, the identity

$$\dim(U_0 + U_1) + \dim(U_0 \cap U_1) = \dim(U_0) + \dim(U_1)$$

shows that if $\dim(U_0) = \dim(U_1)$ and $U_0 + U_1 = V$, then $\dim(U_0 \cap U_1) = 0$ and $\dim V = 2 \dim U_i$, so $U_0 \cap U_1 = \{0\}$ and the sum is direct. We extend the notion of transversality to linear maps $T_0 \in \mathrm{Hom}(U_0, V)$ and $T_1 \in \mathrm{Hom}(U_1, V)$, saying that T_0 is **transversal** to T_1 if their images are transversal subspaces. In this case we write

$$T_0 \pitchfork T_1 \quad \Longleftrightarrow \quad \mathrm{im}\, T_0 + \mathrm{im}\, T_1 = V \tag{2.16}$$

In terms of annihilators $U^0 := \{f \in V^* \mid U \subseteq \ker f\}$, the condition translates to $0 = V^0 = (\mathrm{im}\, T_0 + \mathrm{im}\, T_1)^0 = (\mathrm{im}\, T_0)^0 \cap (\mathrm{im}\, T_1)^0 = \ker T_0^* \cap \ker T_1^*$. ■

Now let $W_1 \in \mathrm{G}(k, V)$ and find a vector space complement $W_2 \in \mathrm{G}(n-k, V)$ for W_1, so that $V = W_1 \oplus W_2$. Denote the set of k-subspaces $U \in \mathrm{G}(k, V)$ intersecting trivially with the given k-subspace $W_2 \in \mathrm{G}(n-k, V)$ by

$$U_{W_2} := \{W \in \mathrm{G}(k, V) \mid W \cap W_2 = \{0\}\} \tag{2.17}$$

Let us demonstrate that U_{W_2} is in bijective correspondence with $\mathrm{Hom}_{\mathbb{R}}(W_1, W_2)$, via

$$\phi_{12} : U_{W_2} \to \mathrm{Hom}_{\mathbb{R}}(W_1, W_2) \tag{2.18}$$

$$\phi_{12}(W) := T = \pi_2 \circ \pi_1|_W^{-1} \tag{2.19}$$

where $\pi_i : V \to W_i$ is the projection onto W_i, $i = 1, 2$. Clearly $\mathrm{im}\, \phi_{12} \subseteq \mathrm{Hom}_{\mathbb{R}}(W_1, W_2)$, since $\pi_1|_W^{-1}$ is an isomorphism from W to W_1 by the transversality condition $W \cap W_1 = \{0\}$. But to see that ϕ_{W_1,W_2} is surjective, let $T \in \mathrm{Hom}_F(W_1, W_2)$, and note that its graph $\Gamma(T)$ is a k-dimensional subspace of V, $\Gamma(T) = \{v + T(v) \mid v \in W_1\} \in \mathrm{G}(k, V)$, since a basis for $\Gamma(T)$ is constructed out of any basis $\beta = (b_1, \dots, b_k)$ for W_1, by $\gamma = (b_1 + T(b_1), \dots, b_k + T(b_k))$. Moreover, $\Gamma(T) \cap W_2 = \{0\}$, since any nonzero $v \in W_1$ would make $v + T(v)$ sit outside W_2, so $\Gamma(T) \in U_{W_2}$. Letting $W = \Gamma(T)$, we then have for all $v \in W_1$ that

$$\phi_{12}(W)(v) = \pi_2 \circ \pi_1|_{\Gamma(T)}^{-1}(v) = \pi_2(v + T(v)) = T(v)$$

so that $\phi_{12}(W) = T$, and ϕ_{12} is surjective. But ϕ_{12} is also injective, for any function $T = \phi_{12}(W)$ uniquely defined on its domain W_1, which means that if $T = \phi_{12}(W) = \phi_{12}(W') = T'$, then $W = \Gamma(T) = \Gamma(T') = W'$.

In summary, we have associated to each trivially intersecting transversal pair (W_1, W_2), where $W_1 \in \mathrm{G}(k, V)$ and $W_2 \in \mathrm{G}(n-k, V)$, a bijection from the set of $W \in \mathrm{G}(k, V)$ transversal to W_2 to $\mathrm{Hom}_{\mathbb{R}}(W_1, W_2)$.

Proposition 11 *The maps* $\phi_{12} : U_{W_2} \to \mathrm{Hom}_{\mathbb{R}}(W_1, W_2) \cong \mathbb{R}^{k(n-k)}$ *together with their domains constitute a smooth atlas on* $\mathrm{G}(k, V)$*,*

$$\mathcal{A} = \Big\{ \big(U_{W_2},\ \phi_{12}\big) \mid W_1 \in \mathrm{G}(k, V),\ W_1 \oplus W_2 = V \Big\}$$

This means that

(1) *The family* $\{U_Z \mid Z \in \mathrm{G}(n-k, V)\}$ *is a cover of* $\mathrm{G}(k, V)$ *(which, by (4), is open).*

(2) *Each image* $\phi_{12}\big(U_{W_2} \cap U_{W_2'}\big)$ *is open in* $\mathrm{Hom}_{\mathbb{R}}(W_1, W_2) \cong \mathbb{R}^{k(n-k)}$.

(3) *The transition maps*

$$\phi_{12,1'2'} = \phi_{1'2'} \circ \phi_{12}^{-1} : \phi_{12}(U_{W_2} \cap U_{W_2'}) \to \phi(U_{W_2} \cap U_{W_2'})$$

are smooth.

(4) *The resulting topology on* $\mathrm{G}(k, V)$ *defined by the atlas* $\mathcal{A}$ *is Hausdorff and second countable. Consequently,* $\mathrm{G}(k, V)$ *is a smooth* $k(n-k)$*-dimensional manifold.*

Remark 10 If we consider a complex vector space V of complex dimension n, with its *complex* Grassmannians $\mathrm{G}_{\mathbb{C}}(k, V)$, the sets of complex k-dimensional subspaces of V, $0 \leq k \leq n$, then the transition maps above are in fact holomorphic, and the atlas $\mathcal{A}$ gives $\mathrm{G}_{\mathbb{C}}(k, V)$ a complex manifold structure. ■

Proof: (1) That $\mathrm{G}(k, V) = \bigcup_{W \in \mathrm{G}(k,V)} U_W$ is clear. However, this cover in fact contains a finite subcover: Let $\beta = (b_1, \ldots, b_n)$ be any basis for V, and consider all partitions of β into two subsets, one of size k and one of size $n-k$. There are $\binom{n}{k}$ such partitions, and each determines a transversal pair $W_1 \oplus W_2 = V$. Given any $W \in \mathrm{G}(k, V)$, there must be at least one choice of W_2 which intersects trivially with W, which is just the statement that W has a complement Z, and any basis for W can be extended to a basis for V in such a way that the remaining $n-k$ basis elements are in β.

(2) Let us now show (following John Lee [73, Example 1.36]) that $\phi_{12}\big(U_{W_2} \cap U_{W_2'}\big)$ is open in $\mathrm{Hom}_{\mathbb{R}}(W_1, W_2) \cong \mathbb{R}^{k(n-k)}$. Now, any T lies in this set iff $\Gamma(T) \cap W_2' = \{0\}$, by the definition of U_{W_2}. Letting $I_T \in \mathrm{Hom}_{\mathbb{R}}(W_1, V)$ be $I_T(v) = v + T(v)$, which is an isomorphism from W_1 to $\Gamma(T)$ by the remarks above, we have $\Gamma(T) = \operatorname{im} I_T$, and since $W_2' = \ker \pi_{1'}$, we see that $\Gamma(T) \cap W_2' = \{0\}$ iff $\pi_{1'} \circ I_T$ has maximal rank. Moreover, the matrix entries of any matrix representation of $\pi_{1'} \circ I_T$ depend continuously on T (since $M_{m,n}(F)$ is linearly isomorphic to $\mathrm{Hom}_F(V, W)$, $n = \dim V$, $m = \dim W$), with respect to any choice of basis (since any change of basis is just conjugation by an invertible matrix, which is a smooth operation). But the set of such maximal rank matrices is open in $M_{m,n}(\mathbb{R})$, by the continuity of the determinant. Hence, T is contained in an open neighborhood in $\mathrm{Hom}_{\mathbb{R}}(W_1, W_2) \cong \mathbb{R}^{k(n-k)}$.

(3) Nex, let us show that the transition maps $\phi_{12,1'2'}$ are smooth. If $T \in \phi_{12}(U_{W_2} \cap U_{W_2'})$ then $T' := \phi_{12,1'2'}(T) = \pi_{2'} \circ \pi_{1'}|_{\Gamma(T)}^{-1}$. Writing I_T for the isomorphism from W_1 to $\Gamma(T)$ as in (1), we can write this as

$$T' = \pi_{2'} \circ I_T \circ (\pi_{1'} \circ I_T)^{-1}$$

However, $\pi_{1'} \circ I_T = \pi_{1'}|_{W_1} + \pi_{1'}|_{W_2} \circ T$ and $\pi_{2'} \circ I_T = \pi_{2'}|_{W_1} + \pi_{2'}|_{W_2} \circ T$, and each of these depends smoothly on T, as does the inverse of the first, which is a smooth map (see for example Rudin [93, Corollaries 1 and 2, p. 353]), as can be seen by any matrix representation of the maps involved—matrix addition, multiplication, and inversion are smooth operations.

(4) The second countability condition follows from the finite subcover of $\big\{U_Z | Z \in \mathrm{G}(n-k, V)\big\}$ as explained in the proof of (1). To see the Hausdorff condition, let $W, W' \in \mathrm{G}(k, V)$ and note that they have a common complement W_2 (which follows from the basic fact that no vector space is the finite union of proper subspaces), so they are both contained in a chart domain, say U_W. Then, a

basic fact of point-set topology (as in John Lee [73, Lemma 1.35] or Jeffrey Lee [72, Proposition 1.32]) implies the result. ■

2.2.3 Homogeneous Space Structure of Grassmannians

Consider the $\mathrm{GL}(V)$-action on the Stiefel manifold of $V_k(V)$ of k-frames,

$$\begin{aligned}\mathrm{GL}(V) \times V_k(V) &\to V_k(V)\\ \big(g, (v_1, \ldots, v_k)\big) &\mapsto \big(g(v_1), \ldots, g(v_k)\big)\end{aligned}$$

or, in shorthand, $(g, \beta) \mapsto g(\beta)$. Note that the linearity of any $g \in \mathrm{GL}(V)$ implies that g commutes with the quotient map $q : V_k(V) \to \mathrm{G}(k, V)$, $q \circ g = g \circ q$, so the $\mathrm{GL}(V)$-action descends to an action on the Grassmann manifold $\mathrm{G}(k, V)$:

$$\begin{aligned}\mathrm{GL}(V) \times \mathrm{G}(k, V) &\to \mathrm{G}(k, V)\\ (g, U) &\mapsto g(U) = \operatorname{im} g|_U\end{aligned}$$

On $V_k(V)$ the $\mathrm{GL}(V)$-action is obviously transitive by the change-of-basis theorem, and therefore the action is transitive on $\mathrm{G}(k, V)$. Put an equivalence relation $\sim$ on $V_k(V)$,

$$\beta \sim \gamma \iff q(\beta) = q(\gamma)$$

and let $\pi : V_k(V) \to V_k(V)/\sim$ be the quotient map, then consider the induced map $f : \big(V_k(V)/\sim\big) \to \mathrm{G}(k, V)$ making the following diagram commute,

$$\begin{tikzcd}
V_k(V) \arrow[rr, "q"] \arrow[dr, "\pi"'] & & \mathrm{G}(k, V) \\
& V_k(V)/\sim \arrow[ur, "f"'] &
\end{tikzcd}$$

f is bijective by construction, for $f([\beta]) = f([\gamma])$ iff $q(\beta) = q(\gamma)$. The orbit of any $\beta \in V_k(V)$ under the $\mathrm{GL}(V)$-action is not necessarily $[\beta]$ (for example the frame $(e_1, e_2) \in V_2(\mathbb{R}^3)$ may be rotated about the x-axis by $\pi/2$) so $g[\beta] = [\beta]$ iff we restrict the action to the stabilizer subgroup

$\mathrm{GL}(V)_{f([\beta])}$, for by the commutativity of q with all $g \in \mathrm{GL}(V)$, we will have $\gamma \in \pi^{-1}(\pi(\beta))$ iff $q(\gamma) = q(\beta)$, and so $g \in G$ satisfies $\gamma = g(\beta)$ iff $q(\gamma) = q(g(\beta)) = g(q(\beta)) = q(\beta)$ iff $g \in \mathrm{GL}(V)_{f([\beta])}$. Thus, there is a bijective correspondence between equivalence classes in $V_k(V)/\sim$ and stabilizers of their image under f,

$$[\beta] \longleftrightarrow \mathrm{GL}(V)_{f([\beta])} = \mathrm{GL}(V)_{q(\beta)}$$

Moreover, $\mathrm{GL}(V)_{q(\beta)} \cdot \beta = [\beta]$ is an embedded submanifold of $V_k(V)$, whose dimension we compute below. By the Lie group orbit-stabilizer theorem we have the diffeomorphism

$$[\beta] = \mathrm{GL}(V)_{q(\beta)} \cdot \beta = \mathrm{GL}(V)_{q(\beta)} / \mathrm{GL}(V)_{q(\beta)_\beta}$$

Now consider the action of the full group $\mathrm{GL}(V)$ on $\mathrm{G}(k,V)$. The orbit-stablizer theorem for Lie groups gives us a diffeomorphism

$$\mathrm{G}(k,V) = GL(V) \cdot q(\beta) \approx \mathrm{GL}(V)/\mathrm{GL}(V)_{q(\beta)}$$

We mention in passing that the bijection $\mathrm{G}(k,V) \approx V_k(V)/\sim$ is a diffeomorphism, for $V_k(V)/\sim$ can be given a quotient manifold topology, though this requires some work to demonstrate. Let us characterize the stabilizer $\mathrm{GL}(V)_{q(\beta)}$, and thus get some topological information about $\mathrm{G}(k,V)$. Note that any $g \in \mathrm{GL}(V)_{q(\beta)}$ can be written in block matrix form as

$$\mathrm{GL}(V)_{q(\beta)} = \left\{ \begin{pmatrix} h & k \\ 0 & \ell \end{pmatrix} \Bigg| \mid h \in \mathrm{GL}(U),\ \ell \in \mathrm{GL}(U^\perp),\ k \in \mathrm{Hom}_F(U^\perp, U) \right\}$$

for if we write $V = U \bigoplus U^\perp$, then g must act on the U component by a sub-operator $g' \in \mathrm{GL}(U)$ and on $U^\perp$ by any $U^\perp$-invarint operator, not necessarily invertible. But g itself must remain invertible, so once we choose $h \in \mathrm{GL}(U)$ we must let the lower left-hand sub-operator be 0. Then, noting that $\det g = (\det h)\det(\ell) \neq 0$ implies $\det \ell \neq 0$, we see that $\ell \in \mathrm{GL}(U^\perp)$, and we let $k \in \mathrm{Hom}_F(U^\perp, U)$ be arbitrary. Thus, the dimension of $\mathrm{GL}(V)_{q(\beta)}$ is $k^2 + (n-k)^2 + k(n-k) = n^2 - nk + k^2$, and therefore

$$\dim \mathrm{G}(k,V) = \dim \mathrm{GL}(V) - \dim \mathrm{GL}(V)_{q(\beta)} = n^2 - (n^2 - nk + k^2) = k(n-k)$$

which agrees with our previous result. By similar reasoning, the stabilizer of $\beta \in V_k(V)$ under the $\mathrm{GL}(V)_{q(\beta)}$-action consists of all $g \in \mathrm{GL}(V)_{q(\beta)}$ which act only on extensions γ of β to bases for V. Therefore, in the block form of g we must have $h = I$ and $k = 0$, so $\dim \mathrm{GL}(V)_\beta = (n-k)^2$ and $\dim[\beta] = (k^2 + (n-k)^2 + k(n-k)) - (n-k)^2 = k^2 + k(n-k) = kn$. Note that the stabilizer $\mathrm{GL}(V)_\beta$ of β under the total group action on $V_k(V)$ must be a subgroup of $\mathrm{GL}(V)_{q(\beta)}$, so in the block form of g we must have $h = I$, and therefore $\mathrm{GL}(V)_{q(\beta)_\beta} = \mathrm{GL}(V)_\beta$.

Let us summarize these results.

Proposition 12 *For any n-dimensional real vector space V and any $0 \leq k \leq n$, the $\mathrm{GL}(V)$-action on $\mathrm{G}(k,V)$ is transitive, and the stabilizer subgroup of any $U \in \mathrm{G}(k,V)$ is an $(n^2 - nk + k^2)$-dimensional Lie subgroup, so the Grassmann manifold is a homogeneous space of dimension $\dim \mathrm{G}(k,V) = k(n-k)$ which is diffeomorphic to the quotient group by the stabilizer group,*

$$\mathrm{G}(k,V) \approx \mathrm{GL}(V)/\,\mathrm{GL}(V)_U \tag{2.20}$$

$\mathrm{GL}(V)$ also acts transitively on $V_k(V)$, and the stabilizer group $\mathrm{GL}(V)_\beta$ of any $\beta \in V_k(V)$ is a Lie subgroup of dimension $n(n-k)$, which gives the diffeomorphism

$$V_k(V) \approx \mathrm{GL}(V)/\,\mathrm{GL}(V)_\beta \tag{2.21}$$

making $V_k(V)$ a smooth manifold of dimension $\dim V_k(V) = nk$. Moreover, the stabilizer subgroup $\mathrm{GL}(V)_{q(\beta)}$ of $U = q(\beta) \in \mathrm{G}(k,V)$ acts transitively on the equivalence class $[\beta]$ of β in $V_k(V)/\sim$, and its stabilizer subgroup $\mathrm{GL}(V)_{q(\beta)_\beta}$ of β is a Lie group of dimension $(n-k)^2 + k(n-k) = n(n-k)$ equal to $\mathrm{GL}(V)_\beta$. Thus, $[\beta]$ is an embedded submanifold of $V_k(V)$ diffeomorphic to the quotient group $\mathrm{GL}(V)_{q(\beta)}/\,\mathrm{GL}(V)_\beta$,

$$[\beta] \approx \mathrm{GL}(V)_{q(\beta)}/\,\mathrm{GL}(V)_\beta \tag{2.22}$$

whose dimension is $\dim[\beta] = kn$. ■

Let us next consider the action $\mathrm{O}(V)$ on $V_k^0(V)$, which is transitive by the change-of-basis theorem (any two orthonormal n-frames are related by an orthogonal transformation, so any two

k-frames can be extended to orthonormal n-frames and then related by a corresponding orthogonal transformation). The stabilizer of any $\beta \in V_k^0(V)$ is

$$\mathrm{O}(V)_\beta = \left\{ \begin{pmatrix} I_k & 0 \\ 0 & \ell \end{pmatrix} \middle| \, | \, \ell \in \mathrm{O}(U^\perp, g|_{U^\perp}) \cong \mathrm{O}(n-k) \right\} \cong \mathrm{O}(n-k)$$

which therefore has dimension $(n-k)(n-k-1)/2$, so that

$$V_k^0(V) = \mathrm{O}(V) \cdot \beta \approx \mathrm{O}(V)/\,\mathrm{O}(V)_\beta \cong \mathrm{O}(n)/\,\mathrm{O}(n-k)$$

and therefore

$$\begin{aligned} \dim V_k^0(V) &= \dim \mathrm{O}(n) - \dim \mathrm{O}(n-k) \\ &= \frac{n(n-1)}{2} - \frac{(n-k)(n-k-1)}{2} \\ &= nk - \frac{k(k-1)}{2} \end{aligned}$$

The analogous result of $V_k^0(V)$ for V complex is achieved by exactly the same methods, which show that $V_k^0(V) \approx \mathrm{U}(n)/\,\mathrm{U}(n-k)$ and has dimension $n^2 - (n-k)^2 = 2nk - k^2$. Restricting the actions of $\mathrm{O}(n)$ and $\mathrm{U}(n)$, respectively, to $\mathrm{SO}(n)$ and $\mathrm{SU}(n)$, we also get $V_k^0(V) \approx \mathrm{SO}(n)/\,\mathrm{SO}(n-k)$ and $V_k^0(V) \approx \mathrm{SU}(n)/\,\mathrm{SU}(n-k)$.

We summarize these results in the next proposition:

Proposition 13 *The Stiefel manifold $V_k^0(V)$ of orthonormal k-frames in a real vector space V is a homogeneous space, and a smooth submanifold of $V_k(V)$ diffeomorphic to* $\mathrm{O}(n)/\,\mathrm{O}(n-k)$,

$$\begin{aligned} V_k^0(V) &\approx \mathrm{O}(n)/\,\mathrm{O}(n-k) && (2.23) \\ &\approx \mathrm{SO}(n)/\,\mathrm{SO}(n-k) && (2.24) \end{aligned}$$

and therefore $\dim V_k^0(V) = nk - \frac{k(k-1)}{2}$. *If V is complex, then we have*

$$\begin{aligned} V_k^0(V) &\approx \mathrm{U}(n)/\,\mathrm{U}(n-k) && (2.25) \\ &\approx \mathrm{SU}(n)/\,\mathrm{SU}(n-k) && (2.26) \end{aligned}$$

making it a smooth manifold of real dimension $\dim_{\mathbb{R}} V_k^0(V) = 2nk - k^2$. ■

Two frames β and γ in $V_n(\mathbb{R}^n)$ are said to have the **same orientation** if the change-of-basis matrix $g \in \mathrm{GL}(n,\mathbb{R})$ changing β to γ coordinates, $\gamma = g\beta$, has positive determinant $\deg g > 0$. This defines an equivalence relation $\sim$ on $V_n(\mathbb{R})$,

$$\beta \sim \gamma \iff \gamma = g\beta \text{ and } \det g > 0$$

The two equivalence classes in the quotient space $V_n(\mathbb{R}^n)/\sim$ define two **orientations** on $\mathbb{R}^n$, and a choice of one of these gives $\mathbb{R}^n$ an orientation, which we denote by $V_n(\mathbb{R}^n)^+$. The remaining equivalence class is denoted $V_n(\mathbb{R}^n)^-$. Recall that $\mathrm{GL}^+(n,\mathbb{R}) = \det^{-1}(0,\infty)$ is the connected component of the identity, and consists of the orientation-preserving change-of-basis matrices, while $\mathrm{GL}^-(n,\mathbb{R}) = \det^{-1}(-\infty,0)$ consists of orientation-reversing matrices.

Corollary 2 *We have the following diffeomorphisms:*

$$V_n(\mathbb{R}^n) \approx \mathrm{GL}(n,\mathbb{R}) \qquad V_n(\mathbb{R}^n)^+ \approx \mathrm{GL}^+(n,\mathbb{R}) \tag{2.27}$$

$$V_n^0(\mathbb{R}^n) \approx \mathrm{O}(n) \qquad V_n^0(\mathbb{R}^n)^+ \approx \mathrm{SO}(n) \tag{2.28}$$

$$V_n(\mathbb{C}^n) \approx \mathrm{GL}(n,\mathbb{C}) \tag{2.29}$$

$$V_n^0(\mathbb{C}^n) \approx \mathrm{U}(n) \tag{2.30}$$

Remark 11 These results can also be understood more simply, by noting, for example, that an n-frame $\beta \in V_n(\mathbb{R}^n)$ is a set of n-linearly independent vectors, $\beta = (\mathbf{v}_1, \dots, \mathbf{v}_n)$, which can be arranged into columns of an $n \times n$ matrix $g = (\mathbf{v}_1 \ \cdots \ \mathbf{v}_n)$. This matrix is the change-of-basis matrix from the standard basis $\rho = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ to β, for $g\mathbf{e}_i = \mathbf{v}_i$. Similarly, if the $\mathbf{v}_i$ are orthonormal, that is $\beta \in V_n^0(\mathbb{R}^n)$, then $g \in \mathrm{O}(n)$ and it is the change-of-basis matrix from ρ to β. In any case, the identifications of $V_n(F^n)$ with $\mathrm{GL}(n,F)$ and $V_n^0(F^n)$ with $\mathrm{O}(n)$ or $\mathrm{U}(n)$, for $F = \mathbb{R}$ or $\mathbb{C}$, respectively, are the reason for identifying frame bundles $F(E) \to M$ of vector bundles $E \to M$ with principal $\mathrm{GL}(n,F)$-bundles, or their $\mathrm{O}(n)$- or $\mathrm{U}(n)$-bundle equivalents. ■

We pause here to consider also the case of **symplectic frames** for $\mathbb{R}^{2n}$, which are elements $\beta \in V_{2n}(\mathbb{R}^{2n})$ giving matrix representation of ω_0 the standard complex structure, $[\omega_0]_\beta = J_0$

(this is equivalent to requiring $\beta = (\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w}_1, \ldots, \mathbf{w}_n)$ to satisfy the relations $\omega_0(\mathbf{v}_i, \mathbf{v}_j) = \omega_0(\mathbf{w}_i, \mathbf{w}_j) = 0$, $\omega_0(\mathbf{v}_i, \mathbf{w}_j) = \delta_{ij}$, see (2.35) below). Consider two such bases, $\beta, \gamma \in V_{2n}(\mathbb{R}^{2n})$ and let $g \in \mathrm{GL}(2n, \mathbb{R})$ be the change-of-basis matrix taking β to γ, that is $\gamma = g\beta$. Since g acts by

$$J_0 = [\omega_0]_\gamma = g^T [\omega_0]_\beta g = g^T J_0 g$$

according to Theorem 2, we have that $g \in \mathrm{Sp}(n, \mathbb{R})$, and conversely, any $g \in \mathrm{Sp}(n, \mathbb{R})$ takes symplectic frames to symplectic frames (Proposition 19 below). Let

$$V_{2n}^{\mathrm{Sp}}(\mathbb{R}^{2n}) := \{\beta \in V_{2n}(\mathbb{R}) \mid [\omega_0]_\beta = J_0\}$$

be the set of all symplectic frames for $\mathbb{R}^{2n}$. Then we can identify this set with the symplectic group $\mathrm{Sp}(n, \mathbb{R})$ by an analog of the above procedure for showing $V_n^0(\mathbb{R}^n) \approx \mathrm{O}(n)$, namely, let $\mathrm{Sp}(n, \mathbb{R})$ act on $V_{2n}^{\mathrm{Sp}}(\mathbb{R}^n)$ by $g \cdot \beta := g(\beta)$. Since the $\mathrm{Sp}(n, \mathbb{R})$-action is free and transitive we have a bijection $V_{2n}^{\mathrm{Sp}}(\mathbb{R}^{2n}) \approx \mathrm{Sp}(n, \mathbb{R})$, by the orbit stabilizer theorem, which can be upgraded to a diffeomorphism once we view $V_{2n}^{\mathrm{Sp}}(\mathbb{R}^{2n})$ as a manifold. This can be achieved as follows: but of course, as in the previous remark on the identification of $\mathrm{GL}(n, \mathbb{R})$ with $V_n(\mathbb{R}^n)$, we can view any $\beta \in V_{2n}^{\mathrm{Sp}}(\mathbb{R}^{2n})$ as a matrix g with columns the vectors in β. But this matrix is the change-of-basis matrix from the standard symplectic basis $\rho = (p_1, \ldots, p_n, q_1, \ldots, q_n)$ to β, since $gp_i = \mathbf{v}_i$ and $gq_j = \mathbf{w}_j$. Thus, the identification is actually just a change of viewpoint, viewing bases as matrices. The manifold structure is then immediate (Proposition 26).

We summarize these results below:

Proposition 14 *The space of symplectic frames $V_{2n}^{\mathrm{Sp}}(\mathbb{R}^{2n})$ may be identified diffeomorphically with the symplectic group $\mathrm{Sp}(n, \mathbb{R})$,*

$$V_{2n}^{\mathrm{Sp}}(\mathbb{R}^{2n}) \approx \mathrm{Sp}(n, \mathbb{R}) \tag{2.31}$$

■

2.3 Symplectic Vector Spaces

In this section we consider the special case of bilinear spaces (V, ω) over the field $\mathbb{R}$ where the bilinear form ω is skew-symmetric and nondegenerate. Necessarily, the (real) dimension of such a space is even, and it is perhaps usurprising that there is an intimate relationship between such spaces and complex vector spaces of complex dimension half that of the real dimension. Indeed, the complex vector spaces are V itself, equipped with different complex structures. There is a close relationship between the different choices of complex structures for V and the collection of metrics (symmetric, positive definite bilinear forms) on V. Indeed, there is an explicit correspondence—not bijective however—between metrics $\mathcal{M}(V)$ on V and a certain subclass of complex structures $\mathcal{J}(V, \omega)$ on V.

2.3.1 Complexifications and Complex Structures

Let V be a real (finite-dimensional) vector space. There are two main ways to turn V into a complex vector space, by complexifying and by use of so-called complex structures. In the first case we take $\mathbb{C}$ and consider only its uderlying real vector space structure, so that we can tensor it with V over $\mathbb{R}$, thereby getting what is called the **complexification** of V,

$$V^{\mathbb{C}} := V \otimes_{\mathbb{R}} \mathbb{C}$$

Complex scalar multiplication on $V^{\mathbb{C}}$ is given on simple tensors by $\alpha(v \otimes z) := v \otimes (\alpha z)$, where $v \in V$ and $\alpha, z \in \mathbb{C}$, and extended by linearity to all of $V^{\mathbb{C}}$.

Remark 12 Note that $V^{\mathbb{C}}$ is also a real vector space, since the tensor product is over $\mathbb{R}$, so if $a \in \mathbb{R}$, then we still have

$$a(v \otimes z) = (av) \otimes z = v \otimes (az)$$

but if $a \in \mathbb{C}$, then we only have

$$a(v \otimes z) = v \otimes (az) \qquad \blacksquare$$

The second way makes use of of a **complex structure** on V, a real-linear automorphism $J \in \mathrm{GL}_{\mathbb{R}}(V)$ which squares to the negative of the identity I on V,

$$J^2 = -I$$

i.e. $J \circ J = -I$. Using J we can define complex-scalar multiplication on V by

$$zv = (a+ib)v := \big(aI + bJ\big)(v) = av + bJ(v) \tag{2.32}$$

for all $z = a+ib \in \mathbb{C}$ and $v \in V$.

Proposition 15 *The complex multiplication (2.32) makes V a complex vector space, which we denote V_J.*

Proof: (1) $(V_J, +)$ is abelian, since $(V, +)$ is.

(2) (*distributivity*) If $z = a+ib, w = c+id \in \mathbb{C}$ and $u, v \in V$, then

$$\begin{aligned}
(z+w)v &= \big((a+c)+i(b+d)\big)v \\
&= (a+c)v + (b+d)J(v) \\
&= (av + bJ(v)) + (cv + dJ(v)) \\
&= (a+ib)v + (c+id)v \\
&= zv + wv
\end{aligned}$$

and

$$\begin{aligned}
z(u+v) &= (a+ib)(u+v) \\
&= a(u+v) + bJ(u+v) \\
&= (au + bJ(u)) + (av + bJ(v)) \\
&= (a+iv)u + (a+iv)v \\
&= zu + zv
\end{aligned}$$

(3) (*compatibility*) If $\zeta = a + ib, \eta = c + id \in \mathbb{C}$ and $v \in V$, then

$$\begin{aligned}
\zeta(\eta v) &= (a+ib)\big((c+id)v\big) \\
&= (a+ib)(cv + dJ(v)) \\
&\overset{(2)}{=} (a+ib)(cv) + (a+ib)(dJ(v)) \\
&= a(cv) + bJ(cv) + a(dJ(v)) + bJ(dJ(cv)) \\
&= (ac)v + (bc)J(v) + (ad)J(v) + (bd)J^2(v) \\
&= (ac)v + (bc)J(v) + (ad)J(v) - (bd)v \\
&= (ac - bd)v + (ad + bc)J(v) \\
&= \big((a+ib)(c+id)\big)v \\
&= (\zeta\eta)v
\end{aligned}$$

(4) (*identity*) $1_{\mathbb{C}}v = (1_{\mathbb{R}} + i0)v = 1_{\mathbb{R}}v + 0J(v) = 1_{\mathbb{R}}v = v$. ■

Proposition 16 *If W is a complex finite-dimensional vector space, then we can look at the underlying real vector space and put a complex structure J on W, namely $J = L_i$, the left-multiplication-by-i map $J(v) := L_i(v) := iv$. Then the space W_J with multiplication given by (2.32) returns the original space W, i.e. $W_J = W$.*

Proof: This is simply the observation that $(a + ib)v = aI(v) + bJ(v) = av + ibv$. ■

It is an immediate corrolary of this result that V_J must have even real dimension if it is to have a complex dimension, or equivalently if it is to admit a complex structure J. We naturally want to understand the relationships between V, V_J, and $V^{\mathbb{C}}$. It is simplest to consider $V = \mathbb{R}^{2n}$ and $V_J = \mathbb{C}^n$, for it seems natural to consider the underlying real vector space $\mathbb{R}^{2n}$ of $\mathbb{C}^n$ as the direct sum $\mathbb{R}^n \oplus \mathbb{R}^n$, with the complex structure J acting on this space as mutiplication by i.

We begin with the following observation: The complexification $V^{\mathbb{C}}$ of V is functorial. On maps between real vector spaces, $T \in \mathrm{Hom}_{\mathbb{R}}(V, W)$, we have

$$T^{\mathbb{C}} \in \mathrm{Hom}_{\mathbb{C}}(V^{\mathbb{C}}, W^{\mathbb{C}})$$
$$T^{\mathbb{C}}(v \otimes z) := T(v) \otimes z$$

on simple tensors, then extending linearly. However, the complexification via J is not immediately functorial, as the next proposition shows:

Theorem 12 *Let $T \in \mathrm{End}_{\mathbb{R}}(V)$. Then $T_J := T \in \mathrm{End}_{\mathbb{C}}(V_J)$ iff $T \circ J = J \circ T$.*

Proof: Let $T \in \mathrm{End}_{\mathbb{R}}(V)$. If $T \in \mathrm{End}_{\mathbb{C}}(V_J)$, i.e. $T((a+ib)v) = (a+ib)T(v)$, then

$$aT(v) + bT(J(v)) = T(av + bJ(v)) = T((a+ib)v) = (a+ib)T(v) = aT(v) + bJ(T(v))$$

so that $T(J(v)) = J(T(v))$ for all $v \in V$. Conversely, if $T \circ J = J \circ T$, then

$$(a+ib)T(v) = aT(v) + bJ(T(v)) = aT(v) + bT(J(v)) = T(av + bJ(v)) = T((a+ib)v)$$

and $T \in \mathrm{End}_{\mathbb{C}}(V_J)$. ■

Example 2 Consider the real-linear map $T = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^2)$. Then,

$$T \circ J_0 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 4 & -3 \end{pmatrix} \neq \begin{pmatrix} -3 & -4 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = J_0 \circ T$$

Therefore, $T_J \notin \mathrm{End}_{\mathbb{C}}(V_J)$. ■

The previous results allow us to better understand the relationship between $\mathbb{C}^n$ and $\mathbb{R}^{2n}_J$. If we start with $\mathbb{C}^n$ and forget the complex structure, we are left with $\mathbb{R}^{2n}$. To regain the complex structure, we complexify via J, that is choose a complex structure $J \in \mathrm{GL}(2n, \mathbb{R})$ and obtain $\mathbb{R}^{2n}_J$. Let us show how this works in $\mathbb{R}^2_J \cong \mathbb{C}$ first, and then generalize to $\mathbb{R}^{2n}_J \cong \mathbb{C}^n$.

Example 3 (Complex Structure on $\mathbb{R}^2$) Consider $V = \mathbb{R}^2$ with $J \in \mathrm{GL}(\mathbb{R}^2) = \mathrm{GL}(2, \mathbb{R})$ given by

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Clearly $J^2 = -I$, so the complex scalar multiplication on $\mathbb{R}^2_J$ takes the form

$$(a + ib)\begin{pmatrix} x \\ y \end{pmatrix} = (aI + bJ)\begin{pmatrix} x \\ y \end{pmatrix} = a\begin{pmatrix} x \\ y \end{pmatrix} + b\begin{pmatrix} -y \\ x \end{pmatrix} = \begin{pmatrix} ax - by \\ bx + ay \end{pmatrix}$$

This is precisely the (real) matrix representation of $\mathbb{C}$ on $\mathbb{R}^2$,

$$\rho : \mathbb{C} \to \mathrm{GL}(2, \mathbb{R})$$

$$\rho(a + ib) := aI + bJ = a\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b\begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

which is clearly faithful and irreducible, and an isomorphism onto its image,

$$\mathbb{C} \cong \rho(\mathbb{C}) = \mathrm{span}_{\mathbb{R}}(I, J)$$

The only difference between $\mathbb{R}^2_J$ and the matrix representation of $\mathbb{C}$ is the latter takes the form of endomorphisms, or matrices, while the former is the associated group action of $\mathbb{C}$ on $\mathbb{R}^2$ via the representation.

In any case, this shows that $\mathbb{R}^2_J \cong \mathbb{C}$, and therefore, since $\mathbb{R}^2_J \cong \mathbb{R}^{\mathbb{C}}$, we have the known result,

$$\mathbb{C} \cong \mathrm{span}_{\mathbb{R}}(I, J) = \mathbb{R}^2_J \cong \mathbb{R}^{\mathbb{C}} = \mathbb{R} \otimes_{\mathbb{R}} \mathbb{C} \qquad \blacksquare$$

Example 4 (Complex Structure on $\mathbb{R}^{2n}$) The identification $\mathbb{C} \cong \mathbb{R}^2_J$ of the previous example can be generalized in an identical way, except for the analogy with the matrix representation of $\mathbb{C}$, since $\mathbb{C}^n$ is not a group for $n \geq 2$:

$$\mathbb{C}^n \cong \mathrm{span}_{\mathbb{R}}(I_{2n}, J) = \mathbb{R}^{2n}_J \cong (\mathbb{R}^n)^{\mathbb{C}} = \mathbb{R}^n \otimes_{\mathbb{R}} \mathbb{C}$$

for any complex structure J. $\blacksquare$

Let us consider again the identification $\mathbb{C}^n \cong \mathbb{R}^{2n}_{J_0}$, particularly the precise way in which the complex operators, or their matrix representatives, relate to their associated real operators, or their matrix representatives. At issue is the dimension: $\mathbb{C}^n$ has complex dimension n, while $\mathbb{R}^{2n}_J$ has real dimension twice that, $2n$, so matrices from $\mathbb{C}^n$ to itself are complex and of size $n \times n$, while matrices from $\mathbb{R}^{2n}_J$ to itself are real and of size $2n \times 2n$. Thus, we would like to embed $M_n(\mathbb{C})$ into $M_{2n}(\mathbb{R})$ in such a way as to preserve the algebraic structure of $\mathrm{End}_{\mathbb{C}}(\mathbb{C}^n) \cong M_n(\mathbb{C})$ yet allow us to view it as a subring of $\mathrm{End}_{\mathbb{R}}(\mathbb{R}^{2n}_J) \cong M_{2n}(\mathbb{R})$. With this in hand, we will be able to view $\mathrm{GL}(n, \mathbb{C})$ as a subgroup of $\mathrm{GL}(2n, \mathbb{R})$. As a side benefit, subject to the commutativity condition of Theorem 12, we will complete the functoriality of the complexification by J.

Proposition 17 *Define the map $f : M_n(\mathbb{C}) \to M_{2n}(\mathbb{R})$ by*

$$f(X) = f(A + iB) := \begin{pmatrix} A & -B \\ B & A \end{pmatrix}$$

where $A = \Re(X)$ and $B = \Im(X)$. Then, f is an injective real-linear ring homomorphism which satisfies $f(X^*) = f(X)^T$.

Proof: First, f is a linear transformation: if $a \in \mathbb{R}$ and $X = A + iB \in M_n(\mathbb{C})$, then

$$f(aX) = f(aA + iaB) = \begin{pmatrix} aA & -aB \\ aB & aA \end{pmatrix} = a \begin{pmatrix} A & -B \\ B & A \end{pmatrix} = af(X)$$

while if $X = A + iB$ and $Y = C + iD$ lie in $M_n(\mathbb{C})$, then

$$\begin{aligned} f(X+Y) &= f((A+iB) + (C+iD)) = f((A+C) + i(B+D)) \\ &= \begin{pmatrix} A+C & -(B+D) \\ B+D & A+C \end{pmatrix} = \begin{pmatrix} A & -B \\ B & A \end{pmatrix} + \begin{pmatrix} C & -D \\ D & C \end{pmatrix} = f(X) + f(Y) \end{aligned}$$

Secondly, f preserves matrix multiplication. To see this, note that

$$XY = (A + iB)(C + iD) = (AC - BD) + i(BC + AD)$$

so that

$$\begin{aligned} f(XY) &= f((AC-BD)+i(BC+AD)) \\ &= \begin{pmatrix} AC-BD & -(BC+AD) \\ BC+AD & AC-BD \end{pmatrix} = \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \begin{pmatrix} C & -D \\ D & C \end{pmatrix} = f(X)f(Y) \end{aligned}$$

Finally,

$$f(X^*) = f((A+iB)^*) = f(A^T - iB^T) = \begin{pmatrix} A^T & B^T \\ -B^T & A^T \end{pmatrix} = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}^T = f(X)^T$$

The injectivity of f is clear, since $\ker f = \{0\}$. ∎

Corollary 3 $\mathrm{GL}(n,\mathbb{C})$ *can be viewed as a subgroup of* $\mathrm{GL}(2n,\mathbb{R})$ *under the identification with its image under* f*. Moreover, under this identification, we have, in view of Theorem 12, that the complex general linear group is isomorphic to the stabilizer of* J_0 *in* $\mathrm{GL}(2n,\mathbb{R})$ *under the adjoint action:*

$$\mathrm{GL}(n,\mathbb{C}) \cong \mathrm{GL}(2n,\mathbb{R})_{J_0} = \{X \in \mathrm{GL}(2n,\mathbb{R}) \mid XJ_0 = J_0X\}$$

As a result, the real dimension of the Lie group $\mathrm{GL}(n,\mathbb{C})$ *is* $2n^2$*.* ∎

Remark 13 We should observe that the first method of extending a linear map $T \in \mathrm{Hom}_{\mathbb{R}}(V,W)$ to its complexification $T^{\mathbb{C}} \in \mathrm{Hom}_{\mathbb{C}}(V^{\mathbb{C}},W^{\mathbb{C}})$ works equally well for *multilinear* maps. If $T \in \mathrm{Hom}^k_{\mathbb{R}}(V_1,\dots,V_k;W)$, then

$$\begin{aligned} &T^{\mathbb{C}} \in \mathrm{Hom}^k_{\mathbb{C}}(V_1^{\mathbb{C}},\dots,V_k^{\mathbb{C}};W^{\mathbb{C}}) \\ &T^{\mathbb{C}}(v_1\otimes z_1,\dots,v_k\otimes z_k) := T(v_1,\dots,v_k)\otimes(z_1\cdots z_k) \end{aligned}$$

on simple tensors, then extending by multilinearity. Moreover, if $T \in \mathrm{Hom}^2_{\mathbb{R},\mathrm{Sym}}(V;\mathbb{R})$ is a symmetric real form, then $T^{\mathbb{C}} \in \mathrm{Hom}^2_{\mathbb{C},\mathrm{Skew}}(V^{\mathbb{C}};\mathbb{C})$ is a symmetric complex form, and similarly with skew-forms.

But there are situations where we would like to extend $T \in \mathrm{Hom}^2_{\mathbb{R}}(V;\mathbb{R})$ to a *sesquilinear* form. That is, we would like it to be conjugate-linear in the first term and $\mathbb{C}$-linear in the second term. This is accomplished in the obvious way, on simple tensors:

$$T^{\mathbb{C}_s}(\alpha(v \otimes z), \beta(w \otimes \zeta)) := \overline{\alpha}\beta T(v,w) \otimes (z\zeta)$$

for $v, w \in V$, $\alpha, \beta, z, \zeta \in \mathbb{C}$. Then we may distinguish between symmetric and skew-symmetric sesquilinear extensions as before. A symmetric sessquilinear form is called a **Hermitian** form. The symmetry of T means $T(u,v) = T(v,u)$ for all $u, v \in V$, where V is a complex vector space. Combined with the skew-symmetry we have for all $a, b \in \mathbb{C}$ that $T(au,bv) = T(bv,au)$ iff $\overline{b}aT(v,u) = \overline{a}bT(u,v)$ iff $\overline{\overline{b}aT(v,u)} = \overline{a}bT(u,v)$ iff $T(v,u) = \overline{T(u,v)}$, which is the usual definition of Hermitian. We say that a Hermitian form is *positive definite* if additionally $T(u,u) > 0$ for all $u \in V\backslash\{0\}$, and this applies equally well to any complexification of a real form. ■

The identity $J^2 = -I$ shows that the only eigenvalues of J are $\pm i$, but these are not real eigenvaues, while J is a real-linear endomorphism. Hence the need to complexify V and extend J to $V^{\mathbb{C}}$. Now, the real vector space V is naturally embedded into $V^{\mathbb{C}}$ by

$$V \hookrightarrow V^{\mathbb{C}}$$

$$v \mapsto v \otimes 1$$

and we can see that, if we define complex conjugation on $V^{\mathbb{C}}$ by

$$\overline{v \otimes z} := v \otimes \overline{z}$$

then V is the subspace of $V^{\mathbb{C}}$ which is left invariant under conjugation. When V is endowed with a complex structure J, we can extend J functorially to $V^{\mathbb{C}}$ by defining it on simple tensors as

$$J^{\mathbb{C}}(v \otimes z) := J(v) \otimes z$$

and extending linearly. Then the identity $(J^{\mathbb{C}})^2 = -I$ shows that the only eigenvalues of $J^{\mathbb{C}}$ are $\pm i$. The $\pm i$ eigenspaces of $J^{\mathbb{C}}$ are denoted (cf Huybrechts [59]),

$$V^{1,0} := \{v \in V^{\mathbb{C}} \mid J^{\mathbb{C}}(v) = iv\}$$

$$V^{0,1} := \{v \in V^{\mathbb{C}} \mid J^{\mathbb{C}}(v) = -iv\}$$

Since $V^{1,0} \cap V^{0,1} = \{0\}$, the map $V^{1,0} \oplus V^{0,1} \to V^{\mathbb{C}}$ is merely inclusion, so is injective. But it is surjective, too, for it has an inverse, via the projections

$$P_{\pm} : V^{\mathbb{C}} \to V^{1,0} \quad \text{or} \quad V^{0,1}$$

$$P_{\pm}(v) := \frac{1}{2}(I \mp iJ^{\mathbb{C}})(v)$$

namely,

$$P_+ + P_- : V^{\mathbb{C}} \to V^{1,0} \oplus V^{0,1}$$

and this shows that

$$V^{\mathbb{C}} = V^{1,0} \oplus V^{0,1}$$

Moreover, complex conjugation on $V^{\mathbb{C}}$ interchanges $V^{1,0}$ and $V^{0,1}$ in an $\mathbb{R}$-linearly isomorphic manner,

$$V^{1,0} \cong_{\mathbb{R}} V^{0,1}$$

since, writing $v = x + iy \in V^{\mathbb{C}}$, we have $\overline{v - iJ^{\mathbb{C}}(v))} = \overline{x + iy - iJ^{\mathbb{C}}(x + iy)} = x - iy + iJ^{\mathbb{C}}(x) - J^{\mathbb{C}}(y) = \overline{x - iy} + iJ^{\mathbb{C}}(\overline{x - iy}) = \overline{v} + iJ^{\mathbb{C}}(\overline{v})$. Consequently,

$$\dim_{\mathbb{R}} V = 2n$$

$$\dim_{\mathbb{C}} V^{\mathbb{C}} = 2n$$

$$\dim_{\mathbb{C}} V^{1,0} = \dim_{\mathbb{C}} V^{0,1} = n$$

We remark that $V^{\mathbb{C}}$ has two complex structures, i and J, which in general differ by a minus sign on $V^{0,1}$, though they agree on $V^{1,0}$.

Notation 2 *In what follows, we will write J for both J and $J^{\mathbb{C}}$, to avoid pedantry. The context will make clear which complex structure we are referring to.* ■

For proofs of the following statements we refer to Huybrechts [59].

(1) If V is a real vector space equipped with a complex structure J, then the dual space $V^* = \mathrm{Hom}_{\mathbb{R}}(V, \mathbb{R})$ has a natural complex structure, the pullback or transpose, J^*, defined by $J^* f := f \circ J$, and in this case

$$(V^*)^{\mathbb{C}} = (V^*)^{1,0} \oplus (V^*)^{0,1}$$

where

$$(V^*)^{1,0} = (V^{1,0})^* = \mathrm{Hom}_{\mathbb{C}}(V_J, \mathbb{C})$$
$$(V^*)^{0,1} = (V^{0,1})^*$$

(2) We also define the exterior algebra of $V^{\mathbb{C}}$ in terms of bi-degree-valued vectors,

$$\bigwedge^{\bullet} V^{\mathbb{C}} := \bigoplus_{k=0}^{d} \bigwedge^{k} V^{\mathbb{C}}$$

where $d = 2n = \dim_{\mathbb{R}} V^{\mathbb{C}}$, and

$$\bigwedge^{p,q} V := \bigwedge^{p} V^{1,0} \otimes_{\mathbb{C}} \bigwedge^{q} V^{0,1}$$

Here $\bigwedge^{p,q} V$ can be viewed as a subspace of $\bigwedge^{p+q} V^{\mathbb{C}}$, and

$$\bigwedge^{k} V = \bigotimes_{p+q=k} \bigwedge^{p,q} V$$

Therefore,

$$\bigwedge^{\bullet} V^{\mathbb{C}} := \bigoplus_{k=0}^{d} \bigoplus_{p+q=k} \bigwedge^{p,q} V$$

Complex conjugation defines a $\mathbb{C}$-anti-linear ismorphism $\bigwedge^{p,q} V \cong \bigwedge^{q,p} V$, namely

$$\overline{\bigwedge^{p,q} V} = \bigwedge^{q,p} V$$

Lastly, exterior multiplication is a bidgree $(0,0)$ map,

$$\bigwedge^{p,q} V \times \bigwedge^{r,s} V \to \bigwedge^{p+r,q+s} V$$
$$(\alpha, \beta) \mapsto \alpha \wedge \beta$$

From the above considerations we can understand the relations between V, V_J, $V^{\mathbb{C}}$, and $V_J^{\mathbb{C}} = V_J^{1,0} \oplus V_J^{0,1}$, namely, choose a $\mathbb{C}$-basis $(z_1, \ldots, z_n)$ for $V_J^{1,0}$, where $z_j = \frac{1}{2}(v_j - iJ(v_j))$ for some $v_j \in V$, and let $w_j := J(v_j)$. Then $(v_1, w_1, \ldots, v_n, w_n)$ is an $\mathbb{R}$-basis for V, and $(v_1, \ldots, v_n)$ is a $\mathbb{C}$ for V_J, and

$$V_J \cong V^{1,0}$$

(see Huybrechts [59, p. 30]).

2.3.2 Symplectic Forms

We recall some basic facts about skew-symmetric spaces (V, B) from Theorem 10 above. First, any such space has an orthogonal direct sum decomposition into hyperbolic planes and it's radical, $V = H_1 \bigoplus \cdots \bigoplus H_k \bigoplus V^\perp$. If B is nondegenerate, we call it a symplectic form and denote it by ω instead. Then $V^\perp = \{0\}$, so $V = H_1 \bigoplus \cdots \bigoplus H_k$. Moreover, (V, ω) has possesses a symplectic basis $\beta = (v_1, z_1, \ldots, v_n, z_n)$ with respect to which ω has a matrix representation

$$[\omega]_{\tilde{\beta}} = \bigoplus_{i=1}^{n} J = \begin{pmatrix} J & O & \cdots & O \\ O & J & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & J \end{pmatrix}$$

where $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. If we re-order the basis as $\gamma = (v_1, \ldots, v_n, z_1, \ldots, z_n)$, then

$$[\omega]_\gamma = \begin{pmatrix} O_n & I_n \\ -I_n & O_n \end{pmatrix} \tag{2.33}$$

while if we order it as $\delta = (z_1, \ldots, z_k, v_1, \ldots, v_k)$, then

$$[\omega]_\delta = J_0 := \begin{pmatrix} O_n & -I_n \\ I_n & O_n \end{pmatrix} \tag{2.34}$$

These representations are equivalent to ω satisfying the relations

$$\omega(v_i, v_j) = \omega(z_i, z_j) = 0 \qquad \text{and} \qquad \omega(v_i, z_j) = \delta_{ij} \tag{2.35}$$

In particular, we see that a symplectic space (V, ω) has to be *even-dimensional* (This also follows from the much stronger statement, Proposition 32 below, that every symplectic space is symplectomorphic to $(\mathbb{R}^{2n}, \omega_0)$, see the next section for the definition of $(\mathbb{R}^{2n}, \omega_0)$). The matrix $J_0 = [\omega]_\delta$ above is called the **standard** or **canonical representation** of ω, though it should be mentioned that some authors take J_0 to be $[\omega]_\gamma$.

Theorem 13 *Let V be a $2n$-dimensional real vector space. Then, a skew-symmetric bilinear form $\omega \in \mathcal{L}^2_{\text{skew}}(V;\mathbb{R})$ is nondegenerate iff the n-fold exterior power is nonzero,*

$$\omega^n = \omega \wedge \cdots \wedge \omega \neq 0$$

Proof: If ω is degenerate, then there is a $v \neq 0$ in V such that $\omega(v, \cdot) = 0 \in V^*$, so if we extend $v = v_1$ to a basis $(v_1, \ldots, v_{2n})$ for V, then $\omega^n(v_1, \ldots, v_{2n}) = 0$, whence $\omega^n = 0$. Conversely, if ω is nondegenerate, then the easiest way to see that $\omega^n \neq 0$ is by use of the symplectomorphism $\Phi : V \to \mathbb{R}^{2n}$ given in Proposition 32 below, which pulls ω back to ω_0 in $\mathbb{R}^{2n}$, $\Phi^*\omega = \omega_0$. For we know that

$$\omega_0^n = \left(\sum_{i=1}^n dq^i \wedge dp^i\right)^2 = \sum_{\sigma \in S_n} q^{\sigma(1)} \wedge p^{\sigma(1)} \wedge \cdots \wedge q^{\sigma(n)} \wedge p^{\sigma(n)}$$

and this sum is a nonzero multiple of the standard volume form $q^1 \wedge \cdots \wedge q^n \wedge p^1 \wedge \cdots \wedge p^n$, since to put any summand $q^{\sigma(1)} \wedge p^{\sigma(1)} \wedge \cdots \wedge q^{\sigma(n)} \wedge p^{\sigma(n)}$ into this form requires an odd number of moves, each contributing a minus sign. ■

2.3.3 The Symplectic Space $(\mathbb{R}^{2n}, \omega_0)$

On $\mathbb{R}^{2n}$ with the standard basis $\rho = (\mathbf{e}_1, \ldots, \mathbf{e}_{2n})$ and dual basis $\rho^* = (d\mathbf{e}_1, \ldots, d\mathbf{e}_{2n})$ we have the **standard symplectic form**

$$\omega_0 := \sum_{i=1}^n d\mathbf{e}_i \wedge d\mathbf{e}_{n+i} \in \bigwedge^2 (\mathbb{R}^{2n})^* \cong \operatorname{Hom}^2_{\mathbb{R},\text{Skew}}(\mathbb{R}^{2n};\mathbb{R}) \tag{2.36}$$

If we relabel the basis vectors to accord with common practice in physics, letting $q_1 = \mathbf{e}_1, \ldots, q_n = \mathbf{e}_n$ and $p_1 = \mathbf{e}_{n+1}, \ldots, p_n = \mathbf{e}_{2n}$, then

$$\omega_0 = \sum_{i=1}^{n} dq_i \wedge dp_i \tag{2.37}$$

In classical mechanics the q_i represent **position** coordinates of a mechanical system and the p_i represent **momentum** coordinates, with $\mathbb{R}^{2n}$ representing the **phase space** of the system, which is thought of as the (trivial) cotangent bundle $T^*\mathbb{R}^n \cong \mathbb{R}^n \times (\mathbb{R}^n)^* \cong \mathbb{R}^{2n}$ on $\mathbb{R}^n$ identified with $\mathbb{R}^{2n}$.

Viewing ω_0 as a form, rather than a 2-covector, though still keeping the notation of the abstract Grassmann 2-covector, we recall that the effect of a wedge product of 1-forms $\alpha, \beta \in (\mathbb{R}^{2n})^*$ on a pair of vectors $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^{2n}$ is a sum involving the alternator

$$\begin{aligned} &\mathrm{Alt} : \mathcal{L}(\mathbb{R}^{2n}; \mathbb{R}) \to \mathcal{L}^2_{\text{skew}}(\mathbb{R}^{2n}; \mathbb{R}) \\ &\mathrm{Alt}(\alpha \otimes \beta)(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{2!} \sum_{\sigma \in S_2} \mathrm{sgn}(\sigma)(\alpha \otimes \beta)(\mathbf{x}_{\sigma(1)}, \mathbf{x}_{\sigma(2)}) \end{aligned}$$

via

$$\begin{aligned} (\alpha \wedge \beta)(\mathbf{x}_1, \mathbf{x}_2) &= 2!\, \mathrm{Alt}(\alpha \otimes \beta)(\mathbf{x}_1, \mathbf{x}_2) \\ &:= \sum_{\sigma \in S_2} \mathrm{sgn}(\sigma)(\alpha \otimes \beta)(\mathbf{x}_{\sigma(1)}, \mathbf{x}_{\sigma(2)}) \\ &= (\alpha \otimes \beta)(\mathbf{x}_1, \mathbf{x}_2) - (\alpha \otimes \beta)(\mathbf{x}_2, \mathbf{x}_1) \\ &= \alpha(\mathbf{x}_1)\beta(\mathbf{x}_2) - \alpha(\mathbf{x}_2)\beta(\mathbf{x}_1) \end{aligned}$$

on simple tensors $\alpha \otimes \beta$. We then extended by bi-linearity. So, for example, if $\mathbf{u} = (u_1, \ldots, u_{2n})$, $\mathbf{v} = (v_1, \ldots, v_{2n}) \in \mathbb{R}^{2n}$, then $(dq^i \wedge dp^i)(\mathbf{u}, \mathbf{v}) = u_i v_{n+i} - u_{n+i} v_i$, so that

$$\omega_0(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^{n} (dq^i \wedge dp^i)(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^{n} u_i v_{n+i} - u_{n+i} v_i \tag{2.38}$$

Consequently, ω_0 satisfies the relations (2.35),

$$\omega_0(q_i, q_j) = \omega_0(p_i, p_j) = 0 \qquad \text{and} \qquad \omega_0(q_i, p_j) = \delta_{ij} \tag{2.39}$$

for all $i, j = 1, \ldots, n$, and therefore the matrix representation of the standard form ω_0 is a complex structure on $\mathbb{R}^{2n}$,

$$[\omega_0]_\rho = \begin{pmatrix} O_n & I_n \\ -I_n & O_n \end{pmatrix} \tag{2.40}$$

We follow Habermann and Habermann [54] and define the **standard complex structure** on $\mathbb{R}^{2n}$ to be instead the negative of this matrix isomorphism (equivalently, the matrix representation of ω_0 with respect to the basis (p, q) instead of (q, p)):

$$J_0 := \begin{pmatrix} O_n & -I_n \\ I_n & O_n \end{pmatrix} \tag{2.41}$$

Since $J_0^T = [\omega_0]_\rho$, we see that ω_0 is related to the standard inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^{2n}$ by J_0 as

$$\omega_0(\mathbf{u}, \mathbf{v}) = [\mathbf{u}]_\rho^T [\omega_0]_\rho [\mathbf{v}]_\rho = \mathbf{u}^T J_0^T \mathbf{v} = \langle J_0 \mathbf{u}, \mathbf{v} \rangle \tag{2.42}$$

Also, we note that

$$J_0 \in \mathrm{Sp}(\mathbb{R}^{2n}, \omega_0), \quad \text{that is} \quad J_0^* \omega_0 = \omega_0 \tag{2.43}$$

for since $J^T = -J_0 = J_0^{-1}$, i.e. J_0 is skew-adjoint, by the above identity $\omega_0 = \langle J_0 \cdot, \cdot \rangle$ we get

$$J_0^* \omega_0(\mathbf{u}, \mathbf{v}) = \omega_0(J_0 \mathbf{u}, J_0 \mathbf{v}) = \left\langle J_0^2 \mathbf{u}, J_0 \mathbf{v} \right\rangle = \langle \mathbf{u}, -J_0 \mathbf{v} \rangle = \langle J_0 \mathbf{u}, \mathbf{v} \rangle = \omega_0(\mathbf{u}, \mathbf{v})$$

for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{2n}$. An immediate consequence is

$$J_0 \in \mathrm{O}(\mathbb{R}^{2n}, \langle \cdot, \cdot \rangle) = \mathrm{O}(2n), \quad \text{that is} \quad J_0^* \langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle \tag{2.44}$$

for if $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{2n}$,

$$\langle \mathbf{u}, \mathbf{v} \rangle = \left\langle -J_0^2 \mathbf{u}, \mathbf{v} \right\rangle = \langle J_0 \mathbf{u}, J_0 \mathbf{v} \rangle$$

As a result, we have

$$\omega_0(\cdot, J_0 \cdot) = \langle J_0 \cdot, J_0 \cdot \rangle = \langle \cdot, \cdot \rangle$$

We summarize these results in the following proposition.

Proposition 18 *If we denote the standard inner product on* $\mathbb{R}^{2n}$ *by* $g_{J_0} = \langle\cdot,\cdot\rangle$, *then we have*

(1) $J_0 \in \mathrm{Sp}(\mathbb{R}^{2n},\omega_0) \cap \mathrm{O}(\mathbb{R}^{2n}, g_{J_0}) = \mathrm{Sp}(n,\mathbb{R}) \cap \mathrm{O}(2n)$

(2) $\omega_0(\cdot,\cdot) := g_{J_0}(J_0\cdot,\cdot) = \sum_{j=0}^n dq^j \wedge dp^j$

(3) $g_{J_0}(\cdot,\cdot) := \omega_0(\cdot, J_0\cdot) = \sum_{j=0}^n dq^j \wedge (dp^j \circ J_0)$

(4) *Furthermore, if we write* $\mathbf{u} = (\mathbf{u}_1,\mathbf{u}_2)$ *where* $\mathbf{u_1} = (q_1,\dots,q_n)$ *and* $\mathbf{u}_2 = (p_1,\dots,p_n)$, *and similarly* $\mathbf{v} = (\mathbf{v}_1,\mathbf{v}_2)$, *then we also get the expression*

$$\omega_0(\mathbf{u},\mathbf{v}) = \langle J_0\mathbf{u},\mathbf{v}\rangle = \langle(-\mathbf{u}_2,\mathbf{u_1}),(\mathbf{v}_1,\mathbf{v}_2)\rangle = \langle\mathbf{u}_1,\mathbf{v}_2\rangle - \langle\mathbf{u}_2,\mathbf{v_1}\rangle \qquad \blacksquare$$

Remark 14 This situation is typical for even-dimensional real vector spaces V. If V is equipped with both an inner product g and a symplectic form ω, there is a complex structure J that is compatible with ω, in the sense that $\omega(u,v) = g(J(u),v)$ for all $u,v \in V$. It is a little trickier to start with only ω and J, for even though we get a nondegenerate symmetric bilinear form g on V, we do not always have that it is positive definite. This is not a big problem, in general, for we can always just specify g first, which is always possible on a finite-dimensional space. The main result we wish to get is that any $2n$-dimensional symplectic space (V,ω) is linearly symplectomorphic to $(\mathbb{R}^{2n},\omega_0)$, and moreover the symplectomorphism $\Phi : \mathbb{R}^{2n} \to V$ can be chosen to be compatible with complex structures $J \in \mathcal{J}(V)$ and $J_0 \in \mathcal{J}(\mathbb{R}^{2n})$, in the sense that $J \circ \Phi = \Phi \circ J_0$. We prove all this in Section 2.3.5 below. $\blacksquare$

We remark that there are $J \in \mathrm{End}(\mathbb{R}^{2n})$ which are not complex structures but which nevertheless define a symplectic form via $\omega := g_{J_0} \circ (J \times I_{2n})$. The following example of such a J is from Guillemin and Sternberg [47] and de Gosson [22].

Example 5 Let $B \in \mathrm{Skew}_n(\mathbb{R})$ and $\pi_1 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ the projection onto the first component, and define the bilinear form

$$\omega_B : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \to \mathbb{R}$$

$$\omega_B(\mathbf{u},\mathbf{v}) := \omega_0(\mathbf{u},\mathbf{v}) - \langle B\pi_1(\mathbf{u}),\pi_1(\mathbf{v})\rangle$$

Then ω_B is clearly bilinear, while skew-symmetry follows from that of B:

$$\langle B\pi_1(\mathbf{v}), \pi_1(\mathbf{u})\rangle = -\langle \pi_1(\mathbf{v}), B\pi_1(\mathbf{u})\rangle = -\langle B\pi_1(\mathbf{u}), \pi_1(\mathbf{v})\rangle$$

Note that if we set

$$J_B := \begin{pmatrix} -B & -I_n \\ I_n & 0 \end{pmatrix} = -B \oplus 0 + J_0$$

then $J_B^2 \neq -I_{2n}$ unless $B = 0$, but nevertheless

$$g_{J_0}(J_B\cdot,\cdot) = g_{J_0}(-(B\oplus 0)\cdot,\cdot) + g_{J_0}(J_0\cdot,\cdot) = -\langle B\circ\pi_1\cdot,\cdot\rangle + \omega_0(\cdot,\cdot) = \omega_B$$

Thus, J_B gives a symplectic form $\omega_B = g_{J_0}\circ(J\times I_{2n})$ though $J_B \in \mathrm{End}(\mathbb{R}^{2n})$ is not a complex structure. The term $\langle B\circ\pi_1\cdot,\cdot\rangle$ is called the *magnetic term*, and the form ω_B appears in electromagnetic theory. ■

2.3.4 The Symplectic Group $\mathrm{Sp}(n,\mathbb{R})$ and the Symplectic Lie Algebra $\mathfrak{sp}(n,\mathbb{R})$

An immediate corollary of the last proposition, 18, is a characterization of symplectic matrices, that is elements of $\mathrm{Sp}(n,\mathbb{R}) \cong \mathrm{Sp}(\mathbb{R}^{2n},\omega_0)$. Recall from Section 2.1.7 that the group of isometries $\mathrm{Isom}(V,B) := \{T \in \mathrm{GL}(V) \mid T^*B = B\}$ of a bilinear space (V,B) is the set of all invertible linear operators which preserve the form. In the case that B is a symplectic form we call the group the symplectic group and denote it $\mathrm{Sp}(V,B)$. In the special case of $(\mathbb{R}^{2n},\omega_0)$, where we identify the group $\mathrm{GL}(\mathbb{R}^{2n})$ with the matrix group $\mathrm{GL}(2n,\mathbb{R})$, we have the **symplectic matrix group**

$$\mathrm{Sp}(n,\mathbb{R}) := \mathrm{Sp}(\mathbb{R}^{2n},\omega_0)$$

The elements of $\mathrm{Sp}(n,\mathbb{R})$ are called **symplectic matrices**. Fix $J_0 := \begin{pmatrix} O_n & -I_n \\ I_n & O_n \end{pmatrix}$.

Proposition 19 *Symplectic matrices are characterized as follows:*

$$\begin{aligned} \mathrm{Sp}(n,\mathbb{R}) &= \{A \in \mathrm{GL}(2n,\mathbb{R}) \mid A^*\omega_0 = \omega_0\} && (2.45)\\ &= \{A \in \mathrm{GL}(2n,\mathbb{R}) \mid A^TJ_0A = J_0\} && (2.46)\\ &= \{A \in \mathrm{GL}(2n,\mathbb{R}) \mid AJ_0A^T = J_0\} && (2.47) \end{aligned}$$

Proof: From equation (2.42), or equivalently Proposition 18 part (2), we have that $\omega_0(\mathbf{u}, \mathbf{v}) = g_{J_0}(J_0\mathbf{u}, \mathbf{v}) = \langle J_0\mathbf{u}, \mathbf{v}\rangle$, so that $A^*\omega_0 = \omega_0$, that is $\omega_0(A\mathbf{u}, A\mathbf{v}) = \omega_0(\mathbf{u}, \mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{2n}$, iff $\langle J_0A\mathbf{u}, A\mathbf{v}\rangle = \langle J_0\mathbf{u}, \mathbf{v}\rangle$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{2n}$ iff $(J_0A\mathbf{u})^T(A\mathbf{v}) = (J_0\mathbf{u})^T\mathbf{v}$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{2n}$ iff $\mathbf{u}^TA^TJ_0^TA\mathbf{v} = \mathbf{u}^TJ_0^T\mathbf{v}$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{2n}$ iff $A^TJ_0A = J_0$, which proves the first equality. For the second equality, note that $J_0^2 = -I$ and $J_0^T = -J_0$, which implies the following set of equivalences: $A^TJ_0A = J_0$ iff $A^TJ_0AJ_0 = -I = -I^T = (A^TJ_0AJ_0)^T = J_0^TA^TJ_0^TA = J_0A^TJ_0A$ iff $AJ_0A^TJ_0A = -A$ iff $AJ_0A^T = -A(J_0A)^{-1} = -AA^{-1}J_0^{-1} = J_0$. ■

2.3.4.1 Description of the Lie Algebra $\mathfrak{sp}(n, \mathbb{R})$

Proposition 20 *The symplectic Lie algebra of* $\mathbb{R}^{2n}$ *is characterized as follows:*

$$
\begin{aligned}
\mathfrak{sp}(n,\mathbb{R}) &= T_I\,\mathrm{Sp}(n,\mathbb{R}) && (2.48)\\
&= \{A \in M_{2n}(\mathbb{R}) \mid \omega_0(A\mathbf{u},\mathbf{v}) + \omega_0(\mathbf{u},A\mathbf{v}) = 0,\ \forall \mathbf{u},\mathbf{v} \in \mathbb{R}^{2n}\} && (2.49)\\
&= \{A \in M_{2n}(\mathbb{R}) \mid \langle J_0A\mathbf{u},\mathbf{v}\rangle + \langle J_0\mathbf{u},A\mathbf{v}\rangle = 0,\ \forall \mathbf{u},\mathbf{v} \in \mathbb{R}^{2n}\} && (2.50)\\
&= \{A \in M_{2n}(\mathbb{R}) \mid J_0A + A^TJ_0 = 0\} && (2.51)\\
&= \{\begin{pmatrix} B & C \\ D & -B^T \end{pmatrix} \in M_{2n}(\mathbb{R}) \mid B \in M_n(\mathbb{R}),\ C, D \in \mathrm{Sym}_n(\mathbb{R})\} && (2.52)
\end{aligned}
$$

Proof: If $A(t)$ is a path in $\mathrm{Sp}(n, \mathbb{R})$ passing through I_{2n} at $t = 0$, then $A(t)^TJ_0A(t) = J_0$ for all t in a neighborhood of 0, so differentiating with respect to t at $t = 0$ gives

$$
\frac{d}{dt}\bigg|_{t=0}\big(A(t)^TJ_0A(t)\big) = \left(\frac{d}{dt}\bigg|_{t=0} A(t)^T\right)J_0A(0) + A(0)J_0\left(\frac{d}{dt}\bigg|_{t=0} A(t)\right) = BJ_0 + J_0B
$$

where $B = \big(\frac{d}{dt}\big|_{t=0}A(t)\big)$, and this equals 0, since the right-hand-side is constant at J_0. Since we can specify such a bath by its derivative at $t = 0$, $\frac{d}{dt}\big|_{t=0}A(t)$, we get all possible $B \in M_n(\mathbb{R})$. Now, note that $J_0A + A^TJ_0 = 0$ iff $\mathbf{u}^T(J_0A + A^TJ_0)\mathbf{v} = 0$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{2n}$ iff $\langle J_0A\mathbf{v}, \mathbf{u}\rangle + \langle J_0\mathbf{v}, A\mathbf{u} = 0\rangle$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{2n}$ iff $\omega_0(A\mathbf{v}, \mathbf{u}) + \omega_0(\mathbf{v}, A\mathbf{u}) = 0,\ \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^{2n}$. Finally, since $J_0A + A^TJ_0 = 0$ iff $A^T = J_0AJ_0$, if

$$
A = \begin{pmatrix} B & C \\ D & E \end{pmatrix} \in \mathfrak{sp}(n, \mathbb{R})
$$

then

$$\begin{pmatrix} B^T & D^T \\ C^T & E^T \end{pmatrix} = \begin{pmatrix} O & -I \\ I & O \end{pmatrix} \begin{pmatrix} B & C \\ D & E \end{pmatrix} \begin{pmatrix} O & -I \\ I & O \end{pmatrix} = \begin{pmatrix} -E & D \\ C & -B \end{pmatrix}$$

so $B^T = -E$, $D^T = D$ and $C^T = C$. ■

Proposition 21 *The symplectic Lie algebra $\mathfrak{sp}(V,\omega)$ of a symplectic space (V,ω) is a Lie subalgebra of the Lie algebra* $\mathrm{End}(V)$ *with Lie bracket $[\sigma,\tau] = \sigma\circ\tau - \tau\circ\sigma$. In particular, $\mathfrak{sp}(n,\mathbb{R})$ is a Lie subalgebra of $M_{2n}(\mathbb{R})$ with Lie bracket $[A,B] = AB - BA$.*

Proof: Let $\sigma,\tau \in \mathfrak{sp}(V,\omega)$ Then,

$$\begin{aligned}
&\omega([\sigma,\tau](u),v) + \omega(u,[\sigma,\tau](v)) \\
&\quad = \omega(\sigma\circ\tau(u) - \tau\circ\sigma(u),v) + \omega(u,\sigma\circ\tau(v) - \tau\circ\sigma(v)) \\
&\quad = \omega(\sigma\circ\tau(u),v) - \omega(\tau\circ\sigma(u),v) + \omega(u,\sigma\circ\tau(v)) - \omega(u,\tau\circ\sigma(v)) \\
&\quad = -\omega(\tau(u),\sigma(v)) + \omega(\sigma(u),\tau(v)) - \omega(\sigma(u),\tau(v)) + \omega(\tau(u),\sigma(v)) \\
&\quad = 0
\end{aligned}$$

where the penultimate equality follows from the fact that $\omega(\sigma(u),v)+\omega(u,\sigma(v)) = 0$ for all $u,v\in V$, and similarly with τ, because $\sigma,\tau\in\mathfrak{sp}(V,\omega)$. Thus, $[\sigma,\tau]\in\mathfrak{sp}(V,\omega)$, and therefore $\mathfrak{sp}(V,\omega)$ is closed under the bracket operation. ■

Proposition 22 *Let $(\mathbb{R}^{2n})^{\odot 2}$ be the vector space of symmetric 2-tensors of $\mathbb{R}^{2n}$. Then we have the vector space isomorphism*

$$\mathfrak{sp}(n,\mathbb{R}) \cong (\mathbb{R}^{2n})^{\odot 2} \tag{2.53}$$

via the map $\varphi : (\mathbb{R}^{2n})^{\odot 2} \to \mathfrak{sp}(n,\mathbb{R})$ which on simple tensors is given by

$$\begin{aligned}
\varphi(\mathbf{u}\odot\mathbf{v})(\cdot) &:= \omega_0(\cdot,\mathbf{u})\mathbf{v} + \omega_0(\cdot,\mathbf{v})\mathbf{u} && (2.54)\\
&= \langle J_0\cdot,\mathbf{u}\rangle\,\mathbf{v} + \langle J_0\cdot,\mathbf{v}\rangle\,\mathbf{u} && (2.55)
\end{aligned}$$

where the dot indicates the slot for the argument of the map (or matrix viewed as a map).

Proof: Certainly $\varphi(\mathbf{u}\odot\mathbf{v})(\cdot)$ is linear, since ω_0 is in the first argument, and moreover we have that $\varphi(\mathbf{u}\odot\mathbf{v})(\cdot)\in\mathfrak{sp}(n,\mathbb{R})$, for

$$\begin{aligned}
&\omega_0\big(\varphi(\mathbf{u}\odot\mathbf{v})(\mathbf{x}),\mathbf{y}\big)+\omega_0\big(\mathbf{x},\varphi(\mathbf{u}\odot\mathbf{v})(\mathbf{y})\big)\\
&\quad= \omega_0\big(\omega_0(\mathbf{x},\mathbf{u})\mathbf{v}+\omega_0(\mathbf{x},\mathbf{v})\mathbf{u},\mathbf{y}\big)+\omega_0\big(\mathbf{x},\omega_0(\mathbf{y},\mathbf{u})\mathbf{v}+\omega_0(\mathbf{y},\mathbf{v})\mathbf{u}\big)\\
&\quad= \omega_0(\mathbf{x},\mathbf{u})\omega_0(\mathbf{v},\mathbf{y})+\omega_0(\mathbf{x},\mathbf{v})\omega_0(\mathbf{u},\mathbf{y})+\omega_0(\mathbf{y},\mathbf{u})\omega_0(\mathbf{x},\mathbf{v})+\omega_0(\mathbf{y},\mathbf{v})\omega_0(\mathbf{x},\mathbf{u})\\
&\quad= \omega_0(\mathbf{x},\mathbf{u})\omega_0(\mathbf{v},\mathbf{y})+\omega_0(\mathbf{x},\mathbf{v})\omega_0(\mathbf{u},\mathbf{y})-\omega_0(\mathbf{u},\mathbf{y})\omega_0(\mathbf{x},\mathbf{v})-\omega_0(\mathbf{v},\mathbf{y})\omega_0(\mathbf{x},\mathbf{u})\\
&\quad= 0
\end{aligned}$$

Moreover, φ is bijective, because $(\mathbb{R}^{2n})^*\cong\mathbb{R}^{2n}$ via the isomorphism $\mathbf{v}\mapsto\omega(\cdot,\mathbf{v})$ (due to the fact that ω_0 is nondegenerate), and using this map we have the isomorphism $\mathbb{R}^{2n}\otimes\mathbb{R}^{2n}\cong\mathrm{End}(V)(\mathbb{R}^{2n})\cong M_{2n}(\mathbb{R})$, via $\mathbf{u}\otimes\mathbf{v}\mapsto\omega_0(\cdot,\mathbf{v})\mathbf{u}$, and therefore we have the isomorphism

$$\omega_0(\cdot,\mathbf{u})\mathbf{v}+\omega_0(\cdot,\mathbf{v})\mathbf{u}\mapsto\mathbf{v}\otimes\mathbf{u}+\mathbf{u}\otimes\mathbf{v}=2\mathbf{u}\odot\mathbf{v}$$

Thus, the map φ is in fact an isomorphism onto its image. However, since the dimension of $(\mathbb{R}^{2n})^{\odot 2}$ is $\binom{2n+2-1}{2}=\frac{2n(2n+1)}{2}=2n^2+n$, which is also the dimension of $\mathfrak{sp}(n,\mathbb{R})$ (as can be seen from (2.52) of Proposition 20, since we can choose $B\in M_n(\mathbb{R})$ and $C,D\in\mathrm{Sym}_n(\mathbb{R})$, that is we have $n^2+2\frac{n(n+1)}{2}=2n^2+n$ choices), φ is an isomorphism between $(\mathbb{R}^{2n})^{\odot 2}$ and $\mathfrak{sp}(n,\mathbb{R})$. ■

What does a typical matrix A in $\mathfrak{sp}(n,\mathbb{R})$ look like as a symmetric 2-tensor?

Proposition 23 *If $A\in\mathfrak{sp}(n,\mathbb{R})$, then A is mapped by $\varphi^{-1}:\mathfrak{sp}(n,\mathbb{R})\to(\mathbb{R}^{2n})^{\odot 2}$ to*

$$\varphi^{-1}(A)=\frac{1}{2}\sum_{j=1}^{n}\Big(Aq_j\odot p_j-q_j\odot Ap_j\Big)\tag{2.56}$$

where $q_1=\mathbf{e}_1,\dots,q_n=\mathbf{e}_n$ and $p_1=\mathbf{e}_{n+1},\dots,p_n=\mathbf{e}_{2n}$ is the standard symplectic basis for $\mathbb{R}^{2n}$

Proof: If $\mathbf{v} = (v_1, \ldots, v_{2n}) = \sum_{i=1}^n v_i q_i + v_{n+i} p_i$, then the relations $\omega_0(q_i, q_j) = \omega_0(p_i, p_j) = 0$ and $\omega_0(q_i, p_j) = \delta_{ij}$, along with the fact that $v_i = \omega_0(v, p_i)$ and $v_{n+i} = -\omega_0(v, q_i)$, imply that

$$\begin{aligned}
2A\mathbf{v} &= 2A\Big(\sum_{i=1}^n v_i q_i + v_{n+i} p_i\Big) \\
&= 2A\Big(\sum_{i=1}^n \omega_0(\mathbf{v}, p_i) q_i - \omega_0(\mathbf{v}, q_i) p_i\Big) \\
&= A\Big(\sum_{i=1}^n \omega_0(\mathbf{v}, p_i) q_i - \omega_0(\mathbf{v}, q_i) p_i\Big) + \sum_{i=1}^n \Big(\omega_0(A\mathbf{v}, p_i) q_i - \omega_0(A\mathbf{v}, q_i) p_i\Big) \\
&= \sum_{i=1}^n \Big(\omega_0(\mathbf{v}, p_i) A q_i - \omega_0(\mathbf{v}, q_i) A p_i - \omega_0(\mathbf{v}, A p_i) q_i + \omega_0(\mathbf{v}, A q_i) p_i\Big)
\end{aligned}$$

where the last equality follows from the fact that $A \in \mathfrak{sp}(n, \mathbb{R})$. Collecting terms we get

$$\begin{aligned}
2A\mathbf{v} &= \sum_{i=1}^n \Big[\Big(\omega_0(\mathbf{v}, A q_i) p_i + \omega_0(\mathbf{v}, p_i) A q_i\Big) - \Big(\omega_0(\mathbf{v}, A p_i) q_i + \omega_0(\mathbf{v}, q_i) A p_i\Big)\Big] \\
&\mapsto \sum_{i=1}^n \Big(A q_i \odot p_i - q_i \odot A p_i\Big)\mathbf{v}
\end{aligned}$$

under φ^{-1}, and this is true for any $\mathbf{v} \in \mathbb{R}^{2n}$. ■

Remark 15 The identification $\mathfrak{sp}(n, \mathbb{R}) \cong (\mathbb{R}^{2n})^{\odot 2}$,

$$\begin{tikzcd}
A \arrow[r, mapsto] \arrow[d, equal] & \frac{1}{2}\sum_{j=1}^n \Big(A q_j \odot p_j - q_j \odot A p_j\Big) \arrow[d, equal] \\
\omega_0(\cdot, \mathbf{u})\mathbf{v} + \omega_0(\cdot, \mathbf{v})\mathbf{u} & \mathbf{u} \odot \mathbf{v} \arrow[l, mapsto]
\end{tikzcd}$$

allso carries the Lie bracket from $\mathfrak{sp}(n, \mathbb{R})$ to $(\mathbb{R}^{2n})^{\odot 2}$, $[A, B] = AB - BA$. To see what the bracket $[\mathbf{u} \odot \mathbf{v},\ \mathbf{x} \odot \mathbf{y}]$ looks like on simple tensors, we use the identification

$$[\mathbf{u} \odot \mathbf{v},\ \mathbf{x} \odot \mathbf{y}] \mapsto \big[\omega_0(\cdot, \mathbf{u})\mathbf{v} + \omega_0(\cdot, \mathbf{v})\mathbf{u},\ \omega_0(\cdot, \mathbf{x})\mathbf{y} + \omega_0(\cdot, \mathbf{y})\mathbf{x}\big]$$

and compute the right-hand side:

$$
\begin{aligned}
\big[\omega_0(\cdot,&\mathbf{u})\mathbf{v} + \omega_0(\cdot,\mathbf{v})\mathbf{u},\ \omega_0(\cdot,\mathbf{x})\mathbf{y} + \omega_0(\cdot,\mathbf{y})\mathbf{x}\big] \\
&= \Big(\omega_0(\cdot,\mathbf{u})\mathbf{v} + \omega_0(\cdot,\mathbf{v})\mathbf{u}\Big) \circ \Big(\omega_0(\cdot,\mathbf{x})\mathbf{y} + \omega_0(\cdot,\mathbf{y})\mathbf{x}\Big) \\
&\qquad - \Big(\omega_0(\cdot,\mathbf{x})\mathbf{y} + \omega_0(\cdot,\mathbf{y})\mathbf{x}\Big) \circ \Big(\omega_0(\cdot,\mathbf{u})\mathbf{v} + \omega_0(\cdot,\mathbf{v})\mathbf{u}\Big) \\
&= \omega_0\big(\omega_0(\cdot,\mathbf{x})\mathbf{y} + \omega_0(\cdot,\mathbf{y})\mathbf{x},\mathbf{u}\big)\mathbf{v} + \omega_0\big(\omega_0(\cdot,\mathbf{x})\mathbf{y} + \omega_0(\cdot,\mathbf{y})\mathbf{x},\mathbf{v}\big)\mathbf{u} \\
&\qquad - \omega_0\big(\omega_0(\cdot,\mathbf{u})\mathbf{v} + \omega_0(\cdot,\mathbf{v})\mathbf{u},\mathbf{x}\big)\mathbf{y} - \omega_0\big(\omega_0(\cdot,\mathbf{u})\mathbf{v} + \omega_0(\cdot,\mathbf{v})\mathbf{u},\mathbf{y}\big)\mathbf{x} \\
&= \omega_0(\cdot,\mathbf{x})\omega_0(\mathbf{y},\mathbf{u})\mathbf{v} + \omega_0(\cdot,\mathbf{y})\omega_0(\mathbf{x},\mathbf{u})\mathbf{v} + \omega_0(\cdot,\mathbf{x})\omega_0(\mathbf{y},\mathbf{v})\mathbf{u} + \omega_0(\cdot,\mathbf{y})\omega_0(\mathbf{x},\mathbf{v})\mathbf{u} \\
&\qquad - \omega_0(\cdot,\mathbf{u})\omega_0(\mathbf{v},\mathbf{x})\mathbf{y} - \omega_0(\cdot,\mathbf{v})\omega_0(\mathbf{u},\mathbf{x})\mathbf{y} - \omega_0(\cdot,\mathbf{u})\omega_0(\mathbf{v},\mathbf{y})\mathbf{x} - \omega_0(\cdot,\mathbf{v})\omega_0(\mathbf{u},\mathbf{y})\mathbf{x} \\
&= -\omega_0(\mathbf{u},\mathbf{x})\Big(\omega_0(\cdot,\mathbf{v})\mathbf{y} + \omega_0(\cdot,\mathbf{y})\mathbf{v}\Big) - \omega_0(\mathbf{u},\mathbf{y})\Big(\omega_0(\cdot,\mathbf{v})\mathbf{x} + \omega_0(\cdot,\mathbf{x})\mathbf{v}\Big) \\
&\qquad - \omega_0(\mathbf{v},\mathbf{x})\Big(\omega_0(\cdot,\mathbf{u})\mathbf{y} + \omega_0(\cdot,\mathbf{y})\mathbf{u}\Big) - \omega_0(\mathbf{v},\mathbf{y})\Big(\omega_0(\cdot,\mathbf{u})\mathbf{x} + \omega_0(\cdot,\mathbf{x})\mathbf{u}\Big)
\end{aligned}
$$

Therefore, the bracket on simple tensors is

$$[\mathbf{u}\odot\mathbf{v},\ \mathbf{x}\odot\mathbf{y}] = -\omega_0(\mathbf{u},\mathbf{x})\mathbf{v}\odot\mathbf{y} - \omega_0(\mathbf{u},\mathbf{y})\mathbf{v}\odot\mathbf{x} - \omega_0(\mathbf{v},\mathbf{x})\mathbf{u}\odot\mathbf{y} - \omega_0(\mathbf{v},\mathbf{y})\mathbf{u}\odot\mathbf{x} \qquad \blacksquare$$

2.3.4.2 Description of the Symplectic Group $\mathrm{Sp}(n,\mathbb{R})$

We now want to investigate some of the topological and geometric properties of $\mathrm{Sp}(n,\mathbb{R})$, particularly its manifold structure and fundamental group, both of which will be needed later. This description is easiest achieved via the polar decomposition of its elements, which will make the underlying structure transparent.

Theorem 14 (Polar Decomposition) *Let $\big(V,\langle\cdot,\cdot\rangle\big)$ be a finite-dimensional complex inner product space and let $T\in\mathrm{End}(V)$ be any linear operator. Then there exist a unique positive operator $P\in\mathrm{End}(V)^+$ and a (not necessarily unique) unitary operator $U\in\mathrm{U}(V)$ such that*

$$T = U\circ P \tag{2.57}$$

Moreover, if $T\in\mathrm{GL}(V)$, then U is also unique.

Proof: It is known (cf. Roman [92]) that $T^* \circ T$ is positive and has a unique positive square root $\sqrt{T^* \circ T}$, which is also positive, for any linear operator T (the square root of a positive operator A is defined as $\sqrt{A} = \sqrt{\lambda_1}\pi_1 + \cdots + \sqrt{\lambda_k}\pi_k$, where $A = \lambda_1\pi_1 + \cdots + \lambda_k\pi_k$ is the spectral resolution of A, provided of course $\lambda_i \geq 0$ for all i, which is true if A is positive). Hence, let us define P as this square root:

$$P := \sqrt{T^* \circ T}$$

Then, let us define U on $\mathrm{im}(P)$ by

$$U(P(v)) := T(v)$$

The only question is, what happens if P is not injective, what if $P(v) = P(w)$ for $v \neq w$? Clearly we'll have $U(P(v)) = U(P(w))$, but is it necessarily true that $T(v) = T(w)$? The answer is yes, because for all $v \in V$ we have

$$\begin{aligned} \|P(v)\|^2 = \langle P(v), P(v)\rangle &= \left\langle P^2(v), v\right\rangle \\ &= \langle (T^* \circ T)(v), v\rangle = \langle T(v), T(v)\rangle = \|T(v)\|^2 \end{aligned} \tag{2.58}$$

so if $P(v) = P(w)$, then

$$0 = \|0\| = \|P(v) - P(w)\| = \|P(v - w)\| = \|T(v - w)\| = \|T(v) - T(w)\|$$

and therefore $T(v) = T(w)$. Consequently, $\mathrm{im}(U \circ P) = U(P(V)) = T(V) = \mathrm{im}(T)$ and U is well defined. Moreover, U is a linear isometry on $\mathrm{im}(P)$, because for all $v \in V$ we have by (2.58) and our definition of V that

$$\|U(P(v))\| = \|T(v)\| = \|P(v)\|$$

Also U is an isomorphism from $\mathrm{im}(P)$ to $\mathrm{im}(T)$, or $\mathrm{im}(U \circ P) \overset{U}{\cong} \mathrm{im}(T)$: if $U(P(v)) = 0$, then $U(P(v)) = T(v) = 0$, so $v \in \ker(T)$, whence by (2.58) $v \in \ker(P)$ as well, or $P(v) = 0$. Thus $\ker(U) = \{0\}$ and T is injective. Consequently, if β is any basis for $\mathrm{im}(P)$, then $U(\beta)$ is a basis for $\mathrm{im}(U \circ P) = \mathrm{im}(T)$, and so

$$\dim(\mathrm{im}(P)) = |\beta| = |U(\beta)| = \dim(\mathrm{im}(T))$$

We conclude that T is bijective and an isometric isomorphism. Consequently, U is unitary on $\mathrm{im}(P)$, and therefore, if $B = (v_1, \dots, v_k)$ is an orthonormal basis for $\mathrm{im}(P)$, then $U(B) = (U(v_1), \dots, U(v_k))$ is an orthonormal basis for $\mathrm{im}(U \circ P) = \mathrm{im}(T)$. Let us extend B and $U(B)$ to orthonormal bases for V, and then extend the definition of U to an isometry on V for which $T = U \circ P$.

Finally, the uniqueness of P follows from the fact that $P^2 = T^* \circ T$ is positive, so it must have a unique postivie square root, and the square of this square root is unique. If T is additionally an isomorphism, then, since $\ker(P) \subseteq \ker(T) = \{0\}$, so is P, and $U = T \circ P^{-1}$ is unique. ■

Corollary 4 *Let $\big(V, \langle\cdot,\cdot\rangle\big)$ be a finite-dimensional complex inner product space and let $T \in \mathrm{End}(V)$ be any nonzero linear operator. Then, if $T = U \circ P$ is the polar decomposition of T into a (not necessarily unique) unitary operator $U \in \mathrm{U}(V)$ and a unique positive operator $P \in \mathrm{End}(V)^+$,*[1] *then there is a (not necessarily unique) self-adjoint operator $S \in \mathrm{End}(V)$ and a unique self-adjoint operator $B \in \mathrm{End}(V)$ such that*

$$U = e^{iS}, \qquad \textit{and} \qquad P = e^B \tag{2.59}$$

so that

$$T = U \circ P = e^{iS} \circ e^B \tag{2.60}$$

Moreover, if T is an isomorphism, then S is unique.

Proof: Suppose $T = P \circ U$, and let $\sigma(T) = \{\lambda_1, \dots, \lambda_k\}$ be the spectrum of U. By the spectral theorem $\sigma(T) \subseteq S^1$ so $\lambda_j = e^{i\theta_j}$ for some $\theta_j \in [0, 2\pi)$. Define an operator $S \in \mathrm{End}(V)$ functionally by $S := \theta_1\pi_1 + \cdots + \theta_k\pi_k$ where the projections π_j are orthogonal, so that $\sigma(S) = \{\theta_1, \dots, \theta_k\}$ is, by our definition, the spectrum of S. Then by the spectral theorem S normal, and S is self-adjoint because $\sigma(S) \subseteq \mathbb{R}$. Finally, $U = e^{iS} := e^{i\theta_1}\pi_1 + \cdots + e^{i\theta_k}\pi_k$, and so $T = U \circ e^{iS}$. Finally, if T is an isomorphism, then by the previous theorem U is unique, whence the θ_j are uniuqe, whence S is unique.

---

[1] $\mathrm{End}(V)^+$ is the set of positive self-adjoint operators on V. When $V = \mathbb{R}^m$ this space is identified with the convex cone $\mathrm{Sym}_m(\mathbb{R})^+$ in $\mathrm{Sym}_m(\mathbb{R})$ consisting of symmetric positive definite matrices.

If P is positive definite, then it is in particular symmetric and all its eigenvalues are positive, so its spectral resolution is $P = \lambda_1\pi_1 + \cdots + \lambda_k\pi_k$ with all $\lambda_i > 0$. Consequently, taking the natural log of each λ_i gives $\theta_i := \ln(\lambda_i)$ and a self-adjoint operator $B = \theta_1\pi_1 + \cdots + \theta_k\pi_k$, and therefore $\lambda_i = e^{\theta_i}$ for all i, which means $P = e^B$. ∎

Remark 16 If $\big(V, \langle\cdot,\cdot\rangle\big)$ is a *real* finite-dimensional inner product space, then U is still the exponential of a linear operator S, but now we have $e^S = \sum_{n=0}^{\infty} \frac{1}{n!} S^n$, or equivalently $e^S = e^{\theta_1}\pi_1 + \cdots + e^{\theta_k}\pi_k$ where $U = \lambda_1\pi_1 + \cdots + \lambda_k\pi_k$ is the spectral resolution of U, so that $\lambda_i = e^{\theta_i}$ for all i. The arguments used in the above corollary are almost identical. Moreover, in that case U will be *orthogonal* instead of unitary, i.e. $U \in \mathrm{O}(V)$, and if $V = \mathbb{R}^n$, then we may consider $n \times n$ matrices instead of operators, and so if $A \in M_n(\mathbb{R})$, we will have $A = UP$ for an orthogonal matrix $U \in \mathrm{O}(n)$ and a unique positive definite matrix P, and if $A \in \mathrm{GL}(n, \mathbb{R})$, U will be unique as well. ∎

Proposition 24 *If $A \in \mathrm{Sp}(n, \mathbb{R})$, then in the unique polar decomposition $A = UP$ we will have $U \in \mathrm{O}(2n) \cap \mathrm{Sp}(n, \mathbb{R})$ and $P \in \mathrm{Sym}_{2n}(\mathbb{R})^+ \cap \mathrm{Sp}(n, \mathbb{R})$, that is U and P will be symplectic as well. Moreover, if we write $P = e^B$ as in Corollary 4, where B is symmetric (which is equivalent to self-adjoint in the real case), then $B \in \mathfrak{sp}(n, \mathbb{R})$ as well.*

Proof: This follows from the expression (2.46) in Proposition 19, namely $AJ_0A^T = J_0$, for then $(A^T)^{-1}J_0A^{-1} = J_0$, and so $(A^T)^{-1} = J_0AJ_0^{-1}$. Therefore, if $A = UP$, we have

$$(P^T)^{-1}(U^T)^{-1} = (A^T)^{-1} = J_0AJ_0^{-1} = J_0(UP)J_0^{-1} = (J_0UJ_0^{-1})(J_0PJ_0^{-1})$$

But $J_0 \in O(2n)$, because its rows and columns are orthonormal, so $J_0UJ_0^{-1} \in O(2n)$ and $J_0PJ_0^{-1} \in M_{2n}(\mathbb{R})^+$ (i.e. is positive, for if $x = (x_1, x_2) \in \mathbb{R}^{2n}$, then $x^TJ_0 = (x_2, -x_1)$, and $J_0^{-1}x = -J_0x = (x_2, -x_1)^T = (x^TJ_0)^T$, so that $x^TJ_0PJ_0^{-1}x = (x^TJ_0)P(x^TJ_0) \geq 0$ and $= 0$ iff $x^TJ_0 = 0$ iff $x = 0$). Consequently, by the uniqueness of the polar decomposition, $(U^T)^{-1} = J_0UJ_0^{-1}$ and $(P^T)^{-1} = J_0PJ_0^{-1}$, which shows that $U, P \in \mathrm{Sp}(n, \mathbb{R})$.

Finally, if $U = e^B$ and $B = B^T$, then

$$e^B = P = U = J_0 P^{-1} J_0^{-1} = J_0 e^{-B} J_0^{-1} = J_0 \left(\sum_{n=0}^{\infty} \frac{1}{n!} (-B)^n \right) J_0^{-1}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \big(J_0 (-B) J_0^{-1} \big)^n = e^{-J_0 B J_0^{-1}}$$

so we must have $B^T = B = -J_0 B J_0^{-1}$, or $B^T J_0 = -J_0 B$, and therefore $J_0 B + B^T J_0 = 0$, which shows that $B \in \mathfrak{sp}(n, \mathbb{R})$ by Proposition 20, equation (2.51). ■

The following result was first stated in Arnol'd [3].

Proposition 25 *If we identify $\mathbb{R}^{2n}$ with $\mathbb{C}^n$ via $(p, q) \mapsto p + iq$, then*

$$\begin{aligned} \mathrm{U}(n) &= \mathrm{Sp}(n, \mathbb{R}) \cap \mathrm{O}(2n) \\ &= \mathrm{Sp}(n, \mathbb{R}) \cap \mathrm{GL}(n, \mathbb{C})) \\ &= \mathrm{O}(2n) \cap \mathrm{GL}(n, \mathbb{C}) \end{aligned}$$

is a maximal compact real Lie subgroup of $\mathrm{Sp}(n, \mathbb{R})$ *of dimension n^2.*

Proof: If K is a subgroup of $\mathrm{Sp}(n, \mathbb{R})$ containing $\mathrm{Sp}(n, \mathbb{R}) \cap \mathrm{O}(2n)$, then for any $A \in K$ with polar decomposition $A = UP$ we also have $U \in K$ and therefore $P = U^{-1} A \in K$. But P is positive definite and $\det P = 1$, since $P^T J_0 P = J_0$, so either $P = I$ or else some eigenvalue of P is greater than 1 (see Corollary 4 and its proof). In the first case $A = U \in \mathrm{O}(2n)$, and in the second case $P^k \in K$ for all $k \in \mathbb{N}$ and $\|P^k\| \nearrow \infty$, in which case K is not compact. That is, if $\mathrm{Sp}(n, \mathbb{R}) \cap \mathrm{O}(2n) \leq K \leq \mathrm{Sp}(n, \mathbb{R})$ and K is compact, then $K = \mathrm{Sp}(n, \mathbb{R}) \cap \mathrm{O}(2n)$.

Recall the results of Section 2.3.1 on complex structures. Complexifying $\mathbb{R}^{2n}$ by the standard complex structure J_0 gives the identification $\mathbb{R}^{2n}_{J_0} \cong \mathbb{C}^n$ (Example 4) and the identification of $\mathrm{GL}(n, \mathbb{C})$ with the subgroup of $\mathrm{GL}(2n, \mathbb{R})$ consisting of matrices X commuting with J_0 (Corollary 3). We then have the following conditions:

(1) $X \in \mathrm{U}(n) \iff XX^* = X^* X = I$

(2) $X = \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \in \mathrm{GL}(n, \mathbb{C}) \subseteq \mathrm{GL}(2n, \mathbb{R}) \iff XJ_0 = J_0 X$ (Corollary 3)

(3) $X \in \mathrm{Sp}(n,\mathbb{R}) \iff X^T J_0 X = J_0$ (Proposition 19)

(4) $X \in \mathrm{O}(2n) \iff X^T X = XX^T = I$

Out of the last three conditions, we claim that any two of them imply the third. For instance, (3) and (4) imply (2): If $XX^T = I$ and $X^T J_0 X = J_0$, then $J_0 X = (XX^T)J_0 X = X(X^T J_0 X) = XJ_0$, so $X \in \mathrm{GL}(n,\mathbb{C})$. Similarly, (2) and (3) imply (4): If $XJ_0 = J_0 X$ and $X^T J_0 X = J_0$, then $J_0 = X^T J_0 X = X^T X J_0$, so multiplying both sides by $-J_0$ gives $I = X^T X$, which means $X \in \mathrm{O}(2n)$. Finally, (2) and (4) imply (3): If $XJ_0 = J_0 X$ and $X^T X = XX^T = I$, then $X^T J_0 X = X^T X J_0 = IJ_0 = J_0$, and $X \in \mathrm{Sp}(n,\mathbb{R})$. This proves the second and third equalities.

Let us now show the first equality, $\mathrm{U}(n) = \mathrm{Sp}(n,\mathbb{R}) \cap \mathrm{O}(n)$. Consider a unitary matrix $X = A + iB \in \mathrm{U}(n) \subseteq \mathrm{GL}(n,\mathbb{C})$ under the identification of $\mathrm{GL}(n,\mathbb{C})$ with $\mathrm{GL}(2n,\mathbb{R})$. The adjoint X^* becomes the transpose, X^T (cf Proposition 17), therefore the defining condition for a unitary matrix $X = A + iB \in \mathrm{U}(n) \subseteq \mathrm{GL}(n,\mathbb{C})$, $X^*X = XX^* = I$, translates to $X^T X = XX^T = I$ in $\mathrm{GL}(2n,\mathbb{R})$, which means $X \in \mathrm{O}(2n)$. But since $X \in \mathrm{U}(n) \subseteq \mathrm{GL}(n,\mathbb{C})$, we also have $XJ_0 = J_0 X$, so $J_0 = IJ_0 = X^T X J_0 = X^T J_0 X$, and $X \in \mathrm{Sp}(n,\mathbb{R})$, too. For the reverse inclusion, we reverse the argument: Since $X^T X = XX^T = I$ and $X^T J_0 X = J_0$ imply $X \in \mathrm{GL}(n,\mathbb{C})$, the identity $X^T X = XX^T = I$ means $X^*X = XX^* = I$ under the identification of $\mathrm{GL}(n,\mathbb{C})$ with it's image in $\mathrm{GL}(2n,\mathbb{R})$, and $X \in \mathrm{U}(n)$.

Lastly, let $\mathcal{H}(n)$ denote the real vector space of $n \times n$ Hermitian matrices. The real dimension of $\mathcal{H}(n)$ is n^2: $(A + iB)^* = A^T - iB^T = A + iB$ iff $A = A^T$ and $B = -B^T$, and $\dim \mathrm{Sym}_n(\mathbb{R}) = n(n+1)/2$ and $\dim \mathrm{Skew}_n(\mathbb{R}) = n(n-1)/2$). Now define the smooth map

$$F : M_n(\mathbb{C}) \to \mathcal{H}(n)$$

$$F(A) := A^*A - I$$

We claim that F is a submersion at each point $A \in \mathrm{U}(n)$: its derivative at $A \in \mathrm{U}(n)$ is given by the product rule,

$$DF(A) = A^*(\cdot) + (\cdot)^*A \in \mathrm{Hom}_{\mathbb{R}}\big(M_n(\mathbb{C}), \mathcal{H}(n)\big)$$

i.e. $DF(A)B = A^*B + B^*A$, and $DF(A)$ is easily seen to be surjective, for if $B \in \mathcal{H}(n)$, let $\frac{1}{2}BA \in M_n(\mathbb{C})$ and since $A \in \mathrm{U}(n)$ we will have $A^*A = I$, so $DF(A)(\frac{1}{2}AB) = \frac{1}{2}(A^*AB + B^*A^*A) = B$. Consequently, (Lee [73, Proposition 4.1, Corollary 5.14]) $F^{-1}(0) = \mathrm{U}(n)$ is a smooth embedded submanifold of $M_n(\mathbb{C})$ of codimension equal to $\dim \mathcal{H}(n) = n^2$, therefore of dimension $2n^2 - n^2 = n^2$.

■

Proposition 26 $\mathrm{Sp}(n,\mathbb{R})$ *is homeomorphic to the product of the unitary operators and the space of self-adjoint operators in* $\mathfrak{sp}(n,\mathbb{R})$*, which is in turn homeomorphic to* $U(n) \times \mathbb{R}^{n^2+n}$*,*

$$\begin{aligned} \mathrm{Sp}(n,\mathbb{R}) &\approx \mathrm{U}(n) \times \big(\mathrm{Sym}_{2n}(\mathbb{R}) \cap \mathfrak{sp}(n,\mathbb{R})\big) && (2.61)\\ &\approx \mathrm{U}(n) \times \mathbb{R}^{n^2+n} && (2.62)\end{aligned}$$

Consequently, $\mathrm{Sp}(n,\mathbb{R})$ *is a connected* $(2n^2+n)$*-dimensional Lie group, and its fundamental group is* $\mathbb{Z}$*,*

$$\pi_1\big(\mathrm{Sp}(n,\mathbb{R})\big) \cong \mathbb{Z} \tag{2.63}$$

Proof: Since all $A \in \mathrm{Sp}(n,\mathbb{R})$ can be uniquely written as $A = UP$ where $U \in \mathrm{U}(n) = \mathrm{Sp}(n,\mathbb{R}) \cap \mathrm{O}(2n)$ and $P \in \mathrm{Sp}(n,\mathbb{R}) \cap \mathrm{Sym}_{2n}(\mathbb{R})^+$, and moreover P is positive iff $P = e^B$ for some $B \in \mathrm{Sym}_{2n}(\mathbb{R}) \cap \mathfrak{sp}(n,\mathbb{R})$, the first result follows. That $\big(SA(\mathbb{R}^{2n}) \cap \mathfrak{sp}(n,\mathbb{R})\big) \approx \mathbb{R}^{n^2+n}$ is seen as follows: If $A \in \mathrm{Sym}_{2n}(\mathbb{R}) \cap \mathfrak{sp}(n,\mathbb{R})$, then $A = A^T$ and $AJ_0 + J_0A^T = 0$. Writing A as a block matrix,

$$A = \begin{pmatrix} B & C \\ D & E \end{pmatrix}$$

these conditions translate into

$$0 = AJ_0 + J_0A^T = \begin{pmatrix} B & C \\ D & E \end{pmatrix}\begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} + \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}\begin{pmatrix} B & C \\ D & E \end{pmatrix} = \begin{pmatrix} C & -B \\ E & -D \end{pmatrix} + \begin{pmatrix} -D & -E \\ B & C \end{pmatrix}$$

which means

$$C = D, \quad B = -E$$

i.e.

$$A = \begin{pmatrix} B & C \\ C & -B \end{pmatrix}$$

But since $A = A^T$, we must also have $C = C^T$ and $B = B^T$, and therefore A is entirely determined by the two symmetric matrices $B, C \in \mathrm{Sym}_n(\mathbb{R})$. Since $\dim\big(\mathrm{Sym}_n(\mathbb{R})\big) = \frac{n(n+1)}{2}$, we have $\dim\big(\mathrm{Sym}_{2n}(\mathbb{R}) \cap \mathfrak{sp}(n,\mathbb{R})\big) = 2\dim\big(\mathrm{Sym}_n(\mathbb{R})\big) = n(n+1)$, and $\mathrm{Sym}_{2n}(\mathbb{R}) \cap \mathfrak{sp}(n,\mathbb{R})$ is therefore homeomorphic to $\mathbb{R}^{n^2+n}$.

Thus, $A \in \mathrm{Sp}(n,\mathbb{R})$ iff $A = Ue^B$ for unique $U \in \mathrm{U}(n)$ and $B \in \mathrm{Sym}_{2n}(\mathbb{R}) \cap \mathfrak{sp}(n,\mathbb{R}) \approx \mathbb{R}^{2n}$, and so

$$\mathrm{Sp}(n,\mathbb{R}) \approx \mathrm{U}(n) \times \big(\mathrm{Sym}_{2n}(\mathbb{R}) \cap \mathfrak{sp}(n,\mathbb{R})\big) \approx U(n) \times \mathbb{R}^{n^2+n}$$

Since $\mathrm{U}(n)$ is connected and its fundamental group is $\mathbb{Z}$, the theorem follows. (This last statement follows from the long exact sequence of homotopy groups arising from the fibration $\mathrm{U}(n-1) \to \mathrm{U}(n) \to S^{2n-1}$ and the observation that the inclusion $\mathrm{U}(1) \to \mathrm{U}(n)$ induces an isomorphism of fundamental groups. See Theorem 4.55, p. 374, and Example 4.55, p. 381, in Hatcher [56]). ■

Lemma 1 *Let $A \in \mathrm{Sp}(n,\mathbb{R})$ and let $\sigma(A)$ denote its spectrum. Then $\lambda \in \sigma(A)$ iff $\lambda^{-1} \in \sigma(A)$ and they both have the same (even) multiplicity. In particular, if $\pm 1 \in \sigma(A)$ then it occurs with even multiplicity. Lastly, if $\lambda, \lambda' \in \sigma(A)$ and $\lambda\lambda' \neq 1$, then for all eigenvectors $\mathbf{u} \in E_\lambda$ and $\mathbf{v} \in E_{\lambda'}$ we have $\omega_0(\mathbf{u},\mathbf{v}) = 0$. In particular ω_0 vanishes on E_λ whenever $\lambda \neq 1$.*

Proof: Any $A \in \mathrm{Sp}(n,\mathbb{R})$ satisfies $A^T J_0 A = J_0$, or equivalently $A^T = J_0 A^{-1} J_0^{-1}$, which shows that A^T is similar to A^{-1} and therefore A and A^{-1} have the same eigenvalues. If $\lambda \in \sigma(A)$, then $\lambda \in \sigma(A^{-1})$, too, so if $\lambda \neq \pm 1$, it must have even multiplicity. If $\lambda = -1 \in \sigma(A)$, then the fact that $1 = \det(A) = \prod_{\lambda \in \sigma(A)} \lambda$ implies that -1 must have even multiplicity, too, and therefore if $\lambda = 1 \in \sigma(A)$, it will also have even multiplicity. The last statement follows from the identity

$$\lambda\lambda' \langle \mathbf{u}, J_0\mathbf{v}\rangle = \langle A\mathbf{u}, J_0 A\mathbf{v}\rangle = \omega_0(A\mathbf{u}, A\mathbf{v}) = \omega_0(\mathbf{u},\mathbf{v}) = \langle \mathbf{u}, J_0\mathbf{v}\rangle$$

Similarly, if $\lambda \neq 1$ and $\mathbf{u}, \mathbf{v} \in E_\lambda$, then $|\lambda|^2 \langle \mathbf{u}, J_0\mathbf{v}\rangle = \langle \mathbf{u}, J_0\mathbf{v}\rangle$, which must therefore equal 0. ■

Lemma 2 *Let $A \in \mathrm{Sp}(n,\mathbb{R}) \cap \mathrm{Sym}_{2n}(\mathbb{R})^+$ be a symplectic, symmetric positive definite matrix. Then $A^\alpha \in \mathrm{Sp}(n,\mathbb{R})$ for all real $\alpha > 0$.*

Proof: We must show that each $(A^\alpha)^*\omega_0 = \omega_0$ for each $\alpha > 0$. Diagonalizing A involves conjugating by a change-of-basis matrix $g \in \mathrm{GL}(2n, \mathbb{R})$, $D = \mathrm{diag}(\lambda_1, \dots, \lambda_{2n}) = gAg^{-1}$, and all $\lambda_j > 0$ (Proposition 5 and Theorem 7), so that $A^\alpha = g^{-1}D^\alpha g = g^{-1}\,\mathrm{diag}(\lambda_1^\alpha, \dots, \lambda_{2n}^\alpha)g$. Suppose $E_{\lambda_j^\alpha}$ is the eigenspace of A^α corresponding to λ_j^α (this space is identical to the eigenspace E_{λ_j} of A corresponding to λ_j, since $A^\alpha = g^{-1}D^\alpha g$). Then whenever $\lambda_i\lambda_j \neq 1$ we will also have $\lambda_i^\alpha\lambda_j^\alpha \neq 1$, so $E_{\lambda_i} \perp_{\omega_0} E_{\lambda_j}$ implies $E_{\lambda_i^\alpha} \perp_{\omega_0} E_{\lambda_j^\alpha}$. That lemma also guarantees that if $\lambda_j^\alpha \neq 1$, which is always the case whenever $\lambda_j \neq 1$, then ω_0 vanishes on E_{λ_j}. Thus, for $\lambda_i\lambda_j \neq 1$ or $\lambda_i = \lambda_j \neq 1$ we will have

$$\omega_0(A^\alpha\mathbf{u}, A^\alpha\mathbf{v}) = (\lambda_i\lambda_j)^\alpha\omega(\mathbf{u}, \mathbf{v}) = 0 = \omega_0(\mathbf{u}, \mathbf{v})$$

whenever $\mathbf{u} \in E_{\lambda_i^\alpha}$ and $\mathbf{v} \in E_{\lambda_j^\alpha}$. The remaining possibility is $\lambda_i = 1$, in which case we trivially have $\omega_0(A^\alpha\mathbf{u}, A^\alpha\mathbf{v}) = \omega_0(\mathbf{u}, \mathbf{v})$ for $\mathbf{u}, \mathbf{v} \in E_1$. ■

Proposition 27 $\mathrm{U}(n)$ *is a deformation retract of* $\mathrm{Sp}(n, \mathbb{R})$*, so the quotient group* $\mathrm{Sp}(n, \mathbb{R})/\mathrm{U}(n)$ *is contractible.*

Proof: In the polar decomposition of A (Proposition 14), $A = UP$, we have $P = (A^*A)^{1/2} = (A^TA)^{1/2} \in \mathrm{Sp}(n, \mathbb{R})$, which actually lies in $\mathrm{Sp}(n, \mathbb{R}) \cap \mathrm{Sym}_{2n}(\mathbb{R})^+ \subseteq \mathrm{GL}(2n, \mathbb{R})$ by the previous lemma, so we can solve for U, namely $U = A(A^TA)^{-1/2} \in \mathrm{U}(n)$. In fact, by that lemma all positive real powers t of $(A^TA)^{-1/2}$ are also symplectic, so $(A^TA)^{-t/2}A \in \mathrm{Sp}(n, \mathbb{R})$ for all $t \geq 0$. This allows us to define the function:

$$\begin{aligned} &H : \mathrm{Sp}(n, \mathbb{R}) \times [0, 1] \to \mathrm{Sp}(n, \mathbb{R}) \\ &H(A, t) := A(A^TA)^{-t/2} \end{aligned}$$

We note that H is continuous, and satisfies $H(A, 0) = A = \mathrm{id}_{\mathrm{Sp}(n,\mathbb{R})}(A)$ and $H(A, 1) = A(A^TA)^{-1/2} = U \in \mathrm{U}(n)$, so H defines a homotopy between the identity on $\mathrm{Sp}(n, \mathbb{R})$ and the retraction $r : \mathrm{Sp}(n, \mathbb{R}) \to \mathrm{U}(n)$, $r(A) = A(A^TA)^{-1/2}$. That is, H is a deformation retract. ■

2.3.5 The Interaction between Inner Products, Symplectic Forms, and Complex Structures

2.3.5.1 The Existence of Compatible Complex Structures

Let V be a $2n$-dimensional real vector space, let g be an inner product on V, and let ω be a symplectic form on V. That is, g is a symmetric bilinear form with signature $(2n, 0)$ and ω is a skew-symmetric, eqivalently alternating, bilnear form, and both g and ω are nondegenerate. We sometimes use the isomorphism $\bigwedge^2(V) \cong \mathrm{Hom}^2_{\mathrm{skew}}(V;\mathbb{R})$ to view ω as an abstract bivector, when that is more convenient.

Proposition 28 *Given an inner product g and a symplectic form ω on a finite dimensional real vector space V, there exists a linear operator $S \in \mathrm{End}(V)$ such that*

(1) *$\omega(u,v) = g(S(u),v)$, for all $u,v \in V$.*

(2) *S is skew self adjoint, that is $S^* = -S$, or equivalently $g(S(u),v) = g(u,-S(v))$ for all $u,v \in V$.*

(3) *S has matrix representation with respect to a symplectic basis β for ω as J_0,*

$$[S]_\beta = J_0 = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}.$$

(4) *S is a complex structure on V, that is $S^2 = -I$.*

(5) *$S \in \mathfrak{sp}(V,\omega)$, that is $\omega(S(u),v) + \omega(u,S(v)) = 0$ for all $u,v \in V$.*

(6) *$S \in \mathrm{Sp}(V,\omega)$, that is $S^*\omega = \omega$, or explicitly $\omega(S(u),S(v)) = \omega(u,v)$ for all $u,v \in V$.*

(7) *$S \in \mathrm{O}(V,g)$, that is $S^*g = g$, or explicitly $g(S(u),S(v)) = g(u,v)$ for all $u,v \in V$.*

Proof: (1) and (3): Let $\beta = (u_1, \ldots, u_n, v_1, \ldots, v_n)$ be a symplectic basis for V with respect to ω. Since $M_{2n}(\mathbb{R}) \cong \mathrm{End}(V)$, there is a unique $S \in \mathrm{End}(V)$ such that $[S]_\beta = J_0 = [\omega]_\beta$. Then,

$$\begin{pmatrix} g(S(u_1), u_1) & g(S(u_2), u_1) & \cdots & g(S(v_n), u_1) \\ g(S(u_1), u_2) & g(S(u_2), u_2) & \cdots & g(S(v_n), u_2) \\ \vdots & \vdots & \ddots & \vdots \\ g(S(u_1), v_n) & g(S(u_2), v_n) & \cdots & g(S(v_n), v_n) \end{pmatrix}$$

$$= [S]_\beta = J_0 = [\omega]_\beta = \begin{pmatrix} \omega(u_1, u_1) & \omega(u_2, u_1) & \cdots & \omega(v_n, u_1) \\ \omega(u_1, u_2) & \omega(u_2, u_2) & \cdots & \omega(v_n, u_2) \\ \vdots & \vdots & \ddots & \vdots \\ \omega(u_1, v_n) & \omega(u_2, v_n) & \cdots & \omega(v_n, v_n) \end{pmatrix}$$

by virtue of β being a symplectic basis. Thus, since $g \circ (S \times \mathrm{id}_V)$ agrees with ω on the basis elements, we must have $g \circ (S \times \mathrm{id}_V) = \omega$ on all of V.

(2) From the skew symmetry of ω, we have

$$g(S(u), v) = \omega(u, v) = -\omega(v, u) = -g(S(v), u) = g(u, -S(v))$$

But we already know that, since V is finite dimensional, the adjoint of S exists, is unique, and satisfies $g(S(u), v) = g(u, S^*(v))$ for all $u, v \in V$, so from the above equation we must have that $S^* = -S$.

(4) Since $[S^2]_\beta = [S]_\beta^2 = J_0^2 = -I_{2n} = [-I]_\beta$, we see that $S^2 = -I$.

(5) By (1), (2) and (4) we have

$$\begin{aligned} \omega(S(u), v) + \omega(u, S(v)) &= g(S^2(u), v) + g(S(u), S(v)) \\ &= g(S^2(u), v) + g(u, -S^2(v)) \\ &= -g(u, v) + g(u, v) \\ &= 0 \end{aligned}$$

(6) Again by (1)-(2)

$$\begin{aligned} \omega(S(u), S(v)) &= g(S^2(u), S(v)) \\ &= -g(u, S(v)) \\ &= -g(S(v), u) \\ &= -\omega(v, u) \\ &= \omega(u, v) \end{aligned}$$

(7) By (1), (2) and (4), $g(S(u), S(v)) = g(-S^2(u), v) = g(u, v)$. ■

The above proposition shows that if we are given a symplectic form ω and an inner product g on V, we get a complex structure J on V which satisfies $g(J(u), v) = \omega(u, v)$ for all $u, v \in V$. The next proposition shows that if we are given ω and J, then, as long as $J \in \mathrm{Sp}(V, \omega)$, we get g and this time $g(u, v) = \omega(u, J(v))$. Since $J \in \mathrm{Sp}(V, \omega)$, we automatically recover the previous formula $g(J(u), v) = \omega(J(u), J(v)) = \omega(u, v)$. Here g will be symmetric and nondegenerate, but not necessarily positive definite. In other words, g may have signature (p, q) with $q \neq 0$.

Proposition 29 *Let (V, ω) be a symplectic vector space and let J be a complex structure on V. If in addition $J \in \mathrm{Sp}(V, \omega)$, i.e. $J^*\omega = \omega$ (that is $\omega(J(u), J(v)) = \omega(u, v)$ for all $u, v \in V$, see the next section 2.3.4), then there is a unique nondegenerate symmetric bilinear form $g_J \in \mathrm{Hom}^2_{\mathrm{sym}}(V; \mathbb{R})$ given by*

$$g_J(u, v) := \omega(u, J(v)) \tag{2.64}$$

for all $u, v \in V$.

Proof: That g_J as defined in (2.64) is nondegenerate follows from the fact that $J \in \mathrm{Sp}(V, \omega) \subseteq \mathrm{GL}(V)$ and the fact that ω is nondegenerate (and therefore a reflexive form). That g_J is symmetric follows from the fact that $J \in \mathrm{Sp}(V, \omega)$:

$$g_J(u, v) = \omega(u, J(v)) = \omega(J(u), J^2(v)) = -\omega(J(u), v) = \omega(v, J(u)) = g(v, u)$$ ■

Remark 17 We remark that g_J may not be positive definite, and so may not be an inner product. For examle, on $(\mathbb{R}^2, \omega_0)$ we can take $J(x,y) = (y, -x)$, i.e.

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Then clearly $J^2 = -I$, but

$$\begin{aligned} g_J\big((x,y),(x,y)\big) &= \omega_0\big((x,y), J(x,y)\big) = \omega_0\big((x,y),(y,-x)\big) \\ &= dx \wedge yx\big((x,y),(y,-x)\big) = dx(x,y)dy(y,-x) - dx(y,-x)dy(x,y) = -x^2 - y^2 \end{aligned}$$

and so g_J is negative definite. ■

Definition 4 When $g_J = \omega \circ (\mathrm{id}_V \times J)$ is positive definite, that is

(1) $g(v,v) = \omega\big(v, J(v)\big) \geq 0$ for all $v \in V$ with equality only when $v = 0$. (J is *tamed by* ω)

then we say, following McDuff and Salamon [76], [77], that J is **tamed** by ω or is ω-**tame**. When additionally $J \in \mathrm{Sp}(V, \omega)$, i.e. when $J^*\omega = \omega$, explicitly

(2) $\omega\big(J(u), J(v)\big) = \omega(u,v)$ for all $u, v \in V$. (*J is **calibrated** by ω, cf Audin [12]*)

we say J is **compatible with** ω or is ω-**compatible**. Berndt [16] calls J satisfying (1)-(2) *positive compatible*, but we will stick with the McDuff and Salamon terminology here. We will also say, in this case, following Berndt, that g is **compatible with** ω. When J is ω-compatible we call the triple (V, ω, J) a **Khäler vector space** (see Berndt [16]).

Let us follow Audin [12] and McDuff and Salamon [76] [77] in denoting the set of all complex structures on V by

$$\mathcal{J}(V)$$

and the set of all ω-tame, ω-calibrated, and ω-compatible complex structures on V, respectively, by

$$\mathcal{J}_\tau(V,\omega), \qquad \mathcal{J}_c(V,\omega), \qquad \mathcal{J}(V,\omega) = \mathcal{J}_\tau(V,\omega) \cap \mathcal{J}_c(V,\omega)$$

Finally, let us also introduce the following notation, respectively, for the set of all inner products (i.e. metrics, or positive definite symmetric bilinear forms) and the set of all symplectic forms on V:

$$\mathcal{M}(V), \qquad \Omega(V) \qquad \blacksquare$$

Remark 18 If J is ω-tame only, the bilinear form $g_J(u,v) := \omega(u,J(v))$ is not necessarily symmetric, since we do not have the last equality in $\omega(u,J(v)) = -\omega(J(v),u) = \omega(v,J(u))$. However, if we average $\omega(u,J(v))$ and $\omega(v,J(u))$, then we *do* get symmetry: define

$$g_J(u,v) = \frac{1}{2}\Big(\omega(u,J(v)) + \omega(v,J(u))\Big)$$

Then clearly $g_J(u,v) = g_J(v,u)$, and g_J is a metric. $\blacksquare$

2.3.5.2 Interlude on Hermitian Inner Products on Symplectic Spaces

We pause here to note that, once we have a compatible almost complex structure $J \in \mathcal{J}(V,\omega)$, the symplectic form ω and the induced metric g_J combine to give a Hermitian inner product H_J on the complexified space V_J, making (V_J, H_J) a Hermitian space.

Proposition 30 (Hermitian Inner Product) *Given $J \in \mathcal{J}(V,\omega)$, denote the associated positive definite metric g_J. From this metric we can construct a* ***hermitian inner product*** *H_J on the complexified space V_J,*

$$H_J : V_J \times V_J \to \mathbb{C}$$

$$H_J(v,w) := g_J(v,w) + i\omega(v,w)$$

Proof: We first show that H_J is sesquilinear and satisfies $H_J(w,v) = \overline{H_J(v,w)}$. Since additivity in each coordinate follows from that of g_J and ω, we only need to show the scalar multiplication property to show that H_J is sesquilinear. Towards this end, let $a+ib, c+id \in \mathbb{C}$, then

$$
\begin{aligned}
& H_J((a+ib)v,(c+id)w) \\
& \quad = \; g_J((a+ib)v,(c+id)w) + i\omega((a+ib)v,(c+id)w) \\
& \quad = \; g_J(av+bJ(v),\ cw+dJ(w)) + i\omega(av+bJ(v),\ cw+dJ(w)) \\
& \quad = \; acg_J(v,w) + bdg_J(J(v),J(w)) + adg_J(v,J(w)) + bcg_J(J(v),w) \\
& \qquad\qquad + iac\omega(v,w) + ibd\omega(v,w) + iad\omega(v,J(w)) + ibc\omega(J(v),w) \\
& \quad = \; acg_J(v,w) + bdg_J(v,w) - ad\omega(v,w) + bc\omega(v,w) \\
& \qquad\qquad + iac\omega(v,w) + ibd\omega(v,w) + iadg_J(v,w) - ibcg_J(v,w) \\
& \quad = \; \big[(ac+bd) + i(ad-bc)\big]\big[g_J(v,w) + i\omega(v,w)\big] \\
& \quad = \; (a-ib)(c+id)H_J(v,w) \\
& \quad = \; \overline{(a+ib)}(c+id)H_J(v,w)
\end{aligned}
$$

Next,

$$
\begin{aligned}
\overline{H_J(w,v)} &= \overline{g_J(w,v) + i\omega(w,v)} \\
&= g_J(w,v) - i\omega(w,v) \\
&= g_J(v,w) + i\omega(v,w) \\
&= H_J(v,w)
\end{aligned}
$$

Finally, we show that H_J is positive definite. If $v \in V_J \backslash \{0\}$, then

$$H_J(v,v) = g_J(v,v) + i\omega(v,v) = g_J(v,v) > 0$$

since ω is skew-symmetric and g_J is positive definite. ■

Example 6 Consider $\mathbb{C}^n$ with its standard Hermitian inner product

$$H(\mathbf{z},\mathbf{w}) = \sum_{j=1}^{n} \overline{z}_j w_j$$

Writing $z_j = a_j + ib_j$, $w_j = c_j + id_j$, we have

$$\overline{z}_j w_j = (a_j - ib_j)(c_j + id_j) = (a_j c_j + b_j d_j) + i(a_j d_j - b_j c_j)$$

so that the sum becomes

$$\begin{aligned} H(\mathbf{z}, \mathbf{w}) &= \sum_{j=1}^{n} (a_j c_j + b_j d_j) + i(a_j d_j - b_j c_j) \\ &= (a_1, \ldots, a_n, b_1, \ldots, b_n) \cdot (c_1, \ldots, c_n, d_1, \ldots, d_n) \\ &\qquad i(-b_1, \ldots, -b_n, a_1, \ldots, a_n) \cdot (c_1, \ldots, c_n, d_1, \ldots, d_n) \\ &= (\mathbf{a}, \mathbf{b}) \cdot (\mathbf{c}, \mathbf{d}) + i\big(J_0(\mathbf{a}, \mathbf{b})\big) \cdot (\mathbf{c}, \mathbf{d}) \\ &= g_{J_0}((\mathbf{a}, \mathbf{b}), (\mathbf{c}, \mathbf{d})) + i\omega_0((\mathbf{a}, \mathbf{b}), (\mathbf{c}, \mathbf{d})) \end{aligned}$$

That is, under the identification $\mathbb{C}^n \cong \mathbb{R}^{2n}_{J_0}$,

$$\mathbf{z} = (a_1 + ib_1, \ldots, a_n + ib_n) \mapsto (a_1, \ldots, a_n, b_1, \ldots, b_n) = (\mathbf{a}, \mathbf{b})$$

we have that

$$H_{J_0} = g_{J_0} + i\omega_0 = H$$

so H_{J_0} is precisely H. ■

Remark 19 Let us consider the unitary group

$$\mathrm{U}(n) = \{X \in \mathrm{GL}(n, \mathbb{C}) \mid X^* H = H\} = \{X \in \mathrm{GL}(n, \mathbb{C}) \mid X^* X = X X^* = I\}$$

under the identification $\mathbb{C}^n \cong \mathbb{R}^{2n}_{J_0}$ with $H_{J_0} \longleftrightarrow H$. Here, $\mathrm{GL}(n, \mathbb{C})$ is identified with the stabilizer subgroup $\mathrm{GL}(2n, \mathbb{R})_{J_0}$ of $\mathrm{GL}(2n, \mathbb{R})$ via the map

$$X = A + iB \mapsto f(X) := \begin{pmatrix} A & -B \\ B & A \end{pmatrix}$$

(see Corollary 3 above) and the unitary group $U(n)$ is identitfied with $\mathrm{O}(2n) \cap \mathrm{Sp}(n, \mathbb{R})$ (Proposition 25 above), so $X \in \mathrm{U}(n)$ iff $f(X)^* g_{J_0} = g_{J_0}$ and $f(X)^* \omega_0 = \omega_0$. We now show that this relation holds for any symplectic space (V,ω) and any fixed $J \in \mathcal{J}(V,\omega)$, where we use the Hermitian inner product $H_J = g_J + i\omega$ on V_J:

$$\mathrm{U}(V_J, H_J) = \mathrm{O}(V, g_J) \cap \mathrm{Sp}(V,\omega) \qquad \blacksquare$$

Definition 5 Given a symplectic vector space (V,ω) and an ω-compatible complex structure $J \in \mathcal{J}(V,\omega)$, we may construct the Hermitian inner product H_J on V_J. Then we say that a $\mathbb{C}$-linear map $T \in \mathrm{End}_{\mathbb{C}}(V_J)$ is H_J**-Hermitian** if $H_J(T\cdot,\cdot)$ is Hermitian, H_J**-skew-Hermitian** if $H_J(T\cdot,\cdot)$ is skew-Hermitian, and H_J**-unitary** if $T^*H_J = H_J$.

We have seen that a given $T \in \mathrm{End}_{\mathbb{R}}(V)$ lies in $\mathrm{End}_{\mathbb{C}}(V_J)$ iff $T \circ J = J \circ T$. Under what conditions is T also H_J-unitary?

Proposition 31 *Let $J \in \mathcal{J}(V,\omega)$ and $T \in \mathrm{End}_{\mathbb{R}}(V)$. Then $T \in \mathrm{End}_{\mathbb{C}}(V_J)$ and is H_J-unitary iff $T \in \mathrm{O}(V, g_J) \cap \mathrm{Sp}(V,\omega)$, i.e.*

$$\mathrm{U}(V_J, H_J) = \mathrm{O}(V, g_J) \cap \mathrm{Sp}(V,\omega) \tag{2.65}$$

which is the abstract analog of Proposition 25.

Proof: If $T \in \mathrm{End}_{\mathbb{C}}(V_J)$ is H_J-unitary, then $T \circ J = J \circ T$ and $T^*H_J = H_J$, so $T^*g_J + iT^*\omega = g_J + i\omega$, which shows that $T^*g_J = g_J$ and $T^*\omega = \omega$, so $T \in \mathrm{O}(V, g_J) \cap \mathrm{Sp}(V,\omega)$. Conversely, if $T \in \mathrm{O}(V, g_J) \cap \mathrm{Sp}(V,\omega)$, then $T^*g_J = g_J$ and $T^*\omega = \omega$, so in order to show that T^*H_J we also need to have that $J \circ T = T \circ J$. $\blacksquare$

In the remainder of this section we will look at the topological and geometric properties of the spaces $\mathcal{J}(V)$, $\mathcal{J}_\tau(V,\omega)$, $\mathcal{J}_c(V,\omega)$, $\mathcal{J}(V,\omega)$, $\mathcal{M}(V)$ and $\Omega(V)$. Our first result, the fact that every symplectic space is linearly symplectomorphic to the standard space $(\mathbb{R}^{2n},\omega_0)$, in fact naturally with respect to each $J \in \mathcal{J}(V)$, means it will suffice to study $\mathcal{J}(\mathbb{R}^{2n})$, $\mathcal{M}(\mathbb{R}^{2n}) = \mathrm{Sym}_{2n}(\mathbb{R})^+$, $\Omega(\mathbb{R}^{2n})$ and related spaces.

2.3.5.3 All Symplectic Vector Spaces (V,ω) are Linearly Symplectormorphic to $(\mathbb{R}^{2n},\omega_0)$

Proposition 32 *Let (V,ω) be a $2n$-dimensional symplectic space. Then, for every complex structure $J \in \mathcal{J}(V)$ on V there is a linear symplectomorphism*

$$\Phi : \mathbb{R}^{2n} \to V$$

such that the following diagram commutes

$$\begin{tikzcd} \mathbb{R}^{2n} \arrow[r, "\Phi"] \arrow[d, "J_0"'] & V \arrow[d, "J"] \\ \mathbb{R}^{2n} \arrow[r, "\Phi"'] & V \end{tikzcd} \qquad i.e. \qquad J \circ \Phi = \Phi \circ J_0$$

and so $(V,\omega) \cong (\mathbb{R}^{2n},\omega_0)$ with the extra compatibility condition on complex structures specified above. Moreover,

$$\Phi^*\omega = \omega_0$$

Proof: Let $V^{1,0} = \ker(J^{\mathbb{C}} - iI)$, $V^{0,1} = \ker(J^{\mathbb{C}} + iI)$ be the eigenspaces of the complexified $J^{\mathbb{C}}$ on $V^{\mathbb{C}} = V^{1,0} \oplus V^{0,1}$, where $\dim(V^{1,0}) = \dim(V^{0,1}) = n$ (see Section 2.3.1). Then any basis $\beta^{\mathbb{C}} = \{w_j := u_j + iv_j \mid j = 1,\dots,n\}$ for $V^{1,0}$ gives a basis $\beta = (u_1,\dots,u_n,v_1,\dots,v_n)$ for V, and

$$J(u_j) = -v_j, \qquad J(v_j) = u_j$$

on this basis, since $J(w_j) = iw_j$ iff $J(u_j + iv_j) = -v_j + iu_j$. Let $\Phi : \mathbb{R}^{2n} \to V$ be given by

$$\Phi(\mathbf{x}) = \Phi(q_1,\dots,q_n,p_1,\dots,p_n) := \sum_{j=1}^{n}(q_ju_j - p_jv_j)$$

where $\mathbf{x} = \sum_{j=1}^{n} q_j\mathbf{e}_j + p_j\mathbf{e}_{n+j} \in \mathbb{R}^{2n}$. Then for all $\mathbf{x} \in \mathbb{R}^{2n}$ we have

$$(\Phi \circ J_0)(\mathbf{x}) = \Phi(-p_1,\dots,-p_n,q_1,\dots,q_n) = -\sum_{j=1}^{n}(p_ju_j + q_jv_j)$$

$$= -J\Bigg(\sum_{j=1}^{n}(p_jv_j - q_ju_j)\Bigg) = J\Bigg(\sum_{j=1}^{n}(q_ju_j - p_jv_j)\Bigg) = (J \circ \Phi)(\mathbf{x})$$

Finally, note that Φ is a symplectomorphism, that is $\Phi^*\omega = \omega_0$, for if $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2)$ where $\mathbf{u_1} = (q_1, \ldots, q_n)$ and $\mathbf{u}_2 = (p_1, \ldots, p_n)$, and similarly $\mathbf{v} = (\mathbf{v}_1, \mathbf{v}_2)$, then, supposing (as we may in light of Theorem 10) that $(u_1, \ldots, u_n, v_1, \ldots, v_n)$ is a symplectic basis for (V, ω),

$$
\begin{aligned}
\Phi^*\omega(\mathbf{u}, \mathbf{v}) &= \omega(\Phi(\mathbf{u}), \Phi(\mathbf{v})) \\
&= \omega\left(\sum_{i=1}^n q_i u_i - p_i v_i, \sum_{j=1}^n q_j' u_j - p_j' v_i\right) \\
&= \sum_{i,j=1}^n \left[q_i q_i' \omega(u_i, u_j) - q_i p_j' \omega(u_i, v_j) - p_i q_j' \omega(v_i, u_j) + p_i p_j' \omega(v_i, v_j)\right] \\
&= \sum_{i,j=1}^n p_i q_j' \omega_0(\mathbf{e}_j, \mathbf{e}_i) - p_j' q_i \omega_0(\mathbf{e}_j, \mathbf{e}_i) \\
&= \langle \mathbf{u}_1, \mathbf{v}_2 \rangle - \langle \mathbf{u_2}, \mathbf{u}_1 \rangle \\
&= \omega_0(\mathbf{u}, \mathbf{v})
\end{aligned}
$$

where the last equality follows from (4) of Proposition 18. ■

In light of the above it is enough to consider $(\mathbb{R}^{2n}, \omega_0, J_0)$ in what follows.

2.3.5.4 The Space of Metrics $\mathcal{M}(\mathbb{R}^{2n})$

Let us begin with the simpler case, that of the space of metrics on $\mathbb{R}^{2n}$,

$$\mathcal{M}(\mathbb{R}^{2n}) \cong \mathrm{Sym}_{2n}(\mathbb{R})^+$$

Understanding this space will motivate a similar study of the other spaces. We refer to Lang [39, pp. 322-338] and Gil-Medrano and Michor [39] for a fuller account of what follows. To begin with, $\mathcal{M}(\mathbb{R}^{2n})$ is a convex cone in the vector space $\mathrm{Sym}_{2n}(\mathbb{R})$ (it is closed under sums and *positive* real scalar multiplication) and contains the identity, so it is trivially path-connected and contractible. To see that it is an open submanifold of $\mathrm{Sym}_{2n}(\mathbb{R})$ of dimension $n(n+1)/2$,[2] note that

[2] $\mathrm{Sym}_{2n}(\mathbb{R})$ is itself a closed vector subspace, hence closed submanifold, of $M_{2n}(\mathbb{R})$, being finite-dimensional.

Proposition 33 *The exponential* $\exp : M_{2n}(\mathbb{R}) \to \mathrm{GL}(2n,\mathbb{R})$, $\exp(A) := \sum_{k=0}^{\infty} \frac{1}{k!}A^k$, *restricts to a* C^∞*-diffeomorphism,*

$$\exp|_{\mathrm{Sym}_{2n}(\mathbb{R})} : \mathrm{Sym}_{2n}(\mathbb{R}) \xrightarrow{\sim} \mathrm{Sym}_{2n}(\mathbb{R})^+$$

Consequently, $\mathrm{Sym}_{2n}(\mathbb{R})^+$ *is a smooth manifold of dimension* $n(n+1)/2$, *which is path connected and contractible.*

Proof: Note that any $A \in \mathrm{Sym}_{2n}(\mathbb{R})$ is diagonalizable (Proposition 5 and Theorem 7) and its eigenvalues are real ($Av = \lambda v$ and $A = A^T$ iff $\lambda\|v\|^2 = \langle \lambda v, v\rangle = \langle Av, v\rangle = \langle v, Av\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\|v\|^2$ implies $\lambda = \overline{\lambda}$). But diagonalization is just conjugating by a change-of-basis matrix $g \in \mathrm{GL}(2n,\mathbb{R})$, $A = g\,\mathrm{diag}(\lambda_1,\dots,\lambda_{2n})g^{-1}$, which makes exponentiation very transparent,

$$e^A = e^{g\,\mathrm{diag}(\lambda_1,\dots,\lambda_{2n})g^{-1}} = ge^{\mathrm{diag}(\lambda_1,\dots,\lambda_{2n})}g^{-1} = g\,\mathrm{diag}(e^{\lambda_1},\dots,e^{\lambda_{2n}})g^{-1}$$

Moreover, this expression gives an explicit diagonalization of e^A, by the same change-of-basis matrix, which shows that the eigenvalues of e^A are positive, and therefore by Proposition 5 and Theorem 7 demonstrates that e^A is symmetric and positive definite. But more is true: $\exp|_{\mathrm{Sym}_{2n}(\mathbb{R})}$ is smooth and invertible, with smooth inverse

$$\begin{aligned}
&\log : \mathrm{Sym}_{2n}(\mathbb{R})^+ \to \mathrm{Sym}_{2n}(\mathbb{R}) \\
&\log(A) = \log(g\,\mathrm{diag}(\alpha_1,\dots,\alpha_{2n})g^{-1}) := g\,\mathrm{diag}(\log\alpha_1,\dots,\log\alpha_{2n})g^{-1}
\end{aligned}$$

where g is the change-of-basis matrix diagonalizing A, which necessarily has positive eigenvalues α_j. ■

Remark 20 There is a natural Riemannian structure to $\mathcal{M}(\mathbb{R}^{2n})$, which uses the Riemannian structure of $\mathrm{Sym}_{2n}(\mathbb{R})$ given by[3]

$$\begin{aligned}
&\langle\cdot,\cdot\rangle_{\mathrm{tr}} : \mathrm{Sym}_{2n}(\mathbb{R}) \times \mathrm{Sym}_{2n}(\mathbb{R}) \to \mathbb{R} \\
&\langle A, B\rangle_{\mathrm{tr}} := \mathrm{tr}(AB)
\end{aligned}$$

[3] It is symmetric and bilinear by the properties of the trace map, and positive definite since if A is diagonalized as $A = gD_Ag^{-1}$, then $\langle A, A\rangle_{\mathrm{tr}} = \mathrm{tr}(A^2) = \mathrm{tr}((gDg^{-}1)(gDg^{-1})) = \mathrm{tr}(gD^2g^{-1}) = \mathrm{tr}(D^2) = \sum_{j=1}^{2n}\lambda_j^2$, and this equals zero iff all $\lambda_j = 0$ iff $A = 0$.

where we identified the tangent space to a point $A \in \mathrm{Sym}_{2n}(\mathbb{R})$ with $\mathrm{Sym}_{2n}(\mathbb{R})$ itself, since the tangent bundle to any finite-dimensional vector space is trivial (vector spaces are parallelizable). The norm induced by $\langle\cdot,\cdot\rangle_{\mathrm{tr}}$ is called the **Frobenius norm**. Using this metric, we can get a metric on $\mathrm{Sym}_{2n}(\mathbb{R})^+$, viewing it as an open submanifold of $\mathrm{Sym}_{2n}(\mathbb{R})$ with the same tangent spaces above each point: if $P \in \mathrm{Sym}_{2n}(\mathbb{R})^+$ and $A, B \in T_P \mathrm{Sym}_{2n}(\mathbb{R})^+ \cong \mathrm{Sym}_{2n}(\mathbb{R})$, define

$$\langle\cdot,\cdot\rangle_P : T_P \mathrm{Sym}_{2n}(\mathbb{R})^+ \times T_P \mathrm{Sym}_{2n}(\mathbb{R})^+ \to \mathbb{R}$$
$$\langle A, B\rangle_P := \big\langle P^{-1}A, P^{-1}B\big\rangle_{\mathrm{tr}} = \mathrm{tr}(P^{-1}AP^{-1}B)$$

The symmetry and biliniearity follow as before, while the positive definiteness follows from

$$\mathrm{tr}(P^{-1}AP^{-1}A) = \mathrm{tr}\big((P^{-1/1}AP^{-1/2})(P^{-1/2}AP^{-1/2})\big)) = \mathrm{tr}(C^2)$$

where $C = P^{-1/2}AP^{-1/2}$, since $\mathrm{tr}(C^2) > 0$ if $A \neq 0$, for we may suppose A diagonalized by an appropriate choice of basis. If $P : (-\varepsilon, \varepsilon) \to \mathrm{Sym}_{2n}(\mathbb{R})^+$ is a curve, then its length from $P(0)$ to $P(t)$ in the induced norm is

$$s(t) = \int_0^t \|P'(s)\|\, ds = \int_0^t \sqrt{\mathrm{tr}\big((P(s)^{-1}P'(s))^2\big)}ds$$

and the differential of this length is

$$\left(\frac{ds}{dt}\right)^2 = \mathrm{tr}\Big((P(t)^{-1}P'(t)\Big), \qquad \text{or, abbreviated,} \qquad ds^2 = \mathrm{tr}\big((P^{-1}dP)^2\big)$$

This is called the **trace metric** on $\mathrm{Sym}_{2n}(\mathbb{R})^+$, and makes it a Riemannian manifold. ■

2.3.5.5 The Space of Symplectic Forms $\Omega(V)$

Recall that $\Omega(V)$ denotes the space of symplectic forms on a $2n$-dimensional real vector space V. We have an analogous description of $\Omega(V)$ as we did of $\mathcal{M}(V)$, via the same methods.

Proposition 34 *The general linear group* $\mathrm{GL}(V)$ *acts transitively on* $\Omega(V)$ *by*

$$\mathrm{GL}(V) \times \Omega(V) \to \Omega(V) \tag{2.66}$$

$$g \cdot \omega := (g^{-1})^*\omega = \omega\big(g^{-1}(\cdot), g^{-1}(\cdot)\big) \tag{2.67}$$

and the stabilizer of any $\omega \in \Omega(V)$ *is the symplectic group* $\mathrm{Sp}(V,\omega)$*, which is the group of isometries on the symplectic vector space* (V,ω)*,*

$$\mathrm{GL}(V)_\omega = \mathrm{Sp}(V,\omega) \tag{2.68}$$

Consequently, by the orbit-stabilizer theorem for Lie groups we have the diffeomorphism

$$\Omega(V) \approx \mathrm{GL}(2n,\mathbb{R})/\,\mathrm{Sp}(n,\mathbb{R}) \tag{2.69}$$

and $\Omega(V)$ *is an embedded submanifold of* $\mathrm{End}(V)$ *of dimension* $\dim \Omega(V) = 2n^2 - n$*.*

Proof: To see that the $\mathrm{GL}(V)$-action is transitive, let $\omega \in \Omega(V)$ and $g \in \mathrm{GL}(V)$, and define $\tilde{\omega} := g^*\omega = \omega\big(g(\cdot), g(\cdot)\big) \in \Omega(V)$ ($g^*\omega$ is clearly skew-symmetric, and it is nondegenerate since both g and ω are nondegenerate). Then, clearly $(g^{-1})^*\tilde{\omega} = \omega$. To understand this transitivity better, note that if $\alpha = (q,p)$ is an ω-symplectic basis for V, then $\beta = g^{-1}\alpha = (q',p')$ is an $\tilde{\omega}$-symplectic basis for V, for the relations (2.35) for ω imply

$$\tilde{\omega}(q_i', q_j') = g^*\omega(q_i', q_j') = \omega(g(q_i'), g(q_j')) = \omega(q_i, q_j) = 0$$

and similarly $\tilde{\omega}(p_i', p_j') = 0$ and $\tilde{\omega}(q_i', p_j') = \delta_{ij}$. Thus, $\alpha = g\beta$, and g is a change-of-basis matrix from β to α, and g^{-1} therefore changes α to β. Finally, it is clear that $g \in \mathrm{GL}(V)_\omega$ iff $g \cdot \omega = \omega$ iff $g^*\omega = \omega$, iff $g \in \mathrm{Sp}(V,\omega)$.

The fact that the orbit-stabilizer bijection is a diffeomorphism is analogous to the situation with $\mathrm{GL}(2n,\mathbb{R})$ acting on $\mathcal{J}(\mathbb{R}^{2n})$, as in Proposition 35 below: View $\Omega(V)$ as a subset of the vector space (hence manifold) $\mathrm{Hom}^2_{\mathbb{R}}(V;\mathbb{R}) \cong \mathrm{End}(V)$, and view the action of $\mathrm{GL}(V)$ as taking place on all of $\mathrm{End}(V)$. Then by the orbit theorem for Lie groups (see Kirillov [64, Theorem 3.29] and Lee [73, Theorems 21.18, 21.20]) we have that the stabilizer $\mathrm{GL}(V)_\omega$ of ω is a closed Lie subgroup of $\mathrm{GL}(V)$ and the map $\varphi : \mathrm{GL}(V)/\,\mathrm{GL}(V)_\omega \to \mathrm{End}(V)$, $\varphi(g \cdot \mathrm{GL}(V)_\omega) \mapsto g \cdot \omega$ is an immersion which is equivariantly diffeomorphic onto its image $\omega(V)$. The dimension of $\Omega(V)$ is therefore $\dim \mathrm{GL}(V) - \dim \mathrm{GL}(V)_\omega = 4n^2 - (2n^2 + n) = 2n^2 - n$. ∎

2.3.5.6 The Space of Complex Structures $\mathcal{J}(\mathbb{R}^{2n})$

Let us now explore the case of $\mathcal{J}(\mathbb{R}^{2n})$. We will tackle $\mathcal{J}_\tau(\mathbb{R}^{2n})$ and $\mathcal{J}_c(\mathbb{R}^{2n})$ in the next section.

Example 7 Consider the symplectic space $(\mathbb{R}^2, \omega_0)$ and the associated spaces $\mathcal{J}(\mathbb{R}^2)$ and $\mathcal{J}(\mathbb{R}^2, \omega_0)$ of complex structures and ω_0-compatible complex structures. We have seen (Remark 17) that for a given $J \in \mathcal{J}(\mathbb{R}^2)$ we do not always have $g_J \in \mathrm{Sym}_2(\mathbb{R})^+$ (the space of *positive definite* 2×2 symmetric matrices), though, by Proposition 29, we *do* have that $g_J \in \mathrm{Sym}_2(\mathbb{R})$ so long as $J \in \mathrm{Sp}(1, \mathbb{R})$. In the 2×2 case this is always true, because

$$\mathcal{J}(\mathbb{R}^{2n}) \subseteq \mathrm{Sp}(1, \mathbb{R})$$

To see this, write $J = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{J}(\mathbb{R}^2)$ and note that the condition $J^2 = -I$ translates to

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^2 = \begin{pmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{pmatrix}$$

which in turn reduces to the equations

$$a^2 = d^2 = -1 - bc, \qquad\qquad d(a+b) = c(a+d) = 0$$

We see from the first equation that neither b nor c can equal 0, so along with the second equation this implies that $a = -d$. We thus have two possibilities, either $d = -a \neq 0$, in which case $d = b = -a$ and $c = (a^2+1)/a$, or else $d = a = 0$ and $c = -1/b$:

$$J = \begin{pmatrix} a & -a \\ \frac{a^2+1}{a} & -a \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & b \\ -\frac{1}{b} & 0 \end{pmatrix}$$

If we use the characterization of $\mathrm{Sp}(1, \mathbb{R})$ given in Proposition 19, that $A \in \mathrm{Sp}(1, \mathbb{R})$ iff $A^T J_0 A = J_0$, then either form of J above is easily seen to satisfy this identity.

A second observation we make is that $\mathrm{GL}(2, \mathbb{R})$ *acts transitively by conjugation on* $\mathcal{J}(\mathbb{R}^2)$. Indeed, $\mathcal{J}(\mathbb{R}^2)$ is the orbit of J_0:

$$\mathcal{J}(\mathbb{R}^2) = \mathrm{GL}(2, \mathbb{R}) \cdot J_0$$

For clearly, if $g \in \mathrm{GL}(2,\mathbb{R})$, then $J := gJ_0g^{-1} \in \mathcal{J}(\mathbb{R}^2)$, since

$$J^2 = (gJ_0g^{-1})^2 = gJ_0g^{-1}gJ_0g^{-1} = gJ_0^2g^{-1} = -gg^{-1} = -I$$

Conversely, if $J \in \mathcal{J}(\mathbb{R}^2)$, let $b_1 = (1,0)$ and $b_2 := J(b_1)$. Then (b_1,b_2) is a basis for $\mathbb{R}^2$, since $J \in \mathrm{GL}(2,\mathbb{R})$ (indeed, $b_2 = (a, \frac{a^2+1}{a})$, if $a \neq 0$, or $b_2 = (0, -\frac{1}{b})$, if $a = 0$ in the expression for J above). Let g be the change-of-basis matrix from the new basis $\beta = (b_1,b_2)$ to the standard basis $\rho = (\mathbf{e}_1,\mathbf{e}_2)$: in the first case, $a \neq 0$, we have

$$g = M_{\beta,\rho} = \big([b_1]_\rho \;\; [b_2]_\rho\big) = \begin{pmatrix} 1 & a \\ 0 & \frac{a^2+1}{a} \end{pmatrix}$$

so that

$$gJ_0g^{-1} = \begin{pmatrix} 1 & a \\ 0 & \frac{a^2+1}{a} \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & \frac{-a^2}{a^2+1} \\ 0 & \frac{a}{a^2+1} \end{pmatrix} = \begin{pmatrix} a & -a \\ \frac{a^2+1}{a} & -a \end{pmatrix} = J$$

In the second case, when $b = 0$, we have

$$g = \begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{b} \end{pmatrix}$$

and similarly

$$gJ_0g^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{b} \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -b \end{pmatrix} = \begin{pmatrix} 0 & b \\ -\frac{1}{b} & 0 \end{pmatrix} = J$$

Finally, observe that the stabilizer of J_0 under the $\mathrm{GL}(2,\mathbb{R})$-action is the space of all real 2×2 matrices of the type

$$g = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

which is isomorphic to $\mathbb{C} \cong \mathrm{GL}(1,\mathbb{C})$, which follows from Corollary 3 above. Thus, by the orbit-stabilizer theorem,

$$\mathcal{J}(\mathbb{R}^2) = \mathrm{GL}(2,\mathbb{R}) \cdot J_0 \approx \mathrm{GL}(2,\mathbb{R})/\,\mathrm{GL}(2,\mathbb{R})_{J_0} \cong \mathrm{GL}(2,\mathbb{R})/\,\mathrm{GL}(1,\mathbb{C}) \qquad \blacksquare$$

In fact, the above example generalizes to all $(\mathbb{R}^{2n}, \omega_0)$:

Proposition 35 *The general linear group* $\mathrm{GL}(2n,\mathbb{R})$ *acts transitively by conjugation on space of complex structures* $\mathcal{J}(\mathbb{R}^{2n})$*, and the stabilizer of the standard complex structure* J_0 *is isomorphic to* $\mathrm{GL}(n,\mathbb{C})$*. Consequently,* $\mathcal{J}(\mathbb{R}^{2n})$ *is an embedded submanifold of* $M_{2n}(\mathbb{R})$ *of dimension* $2n^2$*, and the orbit-stabilizer bijection is a diffeomorphism,*

$$\mathcal{J}(\mathbb{R}^{2n}) = \mathrm{GL}(2n,\mathbb{R})\cdot J_0 \tag{2.70}$$

$$\approx \mathrm{GL}(2n,\mathbb{R})/\,\mathrm{GL}(2n,\mathbb{R})_{J_0} \tag{2.71}$$

$$\cong \mathrm{GL}(2n,\mathbb{R})/\,\mathrm{GL}(n,\mathbb{C}) \tag{2.72}$$

Moreover, the tangent space at $J\in\mathcal{J}(\mathbb{R}^{2n})$ *is*

$$T_J\mathcal{J}(\mathbb{R}^{2n})\cong\mathfrak{g}(2n,\mathbb{R})/\mathfrak{g}(n,\mathbb{C})=M_{2n}(\mathbb{R})/M_n(\mathbb{C})\cong\mathbb{R}^{2n^2}$$

Proof: It is clear that if $J=g\cdot J_0=gJ_0g^{-1}$, then $J^2=-I$, so $J\in\mathcal{J}(\mathbb{R}^{2n})$, which shows that $\mathrm{GL}(2n,\mathbb{R})\cdot J_0\subseteq\mathcal{J}(\mathbb{R}^{2n})$. For the reverse inclusion, let $J\in\mathcal{J}(\mathbb{R}^{2n})$, and consider a J-complex basis $\beta=(b_1,\ldots,b_n,J(b_1),\ldots,J(b_n))$ for $\mathbb{R}^{2n}$, for example let $b_j=\mathbf{e}_j$ be the first n coordinates of the standard orthonormal basis $\rho=(\mathbf{e}_1,\ldots,\mathbf{e}_{2n})$, and let the remaining n coordinates be $J(\mathbf{e}_j)$. Then let $g=M_{\beta,\rho}$ be the change-of-basis matrix changing β-coordinates into ρ-coordinates,

$$g:=M_{\beta,\rho}=\big[[b_1]_\rho\ \cdots\ [b_n]_\rho\ \ [J(b_1)]_\rho\ \cdots\ [J(b_n)]_\rho\big]\in\mathrm{GL}(2n,\mathbb{R})$$

and note that

$$gJ_0g^{-1}=J \tag{2.73}$$

This is because the matrix representation of J in the basis β is

$$[J]_\beta=\big[[J(b_1)]_\beta\ \cdots\ [J(J(b_n))]_\beta\big]=J_0$$

so equation (2.73) means (cf Roman [92, Corollary 2.17])

$$J=[J]_\rho=M_{\beta,\rho}[J]_\beta M_{\beta,\rho}^{-1}$$

which is, therefore, just J represented the standard basis ρ instead of the J-complex basis β. We have thus shown that $\mathcal{J}(\mathbb{R}^{2n})\subseteq\mathrm{GL}(2n,\mathbb{R})\cdot J_0$, and therefore $\mathcal{J}(\mathbb{R}^{2n})=\mathrm{GL}(2n,\mathbb{R})\cdot J_0$.

The statement $\mathrm{GL}(2n,\mathbb{R})_{J_0} \cong \mathrm{GL}(n,\mathbb{C})$ is just Corollary 3, and the bijection $\mathrm{GL}(2n,\mathbb{R}) \cdot J_0 \approx \mathrm{GL}(2n,\mathbb{R})/\,\mathrm{GL}(2n,\mathbb{R})_{J_0}$ is just the orbit-stabilizer theorem. To see the manifold structure of $\mathcal{J}(\mathbb{R}^{2n})$, view $\mathcal{J}(\mathbb{R}^{2n})$ as a subset of the vector space (hence manifold) $M_{2n}(\mathbb{R})$, and view the action of $\mathrm{GL}(2n,\mathbb{R})$ as taking place on all of $M_{2n}(\mathbb{R})$. Then by the orbit theorem for Lie groups (see Kirillov [64, Theorem 3.29] and Lee [73, Theorems 21.18, 21.20]) we have that the stabilizer $\mathrm{GL}(2n,\mathbb{R})_{J_0} \cong \mathrm{GL}(n,\mathbb{C})$ of J_0 is a closed Lie subgroup of $\mathrm{GL}(2n,\mathbb{R})$ and the map $\varphi : \mathrm{GL}(2n,\mathbb{R})/\,\mathrm{GL}(2n,\mathbb{R})_{J_0} \to M_{2n}(\mathbb{R})$, $\varphi(g \cdot \mathrm{GL}(2n,\mathbb{R})_{J_0}) = g \cdot J_0$ is an immersion which is equivariantly diffeomorphic onto its image $\mathcal{J}(\mathbb{R}^{2n})$. The dimension of $\mathcal{J}(\mathbb{R}^{2n})$ is therefore $\dim \mathrm{GL}(2n,\mathbb{R}) - \dim \mathrm{GL}(n,\mathbb{C}) = 4n^2 - 2n^2 = 2n^2$ (recall that $\mathrm{GL}(n,\mathbb{C}) = \det^{-1}(\mathbb{C}^*)$ is open in $M_n(\mathbb{C}) \cong M_n(\mathbb{R}) \otimes M_n(\mathbb{R})$, so has dimension $2n^2$). See Theorem 3.29 in Kirillov for the claim about tangent spaces. ■

Remark 21 Another way to describe the inclusion $\mathrm{GL}(2n,\mathbb{R}) \cdot J_0 \supseteq \mathcal{J}(\mathbb{R}^{2n})$ is via Proposition 32 above, by taking $V = \mathbb{R}^{2n}$ and noting that $AJ = J_0 A$ for some $A \in \mathrm{GL}(2n,\mathbb{R}^{2n})$, which holds iff $J = A^{-1} J_0 A = \varphi(A)$ for that same A. ■

Recall that $\mathrm{GL}(2n,\mathbb{R}) = \mathrm{GL}^+(2n,\mathbb{R}) \sqcup \mathrm{GL}^-(2n,\mathbb{R})$ is disconnected with two connected components, $\mathrm{GL}^+(2n,\mathbb{R}) = \det^{-1}(0,\infty)$ containing the identity I and $\mathrm{GL}^-(2n,\mathbb{R}) = \det^{-1}(-\infty,0)$ not containing it. The diffeomorphism $\varphi : \mathcal{J}(\mathbb{R}^{2n}) \approx \mathrm{GL}(2n,\mathbb{R})/\,\mathrm{GL}(2n,\mathbb{R})_{J_0}$ of the previous proposition therefore implies the disconnectedness of $\mathcal{J}(\mathbb{R}^{2n})$ into two connected components, $\mathcal{J}^{\pm}(\mathbb{R}^{2n})$. Since $\det J_0 = 1$, we see that $J_0 \in \mathrm{GL}^+(2n,\mathbb{R})$, so that

$$J_0 \in \mathrm{GL}^+(2n,\mathbb{R}) \cdot J_0 = \mathcal{J}^+(\mathbb{R}^{2n}) \qquad \text{and} \qquad J_0 \notin \mathrm{GL}^-(2n,\mathbb{R}) \cdot J_0 = \mathcal{J}^-(\mathbb{R}^{2n})$$

We have thus shown:

Corollary 5 *The space of complex structures $\mathcal{J}(\mathbb{R}^{2n})$ is disconnected, with two connected components, $\mathcal{J}^+(\mathbb{R}^{2n})$ containing J_0 and $\mathcal{J}^-(\mathbb{R}^{2n})$ not containing it,*

$$\mathcal{J}(\mathbb{R}^{2n}) = \mathcal{J}^+(\mathbb{R}^{2n}) \sqcup \mathcal{J}^-(\mathbb{R}^{2n})$$

and where the connected components are diffeomeorphic to the two components $\mathrm{GL}^{\pm}(2n,\mathbb{R})$ *of* $\mathrm{GL}(2n,\mathbb{R})$ *as*

$$\mathcal{J}^{+}(\mathbb{R}^{2n}) \approx \mathrm{GL}^{+}(2n,\mathbb{R})/\,\mathrm{GL}(n,\mathbb{C}) \qquad \blacksquare$$

Finally, note that $\mathrm{GL}(2n,\mathbb{R})$ is homotopy equivalent to $\mathrm{O}(2n)$, $\mathrm{GL}^{+}(2n,\mathbb{R})$ is homotopy equivalent to $\mathrm{SO}(2n)$ and $\mathrm{GL}(n,\mathbb{C})$ is homotopy equivalent to $U(n)$, each via Gram-Schmidt orthonormalization. We will demonstrate this on the example of $\mathrm{O}(2n)$. Define the continuous map $h : \mathrm{GL}(2n,\mathbb{R}) \times I \to \mathrm{GL}(2n,\mathbb{R})$ as follows: for $A = (\mathbf{a}_1, \dots, \mathbf{a}_{2n}) \in \mathrm{GL}(2n,\mathbb{R})$, with columns $\mathbf{a}_j$, let $B = (\mathbf{b}_1, \dots, \mathbf{b}_{2n}) \in \mathrm{O}(2n)$ be the result of applying the Gram-Schmidt process to the columns $\mathbf{a}_j$ of A,

$$\begin{aligned}
\mathbf{b}_1 &:= \ell_{11}\mathbf{a}_1 \\
\mathbf{b}_2 &:= \ell_{21}\mathbf{a}_1 + \ell_{22}\mathbf{a}_2 \\
&\ \vdots \\
\mathbf{b}_2 &:= \ell_{2n,1}\mathbf{a}_1 + \cdots + \ell_{2n,2n}\mathbf{a}_{2n}
\end{aligned}$$

where all $\ell_{jj} > 0$, and define the vectors

$$\mathbf{c}_j := t(\ell_{j1}\mathbf{a}_1 + \cdots + \ell_{j,j-1}\mathbf{a}_{j-1}) + (t\ell_{jj} + 1 - t)\mathbf{a}_j$$

Using these, we define the map h by letting $h(A,t)$ be the matrix with columns the coordinates of these vectors. Then, h is clearly continuous and satisfies $h(A,0) = A$ and $h(A,1) = B$, and $h(B,t) = B$ for all $B \in \mathrm{O}(2n)$, so it defines a deformation. As a result:

Corollary 6 *We have the following homotopy equivalence:*

$$\mathcal{J}^{+}(\mathbb{R}^{2n}) \simeq \mathrm{SO}(2n)/U(n)$$

The set $\mathcal{J}^{+}(\mathbb{R}^{2n})$ *is therefore the space of all orientation-preserving complex structures on* $\mathbb{R}^{2n}$*.* $\blacksquare$

2.3.5.7 The Spaces of Tame and Calibrated Complex Structures, $\mathcal{J}_\tau(V,\omega)$ and $\mathcal{J}_c(V,\omega)$

The main statements about the manifold structure and homotopy type of the spaces $\mathcal{J}_\tau(V,\omega)$ and $\mathcal{J}_c(V,\omega)$ will be stated and proven in a single proposition.

Proposition 36

(1) *$\mathcal{J}_\tau(V,\omega)$ is diffeomorphic to the open unit ball B_1 of the $2n^2$-dimensionsl vector space $Y = \{A \in M_{2n}(\mathbb{R}) \,|\, AJ_0 + J_0A = 0\}$, and therefore is a $2n^2$-dimensional smooth contractible manifold.*

(2) *$\mathcal{J}_c(V,\omega)$ is diffeomorphic to the open unit ball B_1 of the $n(n+1)$-dimensionsl vector space $Z = \{A \in \mathrm{Sym}_{2n}(\mathbb{R}) \mid AJ_0 + J_0A = 0\}$, and therefore is a $n(n+1)$-dimensional smooth contractible manifold.*

Proof: There are at least three proofs of the contractibility of $\mathcal{J}_\tau(V,\omega)$, all found in McDuff and Salamon [76, p. 65-67]. The first is due to Gromov [45], and uses fibrations, the second and third are due to Sévennec and are found first in Audin [12] and later in McDuff and Salamon, with the second and third being quite elementary and geometric. The third proof is also geometric, making use of Lagrangian subspaces, but we will follow the second of these proofs, which has the benefit of proving also the calibrated case in one extra step and also giving a manifold structure to both spaces.

(1) By Proposition 32 it is enough to consider $(\mathbb{R}^{2n},\omega_0)$. Let $Y = \{A \in M_{2n}(\mathbb{R}) \,|\, J_0A + AJ_0 = 0\}$, which is a real subspace of $M_{2n}(\mathbb{R})$, and let $B_1 := \{A \in Y \mid \|A\| < 1\}$ be the open unit ball in Y. Consider the map $F : \mathcal{J}_\tau(\mathbb{R}^{2n},\omega_0) \to B_1$ given by

$$\begin{aligned} F(J) &:= (J+J_0)^{-1}(J-J_0) \\ &= (J_0^{-1}J+I)^{-1}(J_0^{-1}J-I) \end{aligned}$$

First, note that $J + J_0 \in \mathrm{GL}(2n,\mathbb{R})$, since for all nonzero $\mathbf{x} \in \mathbb{R}^{2n}$ we have

$$\omega_0(\mathbf{x},(J+J_0)(\mathbf{x})) = \omega_0(\mathbf{x},J(\mathbf{x})) + \omega_0(\mathbf{x},J_0(\mathbf{x})) > 0$$

since both J and J_0 are ω-tame. Thus, $\ker(J + J_0) = \{\mathbf{0}\}$ and F is well defined. To see that its image lies in B_1, note first that $F(J)$ satisfies $\|F(J)\| < 1$ iff $\|J_0^{-1}J - I\|^2 < \|J_0^{-1}J + I\|^2$, i.e.

$$\| - J_0J\mathbf{x} + \mathbf{x}\|^2 - \| - J_0J\mathbf{x} - \mathbf{x}\|^2 = \|J_0J\mathbf{x} - \mathbf{x}\|^2 - \|J_0J\mathbf{x} + \mathbf{x}\|^2 > 0$$

for all nonzero $\mathbf{x} \in \mathbb{R}^{2n}$. But this expression follows from the ω-tameness of J:

$$\begin{aligned}
\|J_0J\mathbf{x} - \mathbf{x}\|^2 - \|J_0J\mathbf{x} + \mathbf{x}\|^2 &= g_{J_0}(J_0J\mathbf{x} - \mathbf{x}, J_0J\mathbf{x} - \mathbf{x}) - g_{J_0}(J_0J\mathbf{x} + \mathbf{x}, J_0J\mathbf{x} + \mathbf{x}) \\
&= -4g_{J_0}(J_0J\mathbf{x}, \mathbf{x}) \\
&= 4g_{J_0}(J\mathbf{x}, J_0\mathbf{x}) \\
&= 4g_{J_0}(J_0\mathbf{x}, J\mathbf{x}) \\
&= 4\omega_0(\mathbf{x}, J\mathbf{x}) \\
&> 0
\end{aligned}$$

so $\|F(J)\| < 1$. Next, we show the identity $J_0F(J) + F(J)J_0 = 0$. First, note that $(J + J_0)J_0 = JJ_0 + J_0^2 = JJ_0 + J^2 = J(J + J_0)$, so that

$$\begin{aligned}
J_0(J + J_0)^{-1} = -J_0^{-1}(J + J_0)^{-1} = -((J + J_0)J_0)^{-1}& \\
= -(J(J + J_0))^{-1} = -(J + J_0)^{-1}J^{-1} = (J + J_0)^{-1}J&
\end{aligned}$$

Similarly, $J(J - J_0) = J^2 - JJ_0 = J_0^2 - JJ_0 = -(J - J_0)J_0$, and therefore

$$J_0F(J) = J_0(J + J_0)^{-1}(J - J_0) = (J + J_0)^{-1}J(J - J_0) = -(J + J_0)^{-1}(J - J_0)J_0 = -F(J)J_0$$

Thus, F is well defined and $\operatorname{im} F \subseteq B_1$.

(2) In fact, we claim $\operatorname{im} F = B_1$ and F is invertible, because we can construct an inverse $G : B_1 \to \mathcal{J}_\tau(\mathbb{R}^{2n}, \omega_0)$ for F, as follows. If $S \in B_1$, then $I - S \in \mathrm{GL}(2n, \mathbb{R})$, so we can define G by

$$G(S) := J = J_0(I + S)(I - S)^{-1}$$

To see that $\operatorname{im} G \subseteq \mathcal{J}_\tau(\mathbb{R}^{2n}, \omega_0)$, note that the condition $J_0 S + S J_0 = 0$ implies

$$\begin{aligned}
J^2 &= J_0(I+S)(I-S)^{-1}J_0(I+S)(I-S)^{-1} \\
&= (J_0 + J_0 S)(I-S)^{-1}(J_0 + J_0 S)(I-S)^{-1} \\
&= (J_0 - S J_0)(I-S)^{-1}(J_0 - S J_0)(I-S)^{-1} \\
&= (I-S)J_0(I-S)^{-1}(I-S)J_0(I-S)^{-1} \\
&= (I-S)J_0^2(I-S)^{-1} \\
&= -I
\end{aligned}$$

Moreover J is ω-tame, which again follows from the polarization identity: writing $B = (I+S) \circ (I-S)^{-1}$ we note that $J = J_0 B$, and J is ω-tame iff B is positive (not necessarily symmetric, however), for $g_{J_0}(\mathbf{u}, B\mathbf{u}) = \omega_0(\mathbf{u}, J\mathbf{u})$. Thus, it is enough to show that B is positive. Towards this end, observe first that

$$B + I = 2(I-S)^{-1}$$

$$B - I = 2S(I-S)^{-1}$$

since $B + I = \big[(I+S)(I-S)^{-1} + I\big] = \big[(I+S) + (I-S)\big](I-S)^{-1} = 2(I-S)^{-1}$, and similarly $B - I = \big[(I+S)(I-S)^{-1} - I\big] = \big[(I+S) - (I-S)\big](I-S)^{-1} = 2S(I-S)^{-1}$. Since by assumption $\|S\| < 1$, applying the polar identity we get for all nonzero $\mathbf{u} \in \mathbb{R}^{2n}$

$$\begin{aligned}
g_{J_0}(\mathbf{u}, B\mathbf{u}) &= \frac{1}{4}\|\mathbf{u} + B\mathbf{u}\|^2 - \frac{1}{4}\|\mathbf{u} - B\mathbf{u}\|^2 \\
&= \frac{1}{4}\|(B+I)\mathbf{u}\|^2 - \frac{1}{4}\|(B-I)\mathbf{u}\|^2 \\
&= \|(I-S)^{-1}\mathbf{u}\|^2 - \|S(I-S)^{-1}\mathbf{u}\|^2 \\
&\geq \|(I-S)^{-1}\mathbf{u}\|^2(1 - \|S\|^2) \\
&> 0
\end{aligned}$$

Lastly, note that $G \circ \frac{1}{4}I$ is the inverse of F:

$$\begin{aligned} G\Big(\frac{1}{4}F(J)\Big) &= G\Big(\frac{1}{4}(J+J_0)^{-1}(J-J_0)\Big) \\ &= G\Big(\frac{1}{4}(B+I)^{-1}J_0^{-1}J_0(B-I)\Big) \\ &= G\Big(\frac{1}{4}2(I-S)2S(I-S)^{-1}\Big) \\ &= G(S) \\ &= J \end{aligned}$$

and similarly

$$\begin{aligned} F\Big(G\Big(\frac{1}{4}S\Big)\Big) &= F\Big(\frac{1}{4}J_0B\Big) \\ &= \Big(\frac{1}{4}\big(J_0(B+I)\big)^{-1}J_0(B-I)\Big) \\ &= (I-S)J_0^{-1}J_0S(I-S)^{-1} \\ &= S \end{aligned}$$

Since F and G are smooth maps, this shows that $\mathcal{J}_\tau(\mathbb{R}^{2n}, \omega_0)$ is diffeomorphic to the open ball B_1 in the vector space Y, which is a convex and therefore contractible set. Thus, $\mathcal{J}_\tau(\mathbb{R}^{2n}, \omega_0)$ is contractible.

(3) To see that $\mathcal{J}_c(\mathbb{R}^{2n}, \omega_0)$ is contractible as well, the same constructions of B_1, F and G apply, with only one modification to the definition of Y, namely we let $Y = \{S \in \mathrm{Sym}_{2n}(\mathbb{R}) \mid SJ_0 + J_0S = 0\}$. For if $S = S^T$ as well, then $(I \pm S)^T = I \pm S$, and it is enough to show that $J^TJ_0J = J_0$ for $J = J_0B = J_0(I+S)(I-S)^{-1}$. But,

$$\begin{aligned} J^TJ_0J &= (J_0B)^TJ_0(J_0B) \\ &= B^TJ_0B \\ &= \big((I+S)(I-S)^{-1}\big)^TJ_0(I+S)(I-S)^{-1} \\ &= (I-S)^{-1}(I+S)(I-S)(I+S)^{-1}J_0 \\ &= (I-S)^{-1}(I-S)(I+S)(I+S)^{-1}J_0 \\ &= J_0 \end{aligned}$$

Conversely, if $J^T J_0 J = J_0$, then $B^T J_0 B = J_0$, so

$$\begin{aligned} J_0 &= B^T J_0 B \\ &= \left((I+S)(I-S)^{-1}\right)^T J_0 (I+S)(I-S)^{-1} \\ &= (I-S^T)^{-1}(I+S^T)(I-S)(I+S)^{-1} J_0 \end{aligned}$$

which shows that $I = (I-S^T)^{-1}(I+S^T)(I-S)(I+S)^{-1}$. Therefore,

$$(I-S^T)(I+S) = (I+S^T)(I-S)$$

and this simplifies to $S = S^T$. This shows that $J \in \mathrm{Sp}(n,\mathbb{R})$ iff $S = S^T$. Finally, note that, since $J^T J_0 J = J_0$, or equivalently $J^T = J_0 J J_0$, we have

$$\begin{aligned} F(J)^T &= \left[(J+J_0)^{-1}(J-J_0)\right]^T \\ &= (J-J_0)^T((J+J_0)^T)^{-1} \\ &= (J^T+J_0)(J^T-J_0)^{-1} \\ &= (J_0 J J_0 + J_0)(J_0 J J_0 - J_0)^{-1} \\ &= (J_0 J + I)(J_0 J - I)^{-1} \\ &= (J_0 J - J^2)(J_0 J + J^2)^{-1} \\ &= (J_0 - J)(J_0 + J)^{-1} \\ &= (J+J_0)^{-1}(J-J_0) \\ &= F(J) \end{aligned}$$

so $\operatorname{im} F = B_1$. The penultimate equality is straightforward: it holds iff $(J+J_0)(J_0-J) = (J-J_0)(J+J_0)$, which is easily seen to hold.

(4) Finally, we note that a matrix $S \in M_{2n}(\mathbb{R})$ satisfies $SJ_0 + J_0 S = 0$ iff, after writing S in block matrix form as

$$S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

we have

$$0 = SJ_0 + J_0S = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} B - C & -A - D \\ A + D & B - C \end{pmatrix}$$

which holds iff $A = -D$ and $B = C$, and therefore iff

$$S = \begin{pmatrix} A & B \\ B & -A \end{pmatrix}$$

where $A, B \in M_n(\mathbb{R})$. Thus, the dimension of the vector space of such S is $2n^2$. If we additionally require S to be symmetric, then we must also require A and B to be symmetric, and the dimension of such matrices is $2n(n+1)/2 = n(n+1)$. ■

Remark 22 Since $\mathcal{J}(V, \omega) = \mathcal{J}_\tau(V, \omega) \cap \mathcal{J}_c(V, \omega)$, the last proposition offers another proof of the contractibility of $\mathcal{J}(V, \omega)$, which we will prove directly below. ■

2.3.5.8 The Space of ω-Compatible Structures $\mathcal{J}(V, \omega)$

We recall that any symplectic space (V, ω) can be endowed with a metric g and an ω-compatible complex structure J. The space of such complex structures is, moreover, path connected, and in fact any two ω-compatible complex structures are *smoothly* homotopic:

Proposition 37 *Every symplectic space (V, ω) can be given a compatible complex structure J and a metric g, that is $\mathcal{J}(V, \omega)$ is nonempty. Moreover, any two such structures J_0 and J_1 are smoothly homotopic in the following sense: there is a map $\varphi : [0, 1] \times \mathcal{J}(V, \omega) \to \mathcal{J}(V, \omega)$, which we denote J_t, and which is differentiable in t and forms a path from J_0 to J_1 in $\mathcal{J}(V, \omega)$.*

Proof: We have already seen (Theorem 28) that given $\gamma \in \mathcal{M}(V)$ (which is nonempty) and ω we can find a $J_\gamma \in \mathcal{J}(V, \omega)$, and (Theorem 29) given $J \in \mathcal{J}(V, \omega)$ we get a $g_J \in \mathcal{M}(V)$ by $g_J = \omega \circ (\mathrm{id}_V \times J)$. Finally, if $J_0, J_1 \in \mathcal{J}(V, \omega)$, find g_0 and g_1 in $\mathcal{M}(V)$ such that $J_i = J_{g_i}$ and define

$$g_t := tg_1 + (1 - t)g_0$$

Then $g_t \in \mathcal{M}(V)$, since $\mathcal{M}(V)$ is a convex cone in $\mathrm{Hom}^2_{\mathrm{Sym}}(\mathbb{R}^{2n}; \mathbb{R})$, and $J_t := J_{g_t} \in \mathcal{J}(V, \omega)$. ■

Proposition 38 *Let (V,ω) be any symplectic vector space. Then the space of ω-compatible complex structures $\mathcal{J}(V,\omega)$ is diffeomorphic to the space of symplectic positive definite forms (or their matrix equivalents),*[4]

$$\mathcal{J}(V,\omega) \approx \mathrm{Sp}(V,\omega) \cap \mathrm{Hom}^2_{\mathbb{R},\mathrm{Sym}}(V;\mathbb{R})^+ \cong \mathrm{Sp}(n,\mathbb{R}) \cap \mathrm{Sym}_{2n}(\mathbb{R})^+ \tag{2.74}$$

Proof: By Proposition 32 it will suffice to consider the case of $(\mathbb{R}^{2n},\omega_0,J_0)$. In this case, J belongs to $\mathcal{J}(\mathbb{R}^{2n},\omega_0)$ iff $J^2 = -I_{2n}$, $J \in \mathrm{Sp}(n,\mathbb{R})$ and $g_J \in \mathrm{Hom}_{\mathbb{R},\mathrm{Sym}}(\mathbb{R}^{2n};\mathbb{R})^+$. By Propositions 19 and 18, these conditions hold true iff the following three conditions hold:

(1) $J^2 = -I_{2n}$

(2) $J^T J_0 J = J_0$

(3) $g_{J_0}(v,(-J_0J)(v)) = g_{J_0}(J_0(v),J(v)) = \omega_0(v,J(v)) = g_J(v,v) > 0,\ \forall v \neq 0$

The matrix $P := -J_0J$ in the last condition is symmetric, for by condition (2) and the fact that $J_0^T = -J_0$ and $J^T = -J$ we have

$$P^T = (-J_0J)^T = -J^TJ_0^T = J^TJ_0^TJ^2 = -(J^TJ_0J)J = -J_0J = P$$

and, by condition (3) P is positive definite. P is also symplectic, because J and J_0 are and because $\mathrm{Sp}(n,\mathbb{R})$ is a group. This defines a map

$$\varphi : \mathcal{J}(\mathbb{R}^{2n},\omega_0) \to \mathrm{Sp}(n,\mathbb{R}) \cap \mathrm{Sym}_{2n}(\mathbb{R})^+$$

$$\varphi(J) := -J_0J$$

This map is injective, for if $-J_0J = -J_0J'$, the invertibility of J_0 implies $J = J'$. It is also surjective, for if $P \in \mathrm{Sp}(n,\mathbb{R}) \cap \mathrm{Sym}_{2n}(\mathbb{R})^+$, then define $J := -J_0^{-1}P = J_0P$ and notice that $J^2 = J_0PJ_0P = J_0(P^TJ_0P) = J_0^2 = -I$ since P is symmetric and symplectic. Moreover, J is symplectic (i.e. calibrates ω), for both J_0 and P are and $\mathrm{Sp}(n,\mathbb{R})$ is a group. Finally, φ is smooth, since it is left-multiplication-by-$-J_0$, and it's inverse, $\varphi^{-1}(P) = J_0P$, is smooth for the same reason.

■

[4] We used the notation of Section 2.1.5 here.

Remark 23 In view of Proposition 36 and the fact that $\mathcal{J}(V,\omega)$ is precisely the intersection of $\mathcal{J}_\tau(V,\omega)$ with $\mathcal{J}_c(V,\omega)$, we can see that the diffeomorphism above can be understood in terms of the diffeomorphisms constructed in the proof of that proposition. There, $\mathcal{J}_c(\mathbb{R}^{2n})$ was mapped diffeomorphically onto the open unit ball B_1 of

$$\{S \in \mathrm{Sym}_{2n}(\mathbb{R}) \mid SJ_0 + J_0S = 0\} = \mathfrak{sp}(n,\mathbb{R}) \cap \mathrm{Sym}_{2n}(\mathbb{R})$$

via an analogue of the Cayley transform. The diffeomorphism $G : B_1 \to \mathcal{J}_c(\mathbb{R}^{2n})$ was given by $J := G(S) = J_0B$, where $B = (I+S)(I-S)^{-1}$. For then we see that $B = -J_0J$, which is precisely the positive symplectic matrix we called P in the proposition above.

To see this from another angle, note that the (necessary and sufficient) condition on B in the case that $S \in \mathfrak{sp}(n,\mathbb{R}) \cap \mathrm{Sym}_{2n}(\mathbb{R})$ was that $B \in \mathrm{Sp}(n,\mathbb{R})$, while the (necessary and sufficient) condition on B in the case that S satisfes only $SJ_0 + J_0S = 0$ was that B be positive (but not necessarily symmetric). However, if we add in the symmetry condition on S, then $B = B^T$, too, for then

$$B^T = [(I+S)(I-S)^{-1}]^T = (I-S)^{-1}(I+S) = (I+S)(I-S)^{-1} = B$$

the penultimate equality following from $(I+S)(I-S) = I - S^2 = (I-S)(I+S)$. Thus, we have $B = -J_0J \in \mathrm{Sym}_{2n}(\mathbb{R})^+ \cap \mathrm{Sp}(n,\mathbb{R})$. And since sending J to B is an invertible operation with inverse $B \mapsto J_0B = J$, diffeomorphically so, we have the following diffeomorphisms:

$$\Big(\mathfrak{sp}(n,\mathbb{R}) \cap \mathrm{Sym}_{2n}(\mathbb{R})\Big)_1 \approx \mathcal{J}(V,\omega) \approx \mathrm{Sp}(n,\mathbb{R}) \cap \mathrm{Sym}_{2n}(\mathbb{R}) \tag{2.75}$$

By (2.52) of Proposition 20 we know that $S = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathfrak{sp}(n,\mathbb{R}) \cap \mathrm{Sym}_{2n}(\mathbb{R})$ iff $A = -D, B = C \in \mathrm{Sym}_n(\mathbb{R})$, so the dimension of the manifold $\mathcal{J}(V,\omega)$ is

$$\dim \mathcal{J}(V,\omega) = n(n+1) \qquad \blacksquare$$

We end this section with a different and quite simple proof of the contractibility of $\mathcal{J}(V,\omega)$.

Proposition 39 (Contractibility of $\mathcal{J}(V,\omega)$) *There is a continuous map*

$$r : \mathcal{M}(V) \to \mathcal{J}(V,\omega)$$
$$r(g_J) := J$$

for all $J \in \mathcal{J}(V,\omega)$ and $g \in \mathcal{M}(V)$. Using this map r we can construct a homotopy between the identity map on $\mathcal{J}(V,\omega)$ to the constant map $c(J) = J_0$, thus demonstrating that the space $\mathcal{J}(V,\omega)$ of ω-compatible complex structures is contractible.

Proof: We follow the proof found in McDuff and Salamon [76, Proposition 2.50], with only minor modifications. By Proposition 32 it will suffice to consider the case of $(\mathbb{R}^{2n}, \omega_0, J_0)$. To see that the map $r : \mathcal{M}(\mathbb{R}^{2n}) \to \mathcal{J}(\mathbb{R}^{2n}, \omega_0)$ is continuous, let $g \in \mathcal{M}(\mathbb{R}^{2n})$ and use Proposition 28 to find a g-skew-adjoint $J \in \mathcal{J}(\mathbb{R}^{2n}, \omega_0)$ such that $\omega_0(\cdot,\cdot) = g(J(\cdot),\cdot)$. If $G = [g]_\rho$ is the matrix representation of g with respect to the standard basis ρ for $\mathbb{R}^{2n}$, then by Proposition 1 we will have for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{2n}$

$$-\mathbf{u}^T GJ\mathbf{v} = g(\mathbf{u}, -J(\mathbf{v})) = g(J(\mathbf{u}), \mathbf{v}) = \omega_0(\mathbf{u}, \mathbf{v}) = g_{J_0}(J_0(\mathbf{u}), \mathbf{v}) = -\mathbf{u}^T J_0 \mathbf{v}$$

which shows that $GJ = J_0$, and consequently $J = G^{-1} J_0$. Therefore, since $J_0 \in \mathrm{O}(2n)$ by Proposition 18, we will have, for any two metrics $g, g' \in \mathcal{M}(\mathbb{R}^{2n})$, that

$$\|J - J'\| = \|G^{-1} J_0 - G'^{-1} J_0\| = \|G^{-1} - G'^{-1}\|$$

Since the inversion map $\mathrm{GL}(2n,\mathbb{R}) \to \mathrm{GL}(2n,\mathbb{R})$, $G \mapsto G^{-1}$, is continuous (it is in fact smooth in $\mathrm{GL}(B)$ for any Banach space B, see for example Rudin [93, Corollaries 1 and 2, p. 353]), we have that the map r, which is the composition of the inversion map with the isometry J_0 and the isometry identifying g with G, is also continuous.

Using r we can define the homotopy

$$H : \mathcal{J}(\mathbb{R}^{2n}, \omega_0) \times I \to \mathcal{J}(\mathbb{R}^{2n}, \omega_0)$$
$$H(J,t) := r\big((1-t)g_{J_0} + t g_J\big)$$

Then clearly $H(J,0) = J_0 = c(J)$ and $H(J,1) = J = \mathrm{id}(J)$, so $H : \mathrm{id} \simeq c$ is the required deformation retraction. ∎

Recall from Proposition 32 that any symplectic vector space (V,ω) is linearly symplectomorphic to $(\mathbb{R}^{2n},\omega_0)$ via a linear isomorphism $\Phi:\mathbb{R}^{2n}\to V$ satisfying $J\circ\Phi=\Phi\circ J_0$, or $J=\Phi^*J_0:=\Phi\circ J_0\circ\Phi^{-1}$. From the construction of Φ in that proposition we see its dependence on $J\in\mathcal{J}(V)$, which was used to find the symplectic basis β for V in terms of which defined Φ. This justifies the notation Φ_J. If we consider $(V,\omega)=(\mathbb{R}^{2n},\omega_0)$, then $\Phi_J\in\mathrm{Sp}(n,\mathbb{R})$, and this shows that the $\mathrm{GL}(2n,\mathbb{R})$-action (by conjugation) of Proposition 35 restricts to a transitive $\mathrm{Sp}(n,\mathbb{R})$-action on $\mathcal{J}(\mathbb{R}^{2n})$. Let us consider the $\mathrm{Sp}(n,\mathbb{R})$-action on the subspace $\mathcal{J}(\mathbb{R}^{2n},\omega_0)$ of ω_0-*compatible* complex structures. We claim that *this* action is transitive, and in particular, for all $J\in\mathcal{J}(\mathbb{R}^{2n},\omega_0)$ we have $J=\Phi^*J_0\in\mathcal{J}(V,\omega)$: recall Lemma 2 and Proposition 27, which tell us that $\Phi\Phi^T,\Phi^T\Phi\in\mathrm{Sp}(n,\mathbb{R})\cap\mathrm{Sym}_{2n}(\mathbb{R})^+$, and Proposition 19, which says that $\Phi J_0=J_0(\Phi^T)^{-1}$ and $J_0\Phi^{-1}=\Phi^TJ_0$. Then,

$$
\begin{aligned}
\omega_0(u,Jv)=\omega_0(u,\Phi^*J_0v)&=\omega_0(u,\Phi J_0\Phi^{-1}v)\\
&=\omega_0(u,\Phi\Phi^TJ_0v)=g_{J_0}(J_0u,\Phi\Phi^TJ_0v)=(J_0u)^T\Phi\Phi^TJ_0v\\
&=-u^TJ_0\Phi\Phi^TJ_0v=-u^TJ_0^2(\Phi\Phi^T)^{-1}v=u^T(\Phi\Phi^T)^{-1}v
\end{aligned}
$$

Since $\Phi\Phi^T\in\mathrm{Sp}(n,\mathbb{R})\cap\mathrm{Sym}_{2n}(\mathbb{R})^+$, we also have $(\Phi\Phi^T)^{-1}\in\mathrm{Sp}(n,\mathbb{R})\cap\mathrm{Sym}_{2n}(\mathbb{R})^+$, and the expression $u^T(\Phi\Phi^T)^{-1}v$ defines a metric $g_J(u,v)$ by Theorems 1 and 7 whose matrix representation is precisely $(\Phi\Phi^T)^{-1}$,

$$
g_J(u,v)=u^T(\Phi\Phi^T)^{-1}v
$$

This shows that $g_J=\omega_0\circ I\times J\in\mathcal{M}(V)$, and therefore that $J\in\mathcal{J}(\mathbb{R}^{2n},\omega_0)$. Now consider the stabilizer of J_0 under the $\mathrm{Sp}(n,\mathbb{R})$-action: it consists of $A\in\mathrm{Sp}(n,\mathbb{R})$ such that $AJ_0A^{-1}=J_0$, which is precisely the complex matrices in $\mathrm{Sp}(n,\mathbb{R})$ by Corollary 3. But $\mathrm{Sp}(n,\mathbb{R})\cap\mathrm{GL}(n,\mathbb{C})=\mathrm{U}(n)$ by Proposition 25. Thus, $\mathrm{Sp}(n,\mathbb{R})_{J_0}=\mathrm{U}(n)$. By the orbit-stabilizer theorem for Lie groups, we therefore have a diffeomorphism,

$$
\mathcal{J}(\mathbb{R}^{2n},\omega_0)\approx\mathrm{Sp}(n,\mathbb{R})/\,\mathrm{U}(n)
$$

and the map

$$F : \mathcal{J}(\mathbb{R}^{2n}, \omega_0) \to \mathrm{Sp}(n, \mathbb{R}) / \mathrm{U}(n)$$

$$F(J) := \Phi_J \cdot \mathrm{U}(n)$$

is the very diffeomorphism of the orbit-stabilizer theorem.

Now consider (V, ω), and let $\mathrm{Sp}(V, \omega)$ act on $\mathcal{J}(V, \omega)$ by conjugation. Under the linear symplectomorphism $(V, \omega) \cong (\mathbb{R}^{2n}, \omega_0)$ we may transfer the whole argument above over to (V, ω), and conclude that

$$\mathcal{J}(V, \omega) \approx \mathrm{Sp}(V, \omega) / \mathrm{U}(V, H_J) \cong \mathrm{Sp}(n, \mathbb{R}) / \mathrm{U}(n)$$

(in the notation of Proposition 31).

We summraize these results in the next proposition:

Proposition 40 *The symplectic group acts transitively on $\mathcal{J}(\mathbb{R}^{2n}, \omega_0)$ by conjugation, and the stabilizer subgroup of J_0 under this action is the unitary group. Consequently, we have the diffeomorphism*

$$\mathcal{J}(\mathbb{R}^{2n}, \omega_0) \approx \mathrm{Sp}(n, \mathbb{R}) / \mathrm{U}(n) \tag{2.76}$$

which is given explicitly by the map

$$F : \mathcal{J}(\mathbb{R}^{2n}, \omega_0) \to \mathrm{Sp}(n, \mathbb{R}) / \mathrm{U}(n) \tag{2.77}$$

$$F(J) := \Phi_J \cdot \mathrm{U}(n) \tag{2.78}$$

where Φ_J was constructed in the proof of Proposition 32. More generally, if (V, ω) is any symplectic vector space, we have

$$\mathcal{J}(V, \omega) \approx \mathrm{Sp}(V, \omega) / \mathrm{U}(V, H_J) \cong \mathrm{Sp}(n, \mathbb{R}) / \mathrm{U}(n) \tag{2.79}$$

using the notation of Proposition 31. This shows again that $\mathcal{J}(\mathbb{R}^{2n}, \omega_0)$ is a smooth manifold and homogeneous space of dimension $\dim \mathrm{Sp}(n, \mathbb{R}) - \dim \mathrm{U}(n) = (2n^2 + n) - n^2 = n(n+1)$. ■

Remark 24 This provides yet another proof of the contractibility of $\mathcal{J}(V,\omega)$, namely via the contractibility of $\mathrm{Sp}(n,\mathbb{R})/\,\mathrm{U}(n)$, Proposition 27 above. ■

We end this section by asking a related question: instead of starting with ω and looking for J which is compatible with ω, what if we started with J and asked for ω which is compatible with J?

Proposition 41 *Let V be a $2n$-dimensional real vector space and let $J \in \mathcal{J}(V)$. Define $\Omega(V,J)$ to be the space of symplectic forms on V compatible with J, meaning that $\omega(\cdot, J\cdot) \in \mathcal{M}(V) = \mathrm{Hom}^2_{\mathbb{R},\mathrm{Sym}}(V;\mathbb{R})^+$ is a metric. Then $\Omega(V,J)$ is a nonempty convex subset of $\Omega(V)$, and therefore contractible.*

Proof: Firstly, $\Omega(V,J)$ is nonempty, because given J, then as in the proof of Proposition 32 we may take a basis $\beta = (u_1 + iv_1, \dots, u_n + iv_n)$ for $V^{1,0} = \ker(J^{\mathbb{C}} - iI)$ and note that $v_j = Ju_j$, which gives a basis $\beta = (u_1, \dots, u_n, v_1, \dots, v_n)$ for V such that $Ju_j = v_j$ and $Jv_j = -u_j$, that is $[J]_\beta = J_0$. Define ω by

$$\omega(x,y) := -[x]_\beta^T [J]_\beta [y]_\beta = -[x]_\beta^T J_0 [y]_\beta$$

Then ω is skew-symmetric, since $J_0^T = -J_0$:

$$\omega(y,x) = -[y]_\beta J_0 [x]_\beta = -[x]_\beta^T J_0^T [y]_\beta = [x]_\beta J_0 [y]_\beta = -\omega(x,y)$$

and it is nondegenerate because J_0 is its matrix representation with respect to β (Theorems 1 and 10). To see the compatibility of ω with J, note that

$$g_J(x,y) = \omega(x,Jy) = -[x]_\beta^T J_0 [Jy]_\beta = -[x]_\beta J_0^2 [y]_\beta = [x]_\beta \cdot [y]_\beta$$

which is symmetric and positive definite. Secondly, (Ω, J) is convex, let $\omega_0, \omega_1 \in \Omega(V,J)$, then for all $t \in [0,1]$ we have that $\omega_t := (1-t)\omega_0 + t\omega_1$ is also skew-symmetric, because $\mathrm{Hom}^2_{\mathbb{R},\mathrm{Skew}}(V;\mathbb{R})$ is a vector space, and $\omega_t(\cdot, J\cdot)$ is symmetric and positive definite because $\omega_0(\cdot, J\cdot)$ and $\omega_1(\cdot, J\cdot)$ are. Finally, ω_t is nondegenerate because $\omega_t(\cdot, J\cdot)$ is. ■

2.3.6 Sums and Subspaces of Symplectic Vector Spaces

In this section we look at direct sums and subspaces of symplectic vector spaces. The most important type of subspace in this setting is the Lagrangian subspace, and we identify the main properties of such subspaces and describe the manifold and homogeneous space structure of the Lagrangian Grassmannian, the subspace $\mathcal{L}(V,\omega)$ of $\mathrm{G}(n,V)$ consisting of all Lagrangian subspaces of (V,ω).

2.3.6.1 Direct Sums of Symplectic Vector Spaces

The direct sum of two symplectic vector spaces (V_1,ω_1) and (V_2,ω_2) can be made into a symplectic vector space,

$$(V_1 \oplus V_2, \omega_1 \oplus \omega_2) \tag{2.80}$$

by defining $\omega_1 \oplus \omega_2$ as

$$\begin{aligned}
&\omega_1 \oplus \omega_2 : (V_1 \oplus V_2) \times (V_1 \oplus V_2) \to \mathbb{R} \\
&\omega_1 \oplus \omega_2\big((v_1,v_2),\ (v_1',v_2')\big) := \omega_1(v_1,v_1') + \omega_2(v_2,v_2')
\end{aligned}$$

Then, $\omega_1 \oplus \omega_2$ is a symplectic form on $V_1 \oplus V_2$, for bilinearity and skew-symmetry follow from those of ω_1 and ω_2, and nondegeneracy follows from the observation that, if $\omega_1 \oplus \omega_2\big((v_1,v_1'),\ (v_2',v_2)\big) = 0$ for all $(v_1,v_2) \in V_1 \oplus V_2$, then $\omega_1(v_1,v_1') + \omega_2(v_2,v_2') = 0$, and therefore (by choosing $v_1 = 0$ or $v_2 = 0$) $\omega_1(v_1,v_1') = \omega_2(v_2,v_2') = 0$ for all $v_1 \in V_1$ and $v_2 \in V_2$, and the nondegeneracy of ω_1 and ω_2 then implies that $v_1 = v_2 = 0$, or $(v_1,v_2) = (0,0)$.

We can also introduce the symplectic form

$$\begin{aligned}
&\omega_1 \ominus \omega_2 : (V_1 \oplus V_2) \times (V_1 \oplus V_2) \to \mathbb{R} \\
&\omega_1 \ominus \omega_2\big((v_1,v_2),\ (v_1',v_2')\big) := \omega_1(v_1,v_1') - \omega_2(v_2,v_2')
\end{aligned}$$

which is nondegenerate by the same argument, making

$$(V_1 \oplus V_2, \omega_1 \ominus \omega_2) \tag{2.81}$$

a symplectic vector space. Of course, the symplectic forms $\omega_1 \oplus \omega_2$ and $\omega_1 \ominus \omega_2$ both belong to $\Omega(V_1 \oplus V_2)$, and as we saw in Proposition 34 above, $\mathrm{GL}(V_1 \oplus V_2) \cong \mathrm{GL}(4n, \mathbb{R})$ acts transitively on $\Omega(V_1 \oplus V_2)$ by pulling back, $g \cdot \omega_1 \oplus \omega_2 = (g^{-1})^* \omega_1 \oplus \omega_2$, so in particular there is a $g \in \mathrm{GL}(V_1 \oplus V_2)$ such that $g \cdot \omega_1 \oplus \omega_2 = \omega_1 \ominus \omega_2$, namely, with respect to symplectic bases β_i for V_i,

$$g = I_{2n} \oplus \begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix}$$

Indeed, if $\beta_i = (q_i, p_i)$ is a symplectic basis for V_i, then (q_1, p_1, q_2, p_2) is a symplectic basis for $V_1 \oplus V_2$, and applying g to this basis interchanges the last two pieces, which has the effect of using the skew-symmetry of ω_2 to put a minus sign in front of it.

2.3.6.2 Subspaces of Symplectic Vector Spaces

Recall the terminology of Section 2.1.2 on orthogonality in a bilinear space. We specialize here to the case of a symplectic vector space (V, ω), and follow McDuff and Salamon [76] in our terminology and notation for this specialization.

Since ω is skew-symmetric, orthogonality is a symmetric relation on V, that is $u \perp v$ (i.e. $\omega(u, v) = 0$) iff $v \perp u$ (i.e. $\omega(v, u) = 0$) for all $u, v \in V$, and therefore ${}^{\perp}W = W^{\perp}$ for any subspace W of V.

Definition 6 We introduce the notation

$$W^{\omega} \tag{2.82}$$

for $W^{\perp}$, the orthogonal complement of W in V with respect to ω, which is called the **symplectic complement** of W in V. We remark that it need not be the case that $W \cap W^{\omega} = \{0\}$, for it can happen that $W \subseteq W^{\omega}$ or $W \supseteq W^{\omega}$.

If $W \subseteq W^{\omega}$, we say that W is **isotropic**. If $W \supseteq W^{\omega}$, we say W is **coisotropic**. W is called **Lagrangian** if $W = W^{\omega}$, i.e. if it is both isotropic and coisotropic. If W is neither isotropic nor coisotropic, i.e. if $W \cap W^{\omega} = \{0\}$, then it is called **symplectic**.

Example 8 Let us give some examples of each of these types of subspace of $(\mathbb{R}^{2n}, \omega_0)$.

(1) Let $(q_1, q_2, q_3, p_1, p_2, p_3) = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4, \mathbf{e}_5, \mathbf{e}_6)$ be the standard symplectic basis for $\mathbb{R}^6$, and consider the subspace $W = \text{span}_{\mathbb{R}}(q_1, q_2)$. Then, the relations (2.35) ensure that

$$\omega_0(q_1, q_1) = \omega_0(q_1, q_2) = \omega_0(q_2, q_2) = 0$$

so that $\omega_0(u, v) = 0$ for all $u, v \in W$, whence $W \subseteq W^\omega$, making W an isotropic subspace. In fact, the inclusion is strict, $W \subsetneq W^\omega$, since $q_3 \in W^\omega \backslash W$.

(2) The subspace $W = \text{span}_{\mathbb{R}}(q_1, q_2, p_1)$ of $\mathbb{R}^4$ is coisotropic. For if $v \in W^\omega$, then $v \perp p_1$, $v \perp p_2$ and $v \perp p_1$. Writing $v = aq_1 + bq_2 + cp_1 + dp_2$ and recalling the relations (2.35), we see that

$$\omega_0(v, q_1) = c, \quad \omega_0(v, q_2) = d, \quad \omega_0(v, p_1) = a$$

which must all equal 0 if $v \in W^\omega$. Therefore $v = bq_2 \in W$, which shows that $W^\omega = \text{span}_{\mathbb{R}}(q_2) \subseteq W$, making W a coisotropic subspace.

(3) The subspace $W = \text{span}_{\mathbb{R}}(q_1, p_1)$ of $\mathbb{R}^4$ is symplectic, again by the relations (2.35), which imply $W^\omega = \text{span}_{\mathbb{R}}(q_2, p_2)$.

(4) The subspace $W = \text{span}_{\mathbb{R}}(q_1, q_2)$ of $\mathbb{R}^4$ is Lagrangian, by the relations (2.35). ■

Proposition 42 *A subspace W of V is symplectic iff the restriction of ω to $W \times W$ is nondegenerate.*

Proof: This is just a special application of the general case for bilinear forms, Proposition 3. ■

Proposition 43 *For any subspace W of V, we have*

(1) $\dim W + \dim W^\omega = \dim V$

(2) $W^{\omega\omega} = W$

Proof: (1) Recall Remark 3, which, for a nondegenerate bilinear form B on a finite-dimensional vector space V gives an isomorphism $B^\flat : V \to V^*$ by $B^\flat(v) := B(\cdot\, , v)$. In our case, the flat operator $\omega^\flat := \omega_L$ identifies W^ω with the annihilator W^0 of W in V^*, as follows: We note first that $V/W \cong W^0$, since any $f \in V^*$ factors through V/W iff $W \subseteq \ker f$. Therefore, since $\omega^\flat$ is an isomorphism between V and V^* with inverse $\omega^\sharp$, and since $W^0 \subseteq V^*$, we have

$$W^\omega = \omega^\sharp(W^0)$$

Consequently, $\dim W^\omega = \dim \omega^\sharp(W^0) = \dim W^0 = \dim(V/W) = \dim V - \dim W$. (2) follows from (1), since clearly $W \subseteq W^{\omega\omega}$ and $\dim W = \dim V - \dim W^\omega = \dim W^{\omega\omega}$. ■

Corollary 7 *Every Lagrangian subspace W of V has dimension*

$$\dim W = \dim W^\omega = \frac{\dim V}{2}$$ ■

Proposition 44 *Let L be a subspace of (V, ω). Then the following are equivalent:*

(1) $L \in \mathcal{L}(V, \omega)$.

(2) *L is a maximal isotropic subspace (it is not properly contained in any isotropic subspace).*

(3) *L is isotropic and* $\dim L = n$.

Proof: (1) $\Longrightarrow$ (2) and (3): If $L \in \mathcal{L}(V, \omega)$, then $L = L^\omega$, so L is isotropic and $\dim L = n$ by the previous corollary. To see that it is maximally isotropic, note that the maximum dimension of an isotropic subspace of V is n, because $\dim L + \dim L^\omega = 2n$ by the previous proposition, and any other isotropic subspace W containing L would also have dimension n, so would have to equal L. (2) $\Longrightarrow$ (3): If L is maximal isotropic, then as we saw $\dim L = n$. (3) $\Longrightarrow$ (1) If L is isotropic and $\dim L = n$, then $L \subseteq L^\omega$ and $\dim L^\omega = 2n - \dim L = n$, so $L = L^\omega$. ■

2.3.6.3 Smooth Manifold and Homogeous Space Structure of $\mathcal{L}(V,\omega)$

Definition 7 We will pay special attention to Lagrangian subspaces, so we introduce the following notation for the set of Lagrangian subspaces of (V,ω) and $(\mathbb{R}^{2n},\omega_0)$, called the **Lagrangian Grassmannian**:

$$\mathcal{L}(V,\omega) := \{L \in \mathrm{G}(n,V) \mid L = L^\omega\} \qquad \text{and} \qquad \mathcal{L}(n) := \mathcal{L}(\mathbb{R}^{2n},\omega_0) \tag{2.83}$$

We single out a special Lagrangian subspace of $\mathbb{R}^{2n}$, the **horizontal Lagrangian**,

$$L_{\mathrm{hor}} := \mathbb{R}^n \oplus \{\mathbf{0}\} \in \mathcal{L}(n) \tag{2.84}$$

As we will see, the orbit of this Lagrangian is the whole Lagrangian Grassmannian $\mathcal{L}(n)$ under the action of the unitary group described below. We will call $L_{\mathrm{vert}} := \{\mathbf{0}\} \oplus \mathbb{R}^n$ the **vertical Lagrangian**. Notice that $L_{\mathrm{hor}} \oplus L_{\mathrm{vert}} = \mathbb{R}^{2n}$. This is an example of a Lagrangian decomposition of $\mathbb{R}^{2n}$. A **Lagrangian decomposition** of any symplectic vector space (V,ω) is a pair of transversal Lagrangian spaces $L_0, L_1 \in \mathcal{L}(n)$, that is, a pair satisfying $V = L_0 \oplus L_1$. ■

Lemma 3 *Fix $J \in \mathcal{J}(V,\omega)$. Then, any $L \in \mathcal{L}(V,\omega)$ admits a Lagrangian complement giving V a Lagrangian decomposition, and one such complement is JL. Moreover, with respect to the induced metric $g_J \in \mathcal{M}(V)$, the direct sum $V = L \oplus JL$ is orthogonal,*

$$V = L \bigoplus\!\!\!\!\!\perp JL$$

Proof: If $v \in JL$, then $v = J(u)$ for some $u \in L$, and $g_J(u,v) = g_J(u,J(u)) = \omega(u,J^2(u)) = -\omega(u,u) = 0$, so $L \perp JL$. Note also that $JL \in \mathcal{L}(V,\omega)$, since $L = L^\omega$ and $J^*\omega = \omega$ imply that $(JL)^\omega = JL$. Explicitly, $v = J(u) \in (JL)^\omega$ iff for all $z = J(w) \in JL$ we have $\omega(v,z) = 0$, which is true iff $u, w \in L$ since $\omega(J(u),J(w)) = \omega(u,w) = 0$ on account of $L = L^\omega$. ■

Remark 25 Note that if H_J is the Hermitian inner product on V_J and $L \in \mathcal{L}(V,\omega)$, then $H_J(L \times L) \subseteq \mathbb{R}$, since if $u, v \in L$, then $H_J(u,v) = g_J(u,v) + i\omega(u,v) = g_J(u,v)$. ■

Proposition 45 *Fix $J \in \mathcal{J}(V,\omega)$. Then, for any pair of Lagrangian subspaces $L_1, L_2 \in \mathcal{L}(V,\omega)$ there exists a $T \in \mathrm{GL}_{\mathbb{C}}(V_J) \cap \mathrm{U}(V_J, H_J)$ such that $L_2 = T(L_1)$.*

Proof: Let $\beta = (b_1, \ldots, b_n)$ be a g_J-orthonormal basis for L_1 and note that since $V = L_1 \bigoplus J L_1^{\omega}$, β is actually a complex basis for V_J, since complex scalar multiplication in V_J is given by $(a+ib)v := av + bJ(v)$. Moreover, since $H_J(L_1 \times L_1) \subseteq \mathbb{R}$, this basis is orthonormal with respect to H_J as well. Similarly, a g_J-orthonormal basis $\gamma = (c_1, \ldots, c_n)$ for L_2 yields a complex basis for V_J, and the map $T : L_1 \to L_2$ given on the basis elements by $T(b_j) := c_j$, $j = 1, \ldots, n$, is a $\mathbb{C}$-linear isomorphism. Moreover, $T \in \mathrm{U}(V_J, H_J)$, since if $u = \sum_j a_j b_j$ and $v = \sum_j \alpha_j b_j$ are vectors in V_J, then the fact that $H_J(b_j, b_k) = \delta_{jk} = H_J(c_j, c_k)$ implies

$$
\begin{aligned}
H_J(Tu, Tv) = \sum_{jk} a_j \alpha_k H_J(T(b_j), T(b_k)) \\
= \sum_{jk} a_j \alpha_k H_J(c_j, c_k) = \sum_{jk} a_j \alpha_k H_J(b_j, b_k) = H_J(u, v)
\end{aligned}
$$

which completes the proof. ■

Corollary 8 *Since $\mathrm{End}_{\mathbb{C}}(V_J) \cap \mathrm{U}(V_J, H_J) = \mathrm{O}(V, g_J) \cap \mathrm{Sp}(V,\omega)$ by Proposition 31, we have that any pair of Lagrangian subspaces $L_1, L_2 \in \mathcal{L}(V,\omega)$ are linearly symplectomorphic. Moreover, by choosing oriented bases β and γ for L_1 and L_2, respectively, the restriction of the linear symplectomorphism $T|_{L_1} : L_1 \to L_2$ may be assumed to be orientation-preserving. Finally, observe that the basis β may be assumed to be adapted to J, so that $\beta \cup J\beta$ is a symplectic basis for V, and therefore we may suppose all the elements V, ω, J, g_J, H_J and L_1 to be put into canonical form.* ■

Take a Lagrangian decomposition (L_0, L_1) of (V,ω) and define the map

$$
\rho_{01} : L_1 \to L_0^* \tag{2.85}
$$

$$
\rho_{01}(v) := v^\flat|_{L_0} = \omega(v, \cdot)|_{L_0} \tag{2.86}
$$

Then ρ_{01} is an isomorphism, because $L_0 \oplus L_1 = V$ and L_1 and L_2 are Lagrangian. Define also the isomorphism, for any Lagrangian subspace $L \in \mathcal{L}(V,\omega)$,

$$\rho_L : V/L \to L^* \tag{2.87}$$

$$\rho_L(v+L) := v^\flat|_L = \omega(v,\cdot)|_L \tag{2.88}$$

For a given Lagrangian decomposition, therefore, we have the following commutative diagram of isomorphisms,

$$\begin{tikzcd}
L_1 \arrow[rr, "\rho_{01}"] \arrow[dr, "\pi|_{L_1}"'] & & L_0^* \\
& V/L_0 \arrow[ur, "\rho_{L_0}"'] &
\end{tikzcd}$$

Remark 26 Start by choosing a basis $\beta = (b_1, \ldots, b_n)$ for L_0, then extend it to a symplectic basis $\beta \cup \gamma = (b_1, \ldots, b_{2n})$ for V. If L_1 is a complementary Lagrangian subspace for L_0, then define $b_{n+j} := -\rho_{01}^{-1}(b_j^*)$ and note that $\gamma = (b_{n+1}, \ldots, b_{2n})$, is a basis for L_1. Consequently, given Lagrangian decompositions (L_0, L_1) and (L_0', L_1') of two symplectic vector spaces (V,ω) and (V',ω'), we can find bases $\beta \cup \gamma$ and $\beta' \cup \gamma'$ for V and V', respectively, such that β and β' are bases for L_0 and L_0', and γ and γ' are bases for L_1 and L_1', respectively. By use of these bases we can take any isomorphism from L_0 to L_0' and extend it to a linear symplectomorphism $T : V \to V'$ in such a way that

$$T(L_i) = L_i', \quad i = 1, 2$$

This shows in particular that every isomorphism of a Lagrangian subspace L extends to a linear symplectomorphism of V. ■

Let us return to our main task of showing that $\mathcal{L}(V,\omega)$ is a submanifold of the Grassmannian manifold $\mathrm{G}(2n, V)$ and a homogeneous space under the action of the unitary and symplectic groups. Recall the definition of the charts for Grassmannians $\mathrm{G}(k, V)$: If $V = W_1 \oplus W_2$ where $W_1 \in \mathrm{G}(k, V)$ and $W_2 \in \mathrm{G}(2n-k, V)$, then

$$U_{W_2} := \{W \in \mathrm{G}(k,V) \mid W \cap W_2 = \{0\}\} \tag{2.89}$$

is the open neighborhood of W_1 serving as the domain of the chart

$$\phi_{12} : U_{W_2} \to \mathrm{Hom}_{\mathbb{R}}(W_1, W_2) \cong \mathbb{R}^{k(2n-k)} \tag{2.90}$$

$$\phi_{12}(W) := T = \pi_2 \circ \pi_1|_W^{-1} \tag{2.91}$$

where $\pi_i : V \to W_i$ is the projection onto W_i, $i = 1, 2$.

Lemma 4 *If (L_0, L_1) is a Lagrangian decomposition of V, then a given n-dimensional subspace $L \in U_{L_1}$ is Lagrangian iff the bilinear form*

$$\rho_{01} \circ \phi_{01}(L) \in \mathrm{Hom}_{\mathbb{R}}(L_0, L_0^*) \cong \mathrm{Hom}^2_{\mathbb{R}}(L_0; \mathbb{R}) \tag{2.92}$$

is symmetric.

Proof: Since $\dim L = n$, we have $L \in \mathcal{L}(V, \omega)$ iff L is isotropic, so letting $T := \phi_{01}(L) \in \mathrm{Hom}_{\mathbb{R}}(L_0, L_1)$ we have that $L = \Gamma(T)$, which means all elements of L are of the form $v + T(v)$ for $v \in L_0$. Since L_0 and L_1 are Lagrangian, and therefore isotropic, $\omega(v, w) = \omega(T(v), T(w)) = 0$, so for $v + T(v),\ w + T(w) \in L$ we have

$$\omega(v + T(v), w + T(w)) = \omega(T(v), w) - \omega(T(w), v)$$

On the other hand, $\rho_{01}(T(v)) = \omega(T(v), \cdot)$, so the form $\rho_{01}(T(v))(w)$ is symmetric in v and w iff $\omega(v + T(v), w + T(w)) = 0$, i.e. iff L is isotropic, and therefore Lagrangian. ■

Let $L_1 \in \mathcal{L}(V, \omega)$ and define the subset Λ_{L_1} of U_{L_1} to be that consisting of Lagrangian subspaces of V transversal to L_1:

$$\Lambda_{L_1} := \mathcal{L}(V, \omega) \cap U_{L_1} \tag{2.93}$$

Then, by the above lemma the charts on the domains Λ_{L_1} are precisely those landing in the subspace $\mathrm{Hom}^2_{\mathbb{R},\mathrm{Sym}}(L_0; \mathbb{R}) \cong \mathrm{Sym}_n(\mathbb{R})$ of $\mathrm{Hom}_{\mathbb{R}}(L_0, L_1)$, for a fixed Lagrangian complement $L_0 \in \Lambda_{L_1}$ of L_1, $V = L_0 \oplus L_1$:

$$\psi_{01} : \Lambda_{L_1} \to \mathrm{Hom}^2_{\mathbb{R},\mathrm{Sym}}(L_0; \mathbb{R}) \cong \mathrm{Sym}_n(\mathbb{R}) \tag{2.94}$$

$$\psi_{01}(L) := \rho_{01} \circ \phi_{01}(L) \tag{2.95}$$

Since ρ_{01} is a linear isomorphism and ϕ_{01} is a chart on $\mathrm{G}(n, V)$, the transition functions are automatically smooth, by Proposition 11 above, so $\mathcal{L}(V, \omega)$ acquires a smooth manifold structure of dimension $n(n+1)/2$. We summarize these results in the following proposition:

Proposition 46 *For any symplectic vector space (V, ω), the Lagrangian Grassmannian $\mathcal{L}(V, \omega)$ is a smooth submanifold of the Grassmannian manifold $\mathrm{G}(n, V)$ of dimension $\dim \mathcal{L}(V, \omega) = n(n+1)/2$.*

■

Let us now describe $\mathcal{L}(V, \omega)$ as a homogeneous space under the $\mathrm{Sp}(n, \mathbb{R})$-action, or, what turns out to be the same thing, its restriction to a $\mathrm{U}(n)$-action. As we saw above, Lemma 4, given a Lagrangian decomposition (L_0, L_1) of $\mathbb{R}^{2n}$, the image of any $L \in \Lambda_{L_1}$ under ρ_{01} is a symmetric matrix $A \in \mathrm{Sym}_n(\mathbb{R})$ whose graph, as a map from L_0 to L_1 is precisely L. We give here another, simpler, proof of this fact, adapted to $\mathbb{R}^{2n}$, which has the benefit of describing the graph/Lagrangian L in terms of the image of a matrix $C \in M_{2n,n}(\mathbb{R})$ whose columns contain some extra information.

Lemma 5 *Let $A, B \in M_n(\mathbb{R})$ and define*

$$C := \begin{pmatrix} A \\ B \end{pmatrix} \in M_{2n,n}(\mathbb{R}) \cong \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^{2n}, \mathbb{R}^n)$$

Then, $\operatorname{im} C \in \mathcal{L}(n)$ iff $\operatorname{rank} C = n$ and $A^T B = B^T A$. Such a C is called a **Lagrangian frame***. Consequently, for any $B \in M_n(\mathbb{R}^n) \cong \mathrm{End}_{\mathbb{R}}(\mathbb{R}^n)$, since we may describe the graph of B as the image of C where $A = I$ and $B = B$,*

$$\Gamma(B) = \{(\mathbf{x}, B\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n\} = \operatorname{im} \begin{pmatrix} I_n \\ B \end{pmatrix}$$

we have

$$\Gamma(B) \in \mathcal{L}(n) \quad \Longleftrightarrow \quad B \in \mathrm{Sym}_n(\mathbb{R}) \tag{2.96}$$

Proof: $\operatorname{im} C \in \mathcal{L}(n)$ iff $\operatorname{im} C = \operatorname{im} C^{\omega_0}$ iff $\omega_0(\mathbf{u}, \mathbf{v}) = 0$ for all $\mathbf{u}, \mathbf{v} \in \operatorname{im} C$. Therefore, letting

$\mathbf{u} = C\mathbf{x}$ and $\mathbf{v} = C\mathbf{y}$, this is equivalent to

$$0 = \omega_0(C\mathbf{x}, C\mathbf{y}) = g_{J_0}(J_0 C\mathbf{x}, C\mathbf{y}) = \mathbf{x}^T \begin{pmatrix} -B \\ A \end{pmatrix}^T \begin{pmatrix} A \\ B \end{pmatrix} \mathbf{y} = \mathbf{x}^T (A^T B - B^T A)\mathbf{y}$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. ■

Remark 27 Suppose C is a Lagrangian frame $\operatorname{im} C \in \mathcal{L}(n)$. The columns of C form an *orthonormal* basis for $\operatorname{im} C$ iff additionally $A^T A = B^T B = I_n$, which is the case iff the matrix

$$V = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}$$

satisfies $V^T V = V V^T = I_{2n}$. But by Proposition 25 and its proof this is equivalent to the complex matrix $U = A + iB \in \mathrm{GL}(n, \mathbb{C})$, which is identified with V by Proposition 17, being unitary. That is,

$$\operatorname{im} \begin{pmatrix} A \\ B \end{pmatrix} \in \mathcal{L}(n) \quad \Longleftrightarrow \quad U := A + iB \in \mathrm{U}(n)$$

In this case we call C a **unitary Lagrangian frame**. ■

Let $\mathrm{Sp}(n, \mathbb{R})$ act on $\mathcal{L}(n)$ by left multiplication,

$$\mathrm{Sp}(n, \mathbb{R}) \times \mathcal{L}(n) \to \mathcal{L}(n) \tag{2.97}$$

$$\Phi \cdot L := \Phi(L) = \operatorname{im} \Phi|_L \tag{2.98}$$

or more generally let (V, ω) be a symplectic space and let $\mathrm{Sp}(V, \omega)$ act on $\mathcal{L}(V, \omega)$ by

$$\mathrm{Sp}(V, \omega) \times \mathcal{L}(V, \omega) \to \mathcal{L}(V, \omega) \tag{2.99}$$

$$\Phi \cdot L := \Phi(L) = \operatorname{im} \Phi|_L \tag{2.100}$$

Clearly $\Phi(L) \in \mathcal{L}(V, \omega)$, by the definition of $\mathrm{Sp}(V, \omega)$, for $\omega(\Phi(u), \Phi(v)) = \Phi^*\omega(u, v) = \omega(u, v)$, and this equals 0 iff $u, v \in L$.

Proposition 47 *The action of* $\mathrm{Sp}(n,\mathbb{R})$ *on* $\mathcal{L}(n)$ *is transitive. In fact, the unitary subgroup* $\mathrm{U}(n)$ *of* $\mathrm{Sp}(n,\mathbb{R})$ *acts transitively on* $\mathcal{L}(n)$*, with*

$$\mathcal{L}(n) = \mathrm{U}(n) \cdot L_{\mathrm{hor}}$$

The stabilizer of L_{hor} *under this action is* $\mathrm{O}(n)$*, where we identify* $\mathrm{O}(n)$ *with the subgroup of* $\mathrm{O}(2n)$ *by* $\mathrm{O}(n) \ni O \mapsto O \oplus O \in \mathrm{O}(2n)$*. Consequetly, we have the diffeomorphism*

$$\mathcal{L}(n) \approx \mathrm{U}(n)/\,\mathrm{O}(n) \tag{2.101}$$

which again confirms that $\dim \mathcal{L}(n) = n^2 - n(n-1)/2 = n(n+1)/2$.

Proof: Let $L \in \mathcal{L}(n)$ and find a unitary Lagrangian frame C as in the remark above, and identify it with the unitary matrix

$$\mathrm{U}(n) \ni U = A + iB \cong \Phi = \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \in \mathrm{O}(2n) \cap \mathrm{Sp}(n,\mathbb{R})$$

as in Proposition 25. Then note that

$$L = \operatorname{im} C = \operatorname{im} \begin{pmatrix} A \\ B \end{pmatrix} = \operatorname{im} \Phi|_{\mathbb{R}^n \times \{\mathbf{0}\}} = \Phi(L_{\mathrm{hor}})$$

Thus, if $L_1, L_2 \in \mathcal{L}(n)$, let $\Phi_1, \Phi_2 \in \mathrm{U}(n)$ be the respective unitary frames so that $L_1 = \Phi_1(L_{\mathrm{hor}})$ and $L_2 = \Phi_2(L_{\mathrm{hor}})$. Then $\Phi_2\Phi_1^{-1} \in \mathrm{U}(n)$ gives $\Phi_2\Phi_1^{-1}(L_1) = L_2$.

Finally, note that if $O \in \mathrm{O}(n)$, then

$$O \oplus O \cdot \Phi = \begin{pmatrix} O & 0 \\ 0 & O \end{pmatrix} \begin{pmatrix} A & -B \\ B & A \end{pmatrix} = \begin{pmatrix} OA & -OB \\ OB & OA \end{pmatrix}$$

which gives another unitary Lagrangian frame OC for the same Λ. In particular, it leaves

note that if $U = A + iB \in \mathrm{U}(n)$ is identified with $\Phi \in \mathrm{Sp}(n,\mathbb{R}) \cap \mathrm{O}(2n)$ and with the unitary Lagrangian framd C as above, then right-multiplying C or its equivalent U by an orthogonal matrix $O \in \mathrm{O}(n)$, CO, gives another unitary Lagrangian frame for the same Lagrangian subspace L. This defines ∎

2.4 Symplectic Manifolds

In this section we give an outline of **nonlinear** symplectic geometry, which consists of attaching the linear symplectic geometric structures of the previous section onto manifolds fiberwise. The process is succinctly encapsulated in the notion of the symplectic vector bundle. Such a vector bundle may potentially also be endowed with other structures, such as Riemannian metrics and Hermitian inner products, but now, as one might suspect, the compatibility of these structures with the symplectic structure may face some topological obstructions. The explication of these intricacies of nonlinearity will occupy us for the remainder of this section. The most interesting case will be that of the symplectic tangent bundle, which is the defining feature of a symplectic manifold. Such manifolds have their origin in classical mechanics, where they figure as the phase space of a classical system, the cotangent bundle to the configuration space of the system.

2.4.1 Symplectic Vector Bundles and Chern Classes

In what follows we put the symplectic linear algebra of the previous section onto a manifold M in the most general way possible, as fibers of a vector bundle E over M. The natural construction for this operation is the symplectic vector bundle. We largely follow the treatment in McDuff and Salamon [76] and Audin [12].

Definition 8 A **symplectic vector bundle** (E, ω) over a smooth manifold M is a real rank $2n$ vector bundle $E \to M$ with typical fiber a $2n$-dimensional symplectic vector space (V, ω). This means:

(1) E is locally trivializable with $2n$-dimensional fiber E_p over each $p \in M$ isomorphic to V.

(2) E comes equipped with a symplectic form $\omega \in \Omega^2(E)$, a nondegenerate section of the vector bundle $\bigwedge^2 E^* \to M$ of 2-covectors in E.

These conditions formalize the intuitive picture of the pair $(E_p, \omega_p) \cong (V, \omega)$ varying smoothly with $p \in M$, or parametrized smoothly by M. Two symplectic bundles (E_1, ω_1) and (E_2, ω_2) over M are said to be **(symplectically) isomorphic** if there is a vector bundle map $F : E_1 \to E_2$ covering the identity of M such that $F^*\omega_2 = \omega_1$.

Let us also introduce the vector bundle analogs of the spaces of metrics, symplectic forms, and (tame, calibrated, and compatible) complex structures. A **(Riemannian) metric** on E is a global section $g \in \Gamma(\bigodot^2 E^*) \cong \mathrm{Hom}^2_{C^\infty(M),\mathrm{Sym}}(\Gamma(E);\mathbb{R})$. Similarly, a symplectic form is a section $\omega \in \Omega^2(E) = \Gamma(\bigwedge^2 E^*)$, and the fiber bundle of such sections will be denoted $\mathrm{Symp}(E) \to M$. A **complex structure** J on E is a bundle endomorphism $J \in \mathrm{End}(E)$ satisfying $J^2 = -I$. Alternatively, a complex structure a field of E-valued complex structures, $J_p^2 = -I_p$ on E_p for all $p \in M$, that is a section of a certain fiber bundle

$$\mathcal{J}(E) \to M$$

whose fibers are $\mathcal{J}(E)_p := \mathcal{J}(E_p)$, the space of complex structures on the vector space $E_p \cong \mathbb{R}^{2n}$. We similarly define tame, calibrated and compatible complex structures, fiberwise, as sections of the fiber bundles $\mathcal{J}_\tau(E)$, $\mathcal{J}_c(E)$ and $\mathcal{J}(E,\omega)$ whose fibers are $\mathcal{J}_\tau(E_p)$, $\mathcal{J}_c(E_p)$ and $\mathcal{J}(E_p,\omega_p)$, respectively. Lastly, a **Hermitian structure** on E is a triple (ω, J, g) with $J \in \mathcal{J}(E,\omega)$ and $g = g_J = \omega \circ I \times J$, from which, when it exists, we can construct a Hermitian inner product $H_J := g + i\omega$ for E as in Proposition 30. The existence of these fiber bundles over an arbitrary $2n$-dimensional manifold M is not guaranteed (see Example 12 below), though by Proposition 48 below we have that whenever M is equipped with a symplectic form ω, $\mathcal{J}(E,\omega)$ is nonempty and contractible, and conversely whenever there exists $J \in \Gamma(\mathcal{J}(E))$, then there is a symplectic form $\omega \in \Omega^2(E)$ such that $J \in \Gamma(\mathcal{J}(E,\omega))$. ■

Theorem 15 *The structure group of any rank* $2n$ *symplectic vector bundle* (E,ω) *may be reduced from* $\mathrm{GL}(2n,\mathbb{R})$ *to* $\mathrm{Sp}(n,\mathbb{R})$.

Proof: The typical fiber V of E may, without loss of generality, be chosen to be $(\mathbb{R}^{2n},\omega_0)$, for locally we can trivialize E by a diffeomorphism $\Phi : E_U \to U \times V$ which restricts to a linear isomorphism on fibers, $E_p \cong \{p\} \times V$, and then follow this with $\mathrm{id}_U \times \alpha$ where α is a linear symplectomorphism between (V,ω) and $(\mathbb{R}^{2n},\omega_0)$ as in Proposition 32. We may rephrase this by saying that locally, over U, there exist sections of a symplectic frame bundle for E, the principal $\mathrm{Sp}(n,\mathbb{R})$-bundle $P_{\mathrm{Sp}(n,\mathbb{R})} \to M$. Namely, let $\beta \in V_{2n}^{\mathrm{Sp}}(\mathbb{R}^{2n}) \approx \mathrm{Sp}(n,\mathbb{R})$ be a symplectic frame for $\mathbb{R}^{2n}$ (cf Proposition 14 of Section 2.2.3), and use $\Phi^{-1}\circ \mathrm{id}_U \times \alpha^{-1}$ to pull it back to a frame $\tilde{\beta} \in V_{2n}^{\mathrm{Sp}}(E_p)$ for E_p. Since this same frame works for all $p \in U$, $\tilde{\beta} \in \Gamma(U, P_{\mathrm{Sp}(n,\mathbb{R})})$ is a local frame for all of E_U.

Now consider two local trivializations $\Phi_i : E_{U_i} \to U_i \times \mathbb{R}^{2n}$ and $\Phi_j : E_{U_j} \to U_j \times \mathbb{R}^{2n}$ with $U_i \cap U_j$ nonempty, and let $s_i, s_j \in \Gamma(U_i \cap U_j, P_{\mathrm{Sp}(n,\mathbb{R})})$ be the corresponding local symplectic frames over U_i and U_j, respectively. Then the linear map taking s_i to s_j fiberwise must be a symplectic matrix by the discussion preceding Proposition 14. Consequently, by using the symplectic frame bundle $P_{\mathrm{Sp}(n,\mathbb{R})}$ instead of the general frame bundle $P_{\mathrm{GL}(2n,\mathbb{R})}$ we may reduce the structure group of E from $\mathrm{GL}(2n,\mathbb{R})$ to $\mathrm{Sp}(n,\mathbb{R})$.[5] ■

Proposition 48 *Let $E \to M$ be a real rank $2n$ vector bundle.*

(1) *For every symplectic form $\omega \in \Omega^2(E)$ on E there exists an ω-compatible complex structure $J \in \mathcal{J}(E,\omega)$. The space $\mathcal{J}(E,\omega)$ is therefore nonempty and contractible.*

(2) *For each complex structure $J \in \mathcal{J}(E)$ there exists a symplectic form $\omega \in \Omega^2(E)$ which is compatible with J. Let us denote by* $\mathrm{Symp}(E,J)$ *the fiber bundle of complex structures compatible with ω. Then* $\mathrm{Symp}(E,J)$ *is nonempty and contractible.*

Proof: These statements follow from their vector space analogs, Propositions 37, 38 and 41 applied fiberwise to E_p and then over open sets U to $E_U = \pi^{-1}(U)$. ■

Recall the construction of characteristic classes via classifying spaces, as described, for example, in Milnor and Stasheff [79]. Fix M a paracompact space (M can safely be taken to be a manifold or CW complex). Then for each topological group G there exists a connected topological

[5] The formalism for this reduction is the existence a principal bundle morphism $\iota : P_{\mathrm{Sp}(n,\mathbb{R})} \to P_{\mathrm{GL}(2n,\mathbb{R})}$ (the inclusion fiberwise of $\mathrm{Sp}(n,\mathbb{R})$ in $\mathrm{GL}(2n,\mathbb{R})$) covering the identity which is equivariant with respect to the Lie group inclusion homomorphism $i : \mathrm{Sp}(n,\mathbb{R}) \hookrightarrow \mathrm{GL}(2n,\mathbb{R})$, i.e. satisfies $\iota(\tilde{\beta}\cdot g) = \iota(\tilde{\beta})\cdot i(g)$. The discussion above gives us a method for constructiong this bundle morphism: ι is the inclusion, which fiberwise is just i, and the equivariance is an easy consequence of the fact that $\mathrm{Sp}(n,\mathbb{R})$ is a subgroup of $\mathrm{GL}(2n,\mathbb{R})$.

space BG, the classifying space for G, and a weakly contractible (all homotopy groups are trivial) principal G-bundle $EG \to BG$ such that the following holds: The set of equivalence classes $\mathrm{Prin}_G(M)$ of principal G-bundles over M is in bijective correspondence with the set $[M, BG]$ of homotopy classes of continuous maps $f : M \to BG$, sending $[f] \in [M, BG]$ to $[f^*EG]$ (in fact the contravariant functors Prin_G and $[\,\cdot\,, BG]$ from **hTop** to **Sets** are naturally equivalent, see Husemoller [58, Proposition 10.4, Theorem 12.12]). If in addition G is any connected or semi-simple Lie group G, then G is topologcially the product of a compact subgroup H and a Euclidean space E, $G \approx H \times E$ (this is the case with $\mathrm{Sp}(n, \mathbb{R})$, according to Proposition 26, which gives $\mathrm{Sp}(n, \mathbb{R}) \approx \mathrm{U}(n) \times \mathbb{R}^{n^2+n}$), which means H is a deformation retract of G. This is the case for $G = \mathrm{GL}(n, \mathbb{R})$ or $\mathrm{GL}(n, \mathbb{C})$ with H the maximal compact subgroup $\mathrm{O}(n)$ or $\mathrm{U}(n)$, respectively, but it applies elsewhere, too, as with $G = \mathrm{Sp}(n, \mathbb{R})$. In such a case, the structure group of a principal G-bundle $P_G \to M$ may be reduced to H (see Remarks 12.13 and 12.14 following Theorem 12.7 in Steenrod [96]). This broad result is due to Iwasawa [60], the original statement holding for G semi-simple and due to Elie Cartan [24] in 1927, with a simplified proof by Mostow [83] in 1949. Moreover, by Theorem 5.1 in Husemoller [58], given a Lie group G and a closed Lie subgroup H, the principal H-bundle reductions of a given principal G-bundle P_G are in bijective correspondence with the homotopy classes of maps $f : M \to BH$ such that $f_0 \circ f \simeq g$, where $g : M \to BG$ and $f_0 : BH \to BG$ is covered by the bundle map $h_0 : P_H \times_i G \to P_G$, with $i : H \hookrightarrow G$ the inclusion. If H is a deformation retract of G, therefore, there is only one such homotopy class, and so there is only one homotopy class of f_0, which means BH and BG are homotopy equivalent.

Now, let us apply these facts to the structure group $G = \mathrm{Sp}(n, \mathbb{R})$ for a symplectic vector bundle (E, ω) over a smooth manifold M, recalling that $\mathrm{Sp}(n, \mathbb{R})$ is topologically the product of its maximal compact subgroup $\mathrm{U}(n)$ and a Euclidean space $\mathbb{R}^{n^2+n}$ (Proposition 26) and deformation retracts onto $\mathrm{U}(n)$ (Proposition 27):

Theorem 16 *The structure group of any symplectic vector bundle (E, ω) may be further reduced from $\mathrm{Sp}(n, \mathbb{R})$ to its maximal compact subgroup $\mathrm{U}(n)$, and these reductions are in one-to-one corre-*

spondence with the space of ω-compatible complex structures $\mathcal{J}(E,\omega)$. As a result, every symplectic vector bundle has an underlying complex structure and a Hermitian inner product, and any two symplectic bundles (E_1,ω_1) and (E_2,ω_2) over M are symplectically isomorphic iff their underlying complex vector bundles are isomorphic.

Proof: Then the set of equivalence classes $\mathrm{Prin}_{\mathrm{Sp}(n,\mathbb{R})}(M)$ of principal $\mathrm{Sp}(n,\mathbb{R})$-bundles is in bijective correspondence with the set of homotopy classes $[M, B\,\mathrm{Sp}(n,\mathbb{R})]$ of maps from M to the classifying space $B\,\mathrm{Sp}(n,\mathbb{R})$. The fact that $\mathrm{Sp}(n,\mathbb{R})$ deformation retracts onto its maximal compact subgroup $\mathrm{U}(n)$ (Proposition 27) means that $B\,\mathrm{U}(n)$ is homotopy equivalent to $B\,\mathrm{Sp}(n,\mathbb{R})$, and any principal $\mathrm{Sp}(n,\mathbb{R})$-bundle is equivalent to a principal $\mathrm{U}(n)$-bundle, whose associated vector bundle $E \cong P_{\mathrm{U}(n)} \times_{\mathrm{U}(n)} \mathbb{C}^n$ is by nature a complex Hermitian vector bundle (transition functions taking values in $\mathrm{U}(n)$ preserve the Hermitian inner product on overlaps). Since principal $\mathrm{Sp}(n,\mathbb{R})$-bundles are in one-to-one correspondence with symplectic vector bundles (send $P_{\mathrm{Sp}(n,\mathbb{R})}$ to its associated vector bundle $P_{\mathrm{Sp}(n,\mathbb{R})} \times_{\mathrm{Sp}(n,\mathbb{R})} \mathbb{R}^{2n}$), we see that E admits a complex Hermitian structure.

Now recall Proposition 40, which gave $\mathcal{J}(\mathbb{R}^{2n},\omega_0)$ the structure of a homogeneus space, $\mathcal{J}(\mathbb{R}^{2n},\omega_0) \approx \mathrm{Sp}(n,\mathbb{R})/\,\mathrm{U}(n)$, sending J to $\Phi_J \cdot \mathrm{U}(n)$, where $\Phi_J \in \mathrm{Sp}(n,\mathbb{R})$ was defined in Proposition 32, using J. The contractibility of $\mathrm{Sp}(n,\mathbb{R})/\,\mathrm{U}(n)$ was shown in Proposition 27, where the explicit homotopy $H : \mathrm{id}_{\mathrm{Sp}(n,\mathbb{R})} \to r$, with $r : \mathrm{Sp}(n,\mathbb{R}) \to \mathrm{U}(n)$ the retract $r(A) := A(A^TA)^{-1/2}$, was given by

$$
\begin{aligned}
&H : \mathrm{Sp}(n,\mathbb{R}) \times [0,1] \to \mathrm{Sp}(n,\mathbb{R}) \\
&H(A,t) := A(A^TA)^{-t/2}
\end{aligned}
$$

Now, each nonunitary $A \in \mathrm{Sp}(n,\mathbb{R})$ corresponds to a unique $J \in \mathcal{J}(\mathbb{R}^{2n},\omega_0)$, via the map $J \mapsto \Phi_J = A$, so the homotopy H above applies to this Φ_J. The argument generalizes in the obvious way to any arbitrary symplectic vector space (V,ω). As a consequence, we have the following important fact: *The ω-compatible complex structures $\mathcal{J}(E,\omega)$ on a symplectic vector bundle (E,ω) over M are in one-to-one correspondence with $\mathrm{U}(n)$-reductions of the structure group $\mathrm{Sp}(n,\mathbb{R})$, each nonunitary symplectic frame $\beta \in P_{\mathrm{Sp}(n,\mathbb{R})}$ being sent to $\gamma = H(\beta,1) \in P_{\mathrm{U}(n)}$ by the unique $J \in \mathcal{J}(E,\omega)$*

satisfying $\Phi_J = \beta$ *(recalling that* $V_{2n}^{\mathrm{Sp}}(\mathbb{R}^{2n}) \approx \mathrm{Sp}(n, \mathbb{R})$*, Proposition 14).*

For the last statement, suppose first that (E_1, ω_1) and (E_2, ω_2) are symplectically isomorphic. Then, there is a map $F : E_1 \to E_2$ such that $F^*\omega_2 = \omega_1$. Choose $J_i \in \mathcal{J}(E_i, \omega_i)$ and note that $J_1, F^*J_2 \in \mathcal{J}(E_1, \omega_1)$ (since $F^*J_2 = F^{-1} \circ J_2 \circ F$ implies $(F^*J_2)^2 = F^{-1} \circ (-I) \circ F = -I$ and $\omega_1(u, F^*J_2 v) = F^*\omega_2(u, (F^{-1} \circ J_2 \circ F)v) = \omega_2(Fu, J_2 F v) \geq 0$), so that by convexity we have a smooth family $J_t \in \mathcal{J}(E_1, \omega_1)$ joining $J_0 := F^*J_2$ and J_1, which by the bundle analog of Propositions 32 and 37 means there is a smooth family of bundle isomorphisms $F_t : E_1 \to E_1$ such that $F_t^* J_t = J_1$. Therefore $F \circ F_0 : E_1 \to E_2$ is a bundle isormorphism which intertwines J_1 and J_2, and this is our complex bundle isomorphism. The converse follows from an application of (2) of the previous proposition. ∎

Remark 28 Since the set of isomorphism classes $\mathrm{Vect}_{\mathrm{Sp}}^{2k}(M)$ of symplectic vector bundles over M coincides with the set of isomorphism classes $\mathrm{Vect}_{\mathbb{C}}^{k}(M)$ of complex vector bundles over M, symplectic vector bundles have the same characteristic classes as complex vector bundles. These are the Chern classes. ∎

Remark 29 (Hermitian Vector Bundles) Of course complex vector bundles are not necessarily holomorphic. They are merely smooth vector bundles with typical fiber a complex vector space and complex linear transition functions. **Holomorphic vector bundles** $\pi : E \to M$ additionally require E and M to be complex manifolds (meaning atlas charts (U, φ) are holomorphic: they are homeomorphisms onto open subsets of $\mathbb{C}^n$ and transition functions $\varphi_i \circ \varphi_j^{-1}$ are holomorphic), π to be holomorphic, and the local trivializations $\Phi : \pi^{-1}(U) \to U \times \mathbb{C}^n$ to be biholomorphic. Analogously to the real vector bundle case, holomorphic vector bundles determine and are determined by cocycles $\varphi_{ij} := \varphi_i \circ \varphi_j^{-1} : U_i \cap U_j \to \mathrm{GL}(n, \mathbb{C})$, which in this case are additionally required to be *holomorphic* (Huybrechts [59, Remark 2.2.2]).

Given a complex vector bundle $E \to M$, its structure group $\mathrm{GL}(n, \mathbb{C})$ can be reduced to its maximal compact subgroup $\mathrm{U}(n)$, which is a deformation retract of $\mathrm{GL}(n, \mathbb{C})$ (via the Gram-Schmidt process, see the remarks following Corollary 5), and thus we may suppose E endowed with a Hermitian inner product H. From this Hermitian inner product we can get both a Riemannian metric g and a symplectic form ω on E, as follows:

$$g := \frac{1}{2}(H + \overline{H})$$
$$\omega := -\frac{i}{2}(H - \overline{H})$$

Then g is symmetric and ω is skew-symmetric:

$$g(u,v) = \frac{1}{2}(H(u,v) + \overline{H(u,v)}) = \frac{1}{2}(\overline{H(v,u)} + H(v,u)) = g(v,u)$$
$$\omega(u,v) = -\frac{i}{2}(H(u,v) - \overline{H(u,v)}) = \frac{i}{2}(H(v,u) - \overline{H(v,u)}) = -\omega(v,u)$$

and both are nondegenerate: to see that ω is nondegenerate, note that it can be represented as J_0 with respect to a well-chosen (symplectic) basis β. For let J be the given complex structure i on E, viewed momentarily as a real $2n$-dimensional real vector bundle, and decompose E into the subbundles $E^{1,0}$ and $E^{0,1}$, the ± 1 eigenbundles of J. Then, as in the proof of Proposition 32, any complex basis $(u_j + iv_j)_{j=1}^n$ for $E^{1,0}$ gives a real basis $(v_1, \dots, v_n, u_1, \dots, u_n)$ for E with the property that $v_j = J(u_j)$ and $J(v_j) = -u_j$, so $[\omega]_\beta = J_0$. Moreover, this J is ω-compatible, for

$$\begin{aligned}\omega(u,Jv) = -\frac{i}{2}(H(u,Jv) - \overline{H(u,Jv)}) &= -\frac{i}{2}(iH(u,v) + i\overline{H(u,v)}) \\ &= \frac{1}{2}(H(u,v) + \overline{H(u,v)}) = g(u,v)\end{aligned}$$

This incidentally also shows that g is nondegenerate, for the symplectic basis will diagonalize g, in view of $g(u_i,u_j) = \omega(u_i,Ju_j) = \omega(u_i,v_j) = \delta_{ij}$, $g(v_i,v_j) = \omega(v_i,Jv_j) = -\omega(v_i,u_j) = \omega(u_j,v_i) = \delta_{ij}$. Lastly, it is clear that

$$H = g + i\omega$$

If we replace $-\frac{i}{2}(H - \overline{H})$ with its positive counterpart $\frac{i}{2}(H - \overline{H})$, which is called the **fundamental form**, then we need to replace $J = i$ with $J = -i$, and in this case H becomes $H = g - i\omega$. ■

Our interest, of course, is in the case of E the tangent bundle TM of an even-dimensional manifold M.

2.4.2 Symplectic Manifolds: Definitions and Examples

2.4.2.1 Definitions and Basic Properties

Definition 9 A **symplectic manifold** (M,ω) is a smooth manifold M of real dimension $2n$ equipped with a **symplectic form**, that is a nondegenerate closed 2-form $\omega \in \Omega^2(M)$. Explicitly, this means ω satisfies

(1) $d\omega = 0$

(2) $\omega_p \in \bigwedge^2 T_p^*M \cong \mathrm{Hom}^2_{\mathbb{R},\mathrm{Skew}}(T_pM;\mathbb{R})$ is symplectic for each $p \in M$, that is ω_p is a skew-symmetric and nondegenerate bilinear form on each tangent space.

Let (M,ω_M) and (N,ω_N) be two symplectic manifolds. A smooth map $f : M \to N$ is said to be **symplectic** or a **morphism of symplectic manifolds** if f preserves the form,

$$
\begin{aligned}
&f^*\omega_N = \omega_M, \text{ i.e.}\\
&\omega_M(X,Y) = \omega_N\big(Tf(X), Tf(Y)\big) \quad \forall X,Y \in \Gamma(TM)
\end{aligned}
$$

This means that, fiberwise, T_pf is a linear symplectic morphism from (T_pM, ω_{M_p}) to $(T_{f(p)}N, \omega_{N_{f(p)}})$. A **symplectomorphism** is a symplectic diffeomorphism.[6] Clearly f is a symplectomorphism iff f^{-1} is, since T_pf is invertible and symplecitc fiberwise. In the case $M = N$ the collection of symplectomorphisms of M forms a group under composition, called the **symplectic group** of M, denoted

$$
\mathrm{Sp}(M,\omega) \quad \text{or} \quad \mathrm{Sp}(M) := \{f \in \mathrm{Diff}(M) \mid f^*\omega = \omega\}
$$

Obviously, if $f \in \mathrm{Sp}(M,\omega)$, then fiberwise T_pf is an element of the group of *linear* symplectomorphisms under the identification $\mathrm{Sp}(T_pM,\omega_p) \cong \mathrm{Sp}(n,\mathbb{R})$. ■

Remark 30 If $\omega \in \Omega^2(M)$ is not closed, then we say (M,ω) is an **almost symplectic manifold**, and in this case (TM,ω) is merely a symplectic vector bundle over M, and consequently a complex vector bundle, by Proposition 48 and Theorem 16. ■

[6] To distinguish the manifold case from the linear case, we shall write 'linear' in front of 'symplectic' and 'symplectomorphism' to clarify the meaning.

Proposition 49

(1) *Any almost symplectic manifold* (M,ω) *is orientable.*

(2) *No symplectic form* ω *on a closed (compact without boundary) manifold is exact. Moreover, in such a case* ω*, being closed, induces a nonzero cohomology class* $[\omega^n] = [\omega]^n \in H^{2n}_{dR}(M;\mathbb{R})$*.*

Proof: Since a 2-form $\omega \in \Omega^2(M)$ is nondegenerate iff $\omega^n \in \Omega^{2n}(M)$ is nonzero (Theorem 13), we see that in such a case ω^n is a volume form and so every almost symplectic manifold (M,ω) is *orientable*. If ω is additionally closed, then it defines a nonzero cohomology class $[\omega] \in H^2_{dR}(M;\mathbb{R})$. When M is a closed manifold, then $[\omega]^n := [\omega] \smile \cdots \smile [\omega] = [\omega^n] \in H^{2n}_{dR}(M;\mathbb{R})$ is also nonzero (Salamon and McDuff [76, p. 83]), and this furnishes us with many examples of non-symplectic even-dimensional manifolds (see Remark 32 below). To see that ω cannot be exact on a closed manifold, use Stokes' Theorem: if $\omega = d\alpha$, then

$$\omega^n = (d\alpha)^n = d\alpha \wedge (d\alpha)^{n-1} = d\alpha \wedge (d\alpha)^{n-1} + d\alpha \wedge d(d\alpha)^{n-1} = d(\alpha \wedge \omega^{n-1})$$

so

$$\int_M \omega^n = \int_{\partial M} d\omega = \int_{\partial M} d^2(\alpha \wedge \omega^{n-1}) = \int_{\partial M} 0 = 0$$

which is impossible since ω^n is a volume form, which must give nonzero volume for M. ■

Definition 10 Let (M,ω) be a symplectic manifold. A bundle endomorphism $J \in \mathrm{End}(TM)$ is called an **almost complex structure** if $J^2 = -I$, where I is the idenitity element of $\mathrm{End}(TM)$. This means that fiberwise, on T_pM, we have $J_p^2 = -I_p$, i.e. J_p is a complex structure on each vector space $T_pM \cong \mathbb{R}^{2n}$. As in the case of vector spaces, we say that J is **tamed** by ω if the bilinear form $g_J := \omega \circ (I \times J)$ is symmetric and positive definite on $\Gamma(TM)$,

$$g_J(X,Y) = g_J(Y,X),$$

$$g_J(X,X) := \omega(X,JX) \geq 0, \quad \forall X \in \Gamma(TM), \text{ and } g_J(X,X) = 0 \text{ iff } X = 0$$

(g_J is always symmetric, but not always positive definite.) We say J is **calibrated** by ω if additionally $J \in \mathrm{Sp}(M,\omega)$, i.e. $J^*\omega = \omega$, or

$$\omega(JX, JY) = \omega(X, Y), \quad \forall X, Y \in \Gamma(TM)$$

If J is both tamed and calibrated by ω, then we say that J is **compatible** with ω. In this case, g_J is a Riemannian metric on M and $J \in \mathrm{Sp}(M,\omega) \cap \mathrm{O}(M, g_J)$, i.e. J also preserves the metric, $J^* g_J = g_J$, for

$$g_J(JX, JY) = \omega(JX, J^2Y) = -\omega(JX, Y) = \omega(Y, JX) = g_J(Y, X) = g_J(X, Y)$$

because these relations hold fiberwise (see Sections 2.3.3 and 2.3.5 above). ∎

Now, fiberwise $J_p \in \mathcal{J}(T_pM)$, so we can also view J as a field of complex structures, or a section of a certain bundle, $p \mapsto J_p \in \mathcal{J}(T_pM)$. To formalize this idea we introduce the following fiber bundles over the symplecic manifold (M,ω), and which are indeed fiber bundles over M by Darboux's theorem, which allows for the construction of local trivializations (see Bieliavsky et al. [17], and Proposition 50 below; these bundles do not necessarily exist over arbitrary even dimensional manifolds).

Definition 11 The **bundle of almost complex structures** is the fiber bundle

$$\mathcal{J}(TM) \to M \tag{2.102}$$

whose typical fiber is the $2n^2$-dimensional manifold $\mathcal{J}(\mathbb{R}^{2n}) \approx \mathrm{GL}(2n,\mathbb{R})/\mathrm{GL}(n,\mathbb{C})$ of complex structures (Proposition 35). We can thus think of $\mathcal{J}(TM)$ as a *principal* $\mathrm{GL}(2n,\mathbb{R})/\mathrm{GL}(n,\mathbb{C})$-bundle, $\mathcal{J}(TM) = P_{\mathrm{GL}(2n,\mathbb{R})/\mathrm{GL}(n,\mathbb{C})}$. We also have the **bundle of ω-tame almost complex structures**,

$$\mathcal{J}_\tau(TM,\omega) \to M \tag{2.103}$$

which is a fiber bundle with typical fiber the $2n^2$-dimensional manifold $\mathcal{J}_\tau(\mathbb{R}^{2n},\omega_0) \approx \{A \in M_{2n}(\mathbb{R}) \mid AJ_0 + J_0A = 0\} \approx \mathbb{R}^{2n^2}$ (Propositon 36), as well as the **bundle of ω-calibrated almost**

complex structures,

$$\mathcal{J}_c(TM,\omega)\to M \tag{2.104}$$

a fiber bundle with typical fiber the $n(n+1)$-dimensional manifold $\mathcal{J}_c(\mathbb{R}^{2n},\omega_0)\cong\mathfrak{sp}(n,\mathbb{R})\cap\mathrm{Sym}_{2n}(\mathbb{R})\cong\mathbb{R}^{n(n+1)}$. Lastly, we have the **bundle of ω-compatible almost complex structures**,

$$\mathcal{J}(TM,\omega)\to M \tag{2.105}$$

which is a fiber bundle with typical fiber the $n(n+1)$-dimensional manifold $\mathcal{J}(\mathbb{R}^{2n})\approx\mathfrak{sp}(n,\mathbb{R})\cap\mathrm{Sym}_{2n}(\mathbb{R})\cong\mathbb{R}^{n(n+1)}$ (Remark 23). ■

Remark 31 The bundle $\mathcal{J}(TM,\omega)$ can also be viewed (Bieliavsky et al [17]) as the associated bundle to the symplectic frame bundle,

$$\mathcal{J}(TM,\omega)\cong P_{\mathrm{Sp}(n,\mathbb{R})}\times_{\mathrm{Sp}(n,\mathbb{R})}J(\mathbb{R}^{2n},\omega_0) \tag{2.106}$$

with typical fiber the space of vector-space complex structures $J(\mathbb{R}^{2n},\omega_0)\approx\mathrm{Sp}(n,\mathbb{R})/\,\mathrm{U}(n)$, by Proposition 40, which specifies the $\mathbb{R}^{2n}$-action on the homogeneous space $J(\mathbb{R}^{2n},\omega_0)$. Indeed, $\mathcal{J}(TM,\omega)$ is also a homogeneous space,

$$\mathcal{J}(TM,\omega)\approx P_{\mathrm{Sp}(n,\mathbb{R})}/\,\mathrm{U}(n) \tag{2.107}$$

$$p\circ J_0\circ p^{-1}\;\longleftrightarrow\;p\cdot\mathrm{U}(n) \tag{2.108}$$

In other words, the conjugation action of $\mathrm{Sp}(n,\mathbb{R})$ on $\mathcal{J}(\mathbb{R}^{2n},\omega_0)$ of Proposition 40 can be lifted to the symplectic frame bundle $\mathcal{J}(TM,\omega)$ fiberwise.

Another observation worth making is that the symplectic frame bundle $P_{\mathrm{Sp}(n,\mathbb{R})}$ has an $\mathbb{R}^{2n}$-valued 1-form called the **soldering form**

$$\theta\in\Omega^1(P_{\mathrm{Sp}(n,\mathbb{R})};\mathbb{R}^{2n})$$

$$\theta_p(X):=p^{-1}(T\pi(X)),\quad p\in\mathrm{Sp}(M,\omega),\ \pi:P_{\mathrm{Sp}(n,\mathbb{R})}\to M$$

To understand this form, consider the **vertical bundle** of $P_{\mathrm{Sp}(n,\mathbb{R})}$,

$$VP_{\mathrm{Sp}(n,\mathbb{R})} := \ker T\pi \to M$$

Then $VP_{\mathrm{Sp}(n,\mathbb{R})} = \ker\theta$. A **horizontal** 1-form is one which vanishes on $VP_{\mathrm{Sp}(n,\mathbb{R})}$, and we note that the components of θ are therefore horizontal forms. Given $X \in \mathfrak{sp}(n,\mathbb{R})$, we may define the **vertical vector field** $\tilde{X}$ on $P_{\mathrm{Sp}(n,\mathbb{R})}$ by

$$\tilde{X}_p := \left.\frac{d}{dt}\right|_{t=0} \big(p \circ \exp(tX)\big)$$

and from this we obtain the trivialization of the vertical bundle,

$$\begin{aligned} &P_{\mathrm{Sp}(n,\mathbb{R})} \times \mathfrak{sp}(n,\mathbb{R}) \cong VP_{\mathrm{Sp}(n,\mathbb{R})} \\ &(p, X) \mapsto \tilde{X}_p \end{aligned}$$

This is a standard construction, of course, and applies to any principal G-bundle $P_G \to M$. Note that the pullback bundle π^*TM is also trivial, via $(p, X) \mapsto (p, p^{-1}X)$, so θ imay be viewed as a map

$$\begin{aligned} &\theta : TP_{\mathrm{Sp}(n,\mathbb{R})} \to P_{\mathrm{Sp}(n,\mathbb{R})} \times \mathbb{R}^{2n} \\ &\theta_p(X) = \big(p, p^{-1}(T\pi(X))\big) \end{aligned}$$

We also get an exact sequence of bundles,

$$0 \longrightarrow P_{\mathrm{Sp}(n,\mathbb{R})} \times \mathfrak{sp}(n,\mathbb{R}) \longrightarrow TP_{\mathrm{Sp}(n,\mathbb{R})} \longrightarrow P_{\mathrm{Sp}(n,\mathbb{R})} \times \mathbb{R}^{2n} \longrightarrow 0$$

Splitting this exact sequence trivializes the tangent bundle of $P_{\mathrm{Sp}(n,\mathbb{R})}$. ■

Proposition 50 *Let M be a $2n$-dimensional manifold.*

(1) *For each nondegenerate 2-form $\omega \in \Omega^2(M)$ there exsits an ω-compatible almost complex structure $J \in \Gamma(\mathcal{J}(TM,\omega))$, and moreover the space $\Gamma(\mathcal{J}(TM,\omega))$ is path connected and contractible.*

(2) *For each almost complex structure* $J \in \Gamma(\mathcal{J}(TM))$ *there exists a nondegenerate 2-form* $\omega \in \Omega^2(M)$ *which is compatible with* J*, i.e.* $J \in \Gamma(\mathcal{J}(TM, \omega))$*. The space* $\Omega_J(M)$ *of such forms is path connected and contractible.*

Proof: This is just a special case of Theorem 48 with $E = TM$. ■

Remark 32 As a result of this proposition, we see that an even dimensional manifold M admits an almost complex structure iff it carries a nondegenerate 2-form (not necessarily closed). That is, M is an almost complex manifold iff it is an almost symplectic manifold. The former condition, however, does not hold for certain even dimensional manifolds. For example, the connected sum $\mathbb{CP}^2\#\mathbb{CP}^2$ does not admit any almost complex structure by a result due to Taubes and Wu (see [12, Proposition 1.5.1, Ex. 1.5.2] and [76, Ex. 4.7]). Less exotic examples are even-dimensional spheres S^{2n} for $n \in \mathbb{N}\backslash\{1,3\}$, a result known since the early 1950s (see Borel and Serre [19], and Wu [111]). Of course, since $H^{2n}_{dR}(S^{2n};\mathbb{R}) = 0$, by we know Proposition 49 that no even-dimensional spheres S^{2n} for $n \geq 2$ carries a symplectic structure, though by the Borel and Serre result we know that S^{2n} carries no nondegenerate 2-form whatsoever. For the cases S^2 and S^6 see Examples 13 and 9 below. ■

Remark 33 In his doctoral dissertation [46] of 1969, Mikhail Gromov proved that any *open* (boundaryless and containing no compact component) even dimensional almost symplectic (equivalently almost complex) manifold M is actually symplectic. The precise statement is: Any open almost complex manifold (M, J) (equivalently almost symplectic manifold (M, ω)) admits a symplectic structure ω which belongs to any prescribed cohomology class $a \in H^2(M)$ and such that $J \in [J_\omega]$ where $[J_\omega]$ is the homotopy class of almost complex structures compatible with ω. The situation for closed (compact without boundary) manifolds was rather different, and Gromov's h-principle could not be applied. Taubes, in his 1994-5 papers [98] and [99] showed by means of Seiberg-Witten invariants that in the case of closed almost symplectc/almost complex manifolds the situation was different. He produced the example of $\mathbb{CP}^2\#\mathbb{CP}^2\#\mathbb{CP}^2$, an almost complex/almost symplectic manifold carrying a cohomology class $a \in H^2(M)$ such that $a^n \neq 0$ and yet having no symplectic structure. ■

Thus, if we are given a symplectic manifold (M, ω), we automatically have an almost complex structure $J \in \Gamma(\mathcal{J}(M, \omega))$ compatible with ω and giving the tangent bundle TM the structure of a complex vector bundle. The obvious questions that arise are: (1) Does J make M a complex manifold, at least under certain circumstances? Or, to put it differently, does J come from a complex manifold structure on M? (2) Suppose M has at least one almost complex structure J. Is M a complex manifold? That is, do *any* of its almost complex structures come from a complex manifold structure? (3) If M *is* a complex manifold, under what conditions is it symplectic or Kähler?

There is a relatively easy answer to (1), namely (M, J) is complex precisely when J's Nijenhuis tensor N_J vanishes. This is the Newlander-Nirenberg theorem, which we describe below. However, the Newlander-Nirenberg theorem is not enough to answer (2), which seems to be an open question. The example of S^6 is the most likely to produce an instance of an almost complex manifold without a complex manifold structure, but is currently unresolved (see Example 9 below). (3) has been answered many times over, there are many non-Kähler symplectic and complex manifolds (see Example below).

Definition 12 An **almost complex manifold** (M, J) is a real $2n$-dimensional smooth manifold M equipped with an almost complex structure $J \in \Gamma(\mathcal{J}(TM))$. By the previous proposition such manifolds always possess a nondegenerate 2-form ω which is compatible with J. If ω were closed, this would make (M, ω) a symplectic manifold. ■

Definition 13 A **complex manifold** of complex dimension n is a real $2n$-dimensional real smooth manifold equipped with a **holomorphic structure**, an equivalence class of holomorphic atlases. A **holomorphic atlas** is an atlas $\{(U_i, \varphi_i)\}_{i \in I}$ on M with each φ_i a homoemorphism onto an open set $\varphi(U_i)$ in $\mathbb{C}^n$, usually taken to be the open unit disk, and such that the transition functions

$$\varphi_{ij} := \varphi_i \circ \varphi_j^{-1} : \varphi_j(U_i \cap U_j) \to \varphi_i(U_i \cap U_j)$$

are *biholomorphic*. Two holomorphic atlases $\{(U_i, \varphi_i)\}_{i\in I}$ and $\{(V_j, \psi_j)\}_{i\in J}$ are called *equivalent* if all maps $\varphi_i \circ \psi_j^{-1} : \psi_j(U_i \cap V_j) \to \varphi_i(U_i \cap V_j)$ are holomorphic.

Define the differential operators

$$\frac{\partial}{\partial z_j} := \frac{1}{2}\left(\frac{\partial}{\partial x_j} - i\frac{\partial}{\partial y_j}\right), \qquad \frac{\partial}{\partial \overline{z}_j} := \frac{1}{2}\left(\frac{\partial}{\partial x_j} + i\frac{\partial}{\partial y_j}\right)$$

on $C^\infty(\mathbb{C}^n, \mathbb{C})$, and their associated dual forms,

$$dz_j := dx_j + idy_j, \qquad d\overline{z}_j := dx_j - idy_j$$

If $U \subseteq \mathbb{C}^n$, then by Huybrechts [59, Proposition 1.3.1] the complexified tangent bundle $TU^{\mathbb{C}} := TU \otimes_{\mathbb{R}} \mathbb{C} := \bigsqcup_{p\in U} T_pU \otimes_{\mathbb{R}} \mathbb{C}$ splits as

$$TU^{\mathbb{C}} = T^{1,0}U \otimes T^{0,1}U$$

where $T^{1,0}$ is the $+1$-eigenbundle and $T^{0,1}$ is the -1-eigenbundle of the natural complex structure i on $TU^{\mathbb{C}}$ (see Section 2.3.1 for the algebraic necessaries), with $(\partial/\partial z_1, \dots, \partial/\partial z_n)$ a global frame for $T^{1,0}$ and $(\partial/\partial \overline{z}_1, \dots, \partial/\partial \overline{z}_n)$ a global frame for $T^{0,1}$. The complexified cotangent bundle $T^*U^{\mathbb{C}}$ splits analogously,

$$T^*U^{\mathbb{C}} = (T^*U)^{1,0} \oplus (T^*U)^{0,1}$$

and $(dz_1, \dots, dz_n)$ is a global frame for $(T^*U)^{1,0}$ and $(d\overline{z}_1, \dots, d\overline{z}_n)$ is a global frame for $(T^*U)^{0,1}$. Let

$$\Omega^k_{\mathbb{C}}(U) := \Gamma(U, \bigwedge^k TU^{\mathbb{C}})$$

denote the space of complex k-forms on U, which splits as

$$\Omega^k_{\mathbb{C}}(U) = \bigoplus_{p+q=k} \Omega^{p,q}(U)$$

where $\Omega^{p,q}(U)=\Gamma\big(U,(\bigwedge^p(T^*U)^{1,0})\otimes(\bigwedge^q(T^*U)^{0,1})\big)$. Let $\Pi^{p,q}:\Omega^k_{\mathbb{C}}(U)\to\Omega^{p,q}(U)$ be the (p,q)th projection, and note that the ordinary real exterior derivative $d:\Omega^\bullet(U)\to\Omega^{\bullet+1}(U)$ extends to the complexification $\Omega^\bullet_{\mathbb{C}}(U)$ by $d(\alpha\otimes z):=(d\alpha)\otimes z$, and satisfies

$$d:\Omega^{p,q}\to\bigoplus_{r+s=p+q+1}\Omega^{r,s}$$

$$d\alpha=\sum_{r+s=p+q+1}\Pi^{r,s}\circ d\alpha$$

By means of the complexified d we may define the **Dolbeault operators**,

$$\partial:=\Pi^{p+1,q}\circ d:\omega^{p,q}_{\mathbb{C}}(U)\to\omega^{p+1,q}_{\mathbb{C}}(U)$$

$$\overline{\partial}:=\Pi^{p,q+1}\circ d:\omega^{p,q}_{\mathbb{C}}(U)\to\omega^{p,q+1}_{\mathbb{C}}(U)$$

which satisfy (Huybrechts [59, Lemma 1.3.6])

$$d=\partial+\overline{\partial}$$

$$\partial^2\overline{\partial}^2=0$$

$$\partial(\alpha\wedge\beta)=\partial\alpha\wedge\beta+(-1)^{p+q}\alpha\wedge\partial\beta$$

$$\overline{\partial}(\alpha\wedge\beta)=\overline{\partial}\alpha\wedge\beta+(-1)^{p+q}\alpha\wedge\overline{\partial}\beta$$

If H is the standard Hermitian inner product on $TU=U\times\mathbb{C}^n$, then letting $g=\frac{1}{2}(H+\overline{H})$ and $\omega:=\frac{i}{2}(H-\overline{H})$ be the induced Riemannian metric and symplectic form, respectively, so that $H=g-i\omega$, then ω may be expressed in complex coordinates as

$$\omega=\frac{i}{2}\sum_{i,j=1}^n H_{ij}dz_j\wedge d\overline{z}_j$$

where $H_{ij}=H(\frac{\partial}{\partial x_i},\frac{\partial}{\partial y_j})$, and $d\omega=0$ iff for any point $\mathbf{z}\in U$ there exist a neighbourhood U' of $0\in\mathbb{C}^n$ and a local biholomorphic map $f:U'=f(U')\subseteq U$ with $f(0)=\mathbf{z}$ and such that f^*g osculates in the origin to order two to the standard metric (Huybrechts [59, Proposition 1.3.12]).

To put all of this on the complex manifold M we need to define the **holomorphic tangent bundle**. Let $\psi_{ij}:U_i\cap U_j\to\mathrm{GL}(n,\mathbb{C})$ be the cocycle gotten from the transition map $\varphi_{ij}=\varphi_i\circ\varphi_j^{-1}$

by application of the Jacobian of φ_{ij} at the point $\varphi_j(z) \in \varphi_j(U_i \cap U_j) \subseteq \mathbb{C}$, $J\varphi_{ij}(\varphi_j(z)) := \big(\frac{\partial \varphi_{ij}^k}{\partial z_\ell}\big)_{k,\ell}$:

$$\psi_{ij} : U_i \cap U_j \to \mathrm{GL}(n, \mathbb{C})$$

$$\psi_{ij}(\varphi_j(z)) := J\varphi_{ij}(\varphi_j(z))$$

These cocycles glue together to form a holomorphic vector bundle $\mathcal{T}_M \to M$, called the holomorphic tangent bundle. It is (complex-)isomorphic to $T^{1,0}M$ (Voisin [106, Proposition 2.13], Huybrechts [59, Proposition 2.6.4]), and in fact equal to it as a subbundle of $TM^{\mathbb{C}}$ if we view $TM^{1,0}$ as a complex vector bundle with complex structure J the natural complex structure i coming from the complexification of TM. For this reason, $TM^{1,0}$ is sometimes itself called the holomorphic tangent bundle. The realification of $\mathcal{T}_M$ is clearly real-isomorphic to the real tangent bundle TM of the underlying real manifold M:

$$\mathcal{T}_M \cong_{\mathbb{C}} TM^{1,0} \cong_{\mathbb{R}} TM$$

Given a complex manifold M with holomorphic tangent bundle $\mathcal{T}_M \cong_{\mathbb{C}} TM^{1,0}$, we know that $TM^{1,0}$ locally, over a chart domain U, possesses complex coordinates $(\partial/\partial z_1, \ldots, \partial/\partial z_n)$, and so over U we can apply the whole algebraic apparatus we constructed over $U \subseteq \mathbb{C}^n$: forms, complexified forms, their bi-grading, and the operators d, ∂, $\overline{\partial}$, as well as the form ω. All of these glue together over M, but some of the conditions on d, ∂, $\overline{\partial}$, J, and ω may fail to hold globally. We will describe these below. The main interest will be the condition on ω that it be *closed*, for in that case (M, ω) will also be symplectic. ■

Definition 14 A complex structure $J \in \Gamma(\mathcal{J}(TM))$ on a $2n$-dimensional real smooth manifold M is called **integrable** if it is induced by a holomorphic structure on M, that is if J is multiplication by i coordinate-wise in a way compatible with changes of coordinates. This is the case precisely when M is equipped with an atlas $\{(U_i, \varphi_i)\}_{i \in I}$ such that J is represented by $J_0 \in \mathcal{J}(\mathbb{R}^{2n})$ locally,

$$T_p\varphi_i \circ J_p = J_0 \circ T_p\varphi_i : T_pM \to \mathbb{R}^{2n}$$

(Salamon and McDuff [76, p. 123]). In this case, the almost complex manifold (M, J) acquires the structure of a complex manifold, and J is called a **complex structure**. ■

In order to understand this integrability condition, we need to define an ancillary tool, the Nijenhuis tensor.

Definition 15 Let $J \in \Gamma(\mathcal{J}(TM))$ be an almost complex structure on M. The **Nijenhuis tensor** N_J is a $(1,2)$-tensor field, given in terms of the Lie bracket on $\Gamma(TM)$,

$$N_J \in \mathcal{T}_2^1(M) = \Gamma(TM \otimes T^*M \otimes T^*M) \cong \mathrm{Hom}^2_{C^\infty(M)}(TM;TM)$$

$$N_J(X,Y) := [JX, JY] - J[JX,Y] - J[X,JY] - [X,Y]$$

More generally, given any $A \in \mathrm{End}\,TM$ we may define N_A as

$$N_A(X,Y) := [AX, AY] - A[AX,Y] - A[X,AY] + A^2[X,Y]$$

which is one-half of the Frölicher-Nijenhuis bracket, $N_A = \frac{1}{2}[A,A]_{FN}$. ■

The following proposition characterizes the integrability condition. It's proof may be found in Propositions 2.6.15 and 2.6.17, Corollary 2.6.18 of Huybrechts [59], except for the Newlander-Nirenberg (the equivalence of statements (1) and (5)), whose proof may be found in the original 1957 article of Newlander and Nirenberg [84], with a more modern proof in, e.g. Webster [108].

Proposition 51 *Let $J \in \Gamma(\mathcal{J}(TM))$ be an almost complex structure on an even-dimensional real smooth manifold M. Then the following are equivalent:*

(1) *J is integrable.*

(2) *$d = \partial + \overline{\partial}$ on $\Omega^\bullet_{\mathbb{C}}(M)$.*

(3) *$\Pi^{0,2} \circ d = 0$ on $\Omega^{1,0}(M)$.*

(4) *$[TM^{1,0}, TM^{1,0}] \subseteq TM^{1,0}$, that is the Lie bracket preserves the bundle $TM^{1,0}$.*

(5) *$N_J \equiv 0$. (Newlander-Nirenberg Theorem)*

Moreover, if $\overline{\partial}^2 = 0$, then J is integrable. Conversely, if J is integrable, then $\partial^2 = \overline{\partial}^2 = 0$ and $\partial \circ \overline{\partial} = -\overline{\partial} \circ \partial$. ■

Example 9 It is well known that the 6-sphere S^6 carries a non-integrable almost complex structure J (Calabi [23], Steenrod [96, Remark 41.21, p. 217]), but it remains an open question whether all $J \in \Gamma(\mathcal{J}(TS^6))$ are non-integrable. We know, of course, that S^6 admits no symplectic structure (Proposition 49), so in any case (S^6, J) is an example of an almost complex manifold, possibly a complex manifold, which is not a symplectic manifold. A recent paper of Gábor Etesi [31] claiming to prove the existence of a an integrable complex structure on S^6 has an erratum pending review at the time of this writing. ■

Definition 16 A **Kähler manifold** is a symplectic manifold (M, ω) which is also a complex manifold (M, J) with $J \in \Gamma(\mathcal{J}(TM, \omega))$ an integrable and compatible almost complex structure. Alternatively, a Kähler manifold is a complex manifold M endowed with a **Kähler structure**: a Riemannian metric g, which is compatible with the induced complex structure $J \in \Gamma(\mathcal{J}(TM))$, multiplication by i, meaning that the **fundamental form** ω is compatible, $\omega := g(J\cdot, \cdot)$, and closed $d\omega = 0$. The Riemannian metric g is here called a **hermitian structure**, and (M, g) a **hermitian manifold**. In this case $g = \omega(\cdot, J\cdot)$, too, by compatibility. ■

Example 10 There are many examples of complex manifolds which are symplectic yet not Kähler, beginning with those in Thurston's article [101]. Gompf [41] has examples of compact symplectic non-Kähler manifolds for every dimension ≥ 4, and Guan [49] has constructed examples of compact simply connected symplectic and complex but non-Kähler manifolds. Angella's recent book [2] contains many more, along with the cohomological theory underlying the production of many of these examples. ■

2.4.2.2 Examples of Symplectic Manifolds

Example 11 The standard vector space $(\mathbb{R}^{2n}, \omega_0)$ of Section 2.3.3 can be viewed as a symplectic manifold with trivial tangent bundle $T\mathbb{R}^{2n} = \mathbb{R}^{2n} \times \mathbb{R}^{2n}$. Here, we think of ω_0 not merely as an element of $\bigwedge^2(\mathbb{R}^{2n})^* \cong \mathrm{Hom}^2_{\mathbb{R},\mathrm{Skew}}(\mathbb{R}^{2n}; \mathbb{R}) \cong \mathrm{Skew}_{2n}(\mathbb{R})$, but as a section of the bundle $\bigwedge^2 T^*\mathbb{R}^{2n} \cong$

$\mathrm{Hom}^2_{\mathbb{R},\mathrm{Skew}}(T\mathbb{R}^{2n};\mathbb{R}) \to \mathbb{R}^{2n}$, i.e.

$$\omega_0 \in \Omega^2(\mathbb{R}^{2n}) = \Gamma\Big(\bigwedge^2 T^*\mathbb{R}^{2n}\Big)$$

In other words, ω_0 sends a point $\mathbf{x} \in \mathbb{R}^{2n}$ to a symplectic form in $T_{\mathbf{x}}\mathbb{R}^{2n} = \{\mathbf{x}\} \times \mathbb{R}^{2n}$,

$$\omega_{0_{\mathbf{x}}} = (\mathbf{x},\ \omega_0) = \Big(\mathbf{x},\ \sum_{j=1}^{n} dq_j \wedge dp_j\Big)$$

as an element of $\{\mathbf{x}\} \times \bigwedge^2(\mathbb{R}^{2n})^* = \bigwedge^2 T^*_{\mathbf{x}}\mathbb{R}^{2n}$. Notice that ω_0 is closed, since the exterior derivative of each $dq_i \wedge dp_j$ is 0.

Considering $\mathbb{R}^{2n}$ as a manifold rather than just a vector space already introduces new features. For example, the class of morphisms is much larger. Let us identify $T_{\mathbf{x}}\mathbb{R}^{2n}$ with $\mathbb{R}^{2n}$ to simplify notation, and consider the following example of a nonlinear symplectic morphism:

$$f : \mathbb{R}^2 \to \mathbb{R}^2$$
$$f(x, y) = (x^2 + x + y,\ x^2 + y)$$

Its tangent map is

$$Df(x, y) = \begin{pmatrix} 2x+1 & 1 \\ 2x & 1 \end{pmatrix}$$

which is easily seen to be symplectic on $\mathbb{R}^2$—it satisfies $Df(x, y)^T J_0 Df(x, y) = J_0$ (see Proposition 19). It is, of course, not a symplectomorphism because it fails to be injective. An example of a symplectomorphism is

$$f : \mathbb{R}^2 \to \mathbb{R}^2$$
$$f(x, y) = (e^x + x + y,\ e^x + y)$$

whose tangent map is

$$Df(x, y) = \begin{pmatrix} e^x+1 & 1 \\ e^x & 1 \end{pmatrix}$$

which is easily verified to be an element of $\mathrm{Sp}(1,\mathbb{R})$ for all $(x,y)\in\mathbb{R}^2$. It is, moreover, invertible, for we can solve the equations $u=e^x+x+y$, $v=e^x+y$, namely $x=u-v$ and $y=v-e^{u-v}$, so

$$f^{-1}(u,v)=(u-v,v-e^{u-v})$$
$$D(f^{-1})(u,v)=\begin{pmatrix}1 & -1\\ -e^{u-v} & 1+e^{u-v}\end{pmatrix}$$

which is also easily seen to be symplectic. This shows that the symplectomorphism group

$$\mathrm{Sp}(\mathbb{R}^2,\omega_0)=\{f\in\mathrm{Diff}(\mathbb{R}^2)\mid Df(\mathbf{x})\in\mathrm{Sp}(n,\mathbb{R}),\ \forall\mathbf{x}\in\mathbb{R}^2\}$$

is strictly larger than the linear symplectomorphism group $\mathrm{Sp}(1,\mathbb{R})$. ■

Example 12 More generally, any symplectic vector space (V,ω) can be viewed as a symplectic manifold with trivial tangent bundle $TV=V\times V$ and $\omega\in\Omega^2(V)=\Gamma(\bigwedge^2 T^*V)\cong\mathrm{Hom}^2_{\mathbb{R},\mathrm{Skew}}(T^*V;\mathbb{R})$, $v\mapsto\omega_v:=\{v\}\times\omega\in T_vV=\{v\}\times V$, and as with $\mathbb{R}^{2n}$, to which V is linearly symplectomorphic, we will have a much larger class of morphisms in the manifold category over the vector space category. ■

Example 13 The 2-sphere S^2 may be endowed with a symplectic form ω as follows. View S^2 as embedded in $\mathbb{R}^3$, so that the tangent space $T_{\mathbf{p}}S^2$ above a point $\mathbf{p}\in S^2$ may be identified with the subspace of the tangent space $T_{\mathbf{p}}\mathbb{R}^3=\{\mathbf{p}\}\times\mathbb{R}^3$, whereby we also get the subsapce normal to $T_{\mathbf{p}}S^2$, $N_{\mathbf{p}}S^2:=(T_{\mathbf{p}}S^2)^\perp$, which can be identified with the span of the vector $(\mathbf{p},\mathbf{p})$ in $\{\mathbf{p}\}\times\mathbb{R}^3$, as

$$N_{\mathbf{p}}S^2=\{\mathbf{p}\}\times\mathrm{span}(\mathbf{p})$$

Then, given any two vectors in $(\mathbf{p},\mathbf{u})$, $(\mathbf{p},\mathbf{v})\in T_{\mathbf{p}}S^2$, we have that $(\mathbf{p},\mathbf{u}\times\mathbf{v})\in N_{\mathbf{p}}S^2$, and we can take the (standard $\mathbb{R}^3$) inner product of this vector with $(\mathbf{p},\mathbf{p})$,

$$\langle(\mathbf{p},\mathbf{p}),(\mathbf{p},\mathbf{u}\times\mathbf{v})\rangle_{\mathbf{p}}:=\langle\mathbf{p},\mathbf{u}\times\mathbf{v}\rangle$$

which is positive whenever $\mathbf{u}$ and $\mathbf{v}$ are nonzero and linearly independent, and represents the volume of the parallelpiped $P(\mathbf{p},\mathbf{u},\mathbf{v})$ spanned by $\mathbf{p}$, $\mathbf{u}$ and $\mathbf{v}$—in fact it equals the absolute value of the determinant of the 3×3 matrix A_P with columns are these three vectors,

$$\langle \mathbf{p},\mathbf{u}\times\mathbf{v}\rangle = \mathrm{Vol}(P(\mathbf{p},\mathbf{u},\mathbf{v})) = |\det A_P| = \det\sqrt{A_P^TA_P} = \|\mathbf{u}\times\mathbf{v}\| = \mathrm{Area}(\mathbf{u},\mathbf{v})$$

which, since $\mathbf{p}\perp\mathrm{span}(\mathbf{u},\mathbf{v})$ and $\|\mathbf{p}\|=1$, also gives the area of the parallelogram spanned by $\mathbf{u}$ and $\mathbf{v}$. Using this inner product we define the symplectic form

$$\omega\in\Omega^2(S^2)$$
$$\omega_{\mathbf{p}}\big((\mathbf{p},\mathbf{p}),\ (\mathbf{p},\mathbf{u}\times\mathbf{v})\big) := \langle(\mathbf{p},\mathbf{p}),(\mathbf{p},\mathbf{u}\times\mathbf{v})\rangle_{\mathbf{p}} = \big|\det\big((\mathbf{p}\ \ \mathbf{u}\ \ \mathbf{v})\big)\big| = \|\mathbf{u}\times\mathbf{v}\|$$

and we notice that the skew-symmetry of the cross product ensures the skew-symmetry of ω, while nondegeneracy can be shown by choosing a symplectic basis: let $\mathbf{u}\in\mathbf{p}^\perp$ be any unit vector, and define $\mathbf{v}=\mathbf{u}\times\mathbf{p}$. Then, $\omega(\mathbf{u},\mathbf{u})=\omega(\mathbf{v},\mathbf{v})=0$, while $\omega(\mathbf{u},\mathbf{v})=\langle\mathbf{p},\mathbf{u}\times(\mathbf{u}\times\mathbf{p})\rangle=\langle\mathbf{p},-\mathbf{p}\rangle=-1$, whence $[\omega]_{(\mathbf{u},\mathbf{v})}=J_0$. Note that ω is just the standard area form, whose integral gives the surface area of the sphere,

$$\int_{S^2}\omega = 4\pi$$

For if we choose $\mathbf{u}$ latitudinal, that is tangent to the latitudinal circles and so represented by the vector field ∂_θ, then $\mathbf{v}=\mathbf{u}\times\mathbf{p}$ is the tangent vector to a longitudinal great circles and so represented by ∂_φ, where (θ,φ) are the standard spherical coordinates, $0<\theta<2\pi$, $0<\varphi<\pi$. Of course, the collection of points $(\sin\varphi,0,\cos\varphi)$ for $0\le\varphi\le\pi$ are not covered by this chart, but this is a set of measure zero so does not affect the integral.

Now, consider a symplectomorphism $\varphi\in\mathrm{Sp}(S^2,\omega)$. The requirement $\varphi^*\omega=\omega$ means $\omega_{\varphi(\mathbf{p})}(T_{\mathbf{p}}\varphi(\mathbf{u}),T_{\mathbf{p}}\varphi(\mathbf{v}))=\omega_{\mathbf{p}}(\mathbf{u},\mathbf{v})$, which is equivalent to the area of the parallelogram spanned by $T_{\mathbf{p}}\varphi(\mathbf{u})$ and $T_{\mathbf{p}}\varphi(\mathbf{v})$ being the same as the area of the parallelogram spanned $\mathbf{u}$ and $\mathbf{v}$. Thus,

$$\mathrm{Sp}(S^2,\omega)=\{\text{area- and orientation-preserving }\varphi\in\mathrm{Diff}(S^2)\}$$

Of course, $\mathrm{SO}(3)\subseteq\mathrm{Sp}(S^2,\omega)$, since any special orthogonal matrix A is just a rotation matrix, so is its own derivative and therefore preserves areas and orientation. Indeed, since

$$T_{\mathbf{p}}A : T_{\mathbf{p}}S^2 \to T_{A\mathbf{p}}S^2$$

$$T_{\mathbf{p}}A(\mathbf{p},\mathbf{u}) = (A\mathbf{p}, A\mathbf{u})$$

and A preserves the dot product, and therefore angles and lengths, we have

$$A^*\omega_{\mathbf{p}}(\mathbf{u},\mathbf{v}) = \omega_{A\mathbf{p}}(A\mathbf{u},A\mathbf{v}) = \|(A\mathbf{u})\times(A\mathbf{v})\|$$

$$= \|A\mathbf{u}\|\|A\mathbf{v}\|\sin\theta = \|\mathbf{u}\|\|\mathbf{u}\|\sin\theta = \|\mathbf{u}\times\mathbf{v}\| = \omega_{\mathbf{p}}(\mathbf{u},\mathbf{v})$$

However, the group $\mathrm{Sp}(S^2,\omega)$ is strictly larger than $\mathrm{SO}(3)$. ■

Example 14 The sphere was an example of an orientable surface. More generally, any oriented surface S is a symplectic manifold (S, dA), with symplectic form its area form dA. For $dA \in \Omega^2(S)$ is trivially closed by dimension count, and it is nondegenerate by reason of orientability, which requires the area form to be a nonvanishing top-form. This produces a large class of 2-dimensional examples. ■

Example 15 Consider the 2-torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$. We know from the previous example that it is a symplectic manifold, being an orientable surface, whether we view it as a quotient $\mathbb{R}^2/\mathbb{Z}^2$, a product $S^1 \times S^1$, or an embedding into $\mathbb{R}^3$. But viewing it as a quotient gives an explicit description of the symplectic form ω. Namely, take the standard symplectic form ω_0 on $\mathbb{R}^2$ and note that it descends to the quotient because it is translation-invariant on $\mathbb{R}^2$, which has trivial tangent bundle on which ω_0 is constant, by Example 11. Alternatively, viewing $\mathbb{T}^2$ as $S^1 \times S^1$, we may define $\omega = d\theta \wedge d\varphi$[7] , though note that this expression suffers from the restrictions on θ and φ which require them to lie in $(0, 2\pi)$ and so not covering the points $(1, e^{i\varphi})$, $(e^{i\theta}, 1) \in \mathbb{T}^2$.

The same constructions work to make the $2n$-torus $\mathbb{T}^{2n} = \mathbb{R}^{2n}/\mathbb{Z}^{2n} \approx S^1 \times \cdots \times S^1$ symplectic. Viewing it as a lattice, we let ω_0 descend from $\mathbb{R}^{2n}$ to the quotient, while viewing it as the product

[7] Strictly speaking, $\omega_{(e^{i\theta}, e^{i\varphi})} = \big((e^{i\theta}, e^{i\varphi}),\ d\theta \wedge d\varphi\big)$

of $2n$ circles, $\mathbb{T}^{2n} = (\mathbb{T}^2)^n$, we let $\omega = \pi_1^*(d\theta_1 \wedge d\varphi_1) + \cdots + \pi_n^*(d\theta_n \wedge d\varphi_n)$, where $\pi_j : \mathbb{T}^{2n} \to \mathbb{T}_j^2$ is the projection onto the jth factor. This is a general type of construction for products of symplectic manifolds, which we describe next. ■

Example 16 Let (M, ω_M) and (N, ω_N) be symplectic manifolds of dimensions $2m$ and $2n$, respecively. Then their product $M \times N$ is also a symplectic manifold, with **product symplectic form**

$$\omega_M \oplus \omega_N := \pi_1^* \omega_M + \pi_2^* \omega_N$$

where $\pi_1 : M \times N \to M$ and $\pi_2 : M \times N \to N$ are the projections onto each factor. Skew-symmetry is clear, and nondegeneracy follows from the observation that

$$\frac{1}{(m+n)!} \omega_M \oplus \omega_N = \frac{1}{(m)!(n)!} (\pi_1^* \omega_M)^m \wedge (\pi_2^* \omega_N)^n$$

and this latter form is nondegenerate because each of $(\pi_1^* \omega_M)^m$ and $(\pi_2^* \omega_N)^n$ are.

We also have a **twisted product symplectic form**

$$\omega_M \ominus \omega_N := \pi_1^* \omega_M - \pi_2^* \omega_N$$

which will be needed in the study of symplectomorphisms of a symplectic manifold.

These forms allows us to construct new symplectic manifolds out of old. For example, $S^2 \times \mathbb{T}^2$ and $S^2 \times S^2 \times \mathbb{T}^2$ are symplectic manifolds. ■

Example 17 Not every complex manifold M is a symplectic manifold. If M is complex, then it is also Hermitian (Remark 29), so TM has a Hermitian inner product H from which we get a Riemannian metric $g := \frac{1}{2}(H + \overline{H})$ and a nondegenerate skew-symmetric 2-form $\omega := -\frac{i}{2}(H - \overline{H})$, called the **fundamental form**, and clearly satisfying $H = g + i\omega$. However, ω is not necessarily closed.

Example 18 (The Cotangent Bundle) The canonical example of a symplectic manifold is the cotangent bundle T^*M of an n-manifold M. Consider the projection map $\pi : T^*M \to M$ and it's tangent map $T\pi : T(T^*M) \to TM$. Define a 1-form $\eta \in \Omega^1(T^*M)$, called the **tautological 1-form** (or **canonical** or **symplectic** or **Liouville** or **Poincaré 1-form**), as follows: if $\xi \in T(T^*M)$ and $\alpha \in T^*M$ is a 'point' in the cotangent bundle, then $\xi_\alpha \in T_\alpha(T^*M)$, and $T_\alpha\pi : T_\alpha(T^*M) \to T_{\pi(\alpha)}M$, so let η act on ξ_α as

$$\eta_\alpha(\xi_\alpha) := (\pi^*\alpha)_\alpha(\xi_\alpha) = \alpha(T_\alpha\pi(\xi_\alpha))$$

If (U, φ) be a chart on M, with $\varphi = (q_1, \dots, q_n)$ in components (we think of the q_i as position coordinates), then $\Phi = (q_1, \dots, q_n, p_1, \dots, p_n)$ is a chart on T^*M. A point $\alpha \in T^*M$ then looks locally like $\alpha = (a, b) := \sum_{i=1}^n a_i q_i + b_i p_i$, where $a_i, b_j \in \mathbb{R}$, and a tangent vector $\xi_\alpha \in T_\alpha(T^*M)$ then is expressed locally as $\xi_\alpha = (A, B) := \sum_{i=1}^n A_i \frac{\partial}{\partial q^i} + B_i \frac{\partial}{\partial p^i}$. Consequently, $T_\alpha\pi(\xi_\alpha) = \sum_{i=1}^n A_i \frac{\partial}{\partial q^i}$, and so, since $p_i = dq^i$, we have

$$\eta_\alpha(\xi_\alpha) = \alpha\left(\sum_{i=1}^n A_i \frac{\partial}{\partial q^i}\right) = \sum_{j=1}^n (a_j q_j + b_j dq^j)\left(\sum_{i=1}^n A_i \frac{\partial}{\partial q^i}\right) = \sum_{i=1}^n b_i A_i$$

Therefore,

$$\eta = p\, dq := \sum_{i=1}^n p_i\, dq^i$$

is the coordinate expression of the tautological 1-form. ■

2.4.3 Darboux's Theorem

Recall Proposition 32, which says that all symplectic vector spaces (V, ω) of dimension $2n$ are linearly symplectomorphic to $(\mathbb{R}^{2n}, \omega_0)$. Our first major result is a generalization of this statement to manifolds: all symplectic manifolds are *locally* symplectomorphic to the symplectic *manifold* $(\mathbb{R}^{2n}, \omega_0)$. This means that about every point p in M there are *canonical* or *symplectic coordinates*, i.e. there is a chart (U, φ) with $\varphi = (x_1, \dots, x_n, y_1, \dots, y_n) : U \to \varphi(U) \subseteq \mathbb{R}^{2n}$ such that ω has the local expression $\omega = \sum_{i=1}^n dx^i \wedge dy^i$ in these coordinates, and moreover φ is a symplectomorphism, i.e. $\varphi^*\omega_0 = \omega$. This is the content of Darboux's theorem.

Lemma 6 *Let M and N be smooth manifolds and let $\{F_t : M \to N \mid t \in \mathbb{R}\}$ be a smooth 1-parameter family of diffeomeorphisms ($F : M \times \mathbb{R} \to N$ is a smooth map). Let $X_t \in \mathfrak{X}_{F_t}(M)$ be the tangent field along F_t, that is find X_t making the following diagram commute,*

$$\begin{tikzcd} & TN \arrow[d, "\pi_N"] \\ M \arrow[ur, "X_t"] \arrow[r, "F_t"'] & N \end{tikzcd}$$

That is, $X_t(p)$ is the tangent vector to the curve $c : \mathbb{R} \to N$ given by $c(s) := F_s(p)$ at the point $q = F_t(p)$ in N. Then, if $(\alpha_t)_{t\in\mathbb{R}} \in \Omega(N) \times \mathbb{R}$ is a 1-parameter family of differential forms on N, we have

$$\begin{aligned} \frac{d}{dt}(F_t^*\alpha_t) &= F_t^*\left(\frac{d\alpha_t}{dt} + i_{X_t} d\alpha_t\right) + d\big(F_t^*(i_{X_t}\alpha_t)\big) \\ &= F_t^*\left(\frac{d\alpha_t}{dt} + (i_{X_t} \circ d + d \circ i_{X_t})\alpha_t\right) \end{aligned}$$

When $M = N$, F_t is the flow of $X \in \mathfrak{X}(M)$, and α_t is independent of t, i.e. $\alpha_t = \alpha \in \Omega(M)$, the last equality at $t = 0$ may be formulated in terms of the Lie derivative as Cartan's magic formula, $\frac{d}{dt}\big|_{t=0}(F_t^*\alpha) = \mathcal{L}_X\alpha = (i_X \circ d + d \circ i_X)\alpha$.

Proof: This is Lemma 2.3, Berndt [16]. ■

Theorem 17 (Darboux's Theorem) *Let M be a $2n$-dimensional smooth manifold and let $\omega_0, \omega_1 \in \Omega^2(M)$ be symplectic forms agreeing at some point $p \in M$. Then there is a neighborhood U of p and a diffeomorphism $F : U \to F(U)$ fixing p and satisfying $F^*\omega_1 = \omega_0$ on U.*

Proof: Define $\omega_t := (1-t)\omega_0 + t\omega_1$ for all $t \in I = [0,1]$, and note that it is closed since ω_0 and ω_1 are. Define $\sigma := \omega_0 - \omega_1 = -\frac{d}{dt}\omega_t$. Since σ is also closed, Poincaré's Lemma implies that it is locally exact, i.e. there exists a neighborhood U_1 of p and a 1-form $\alpha \in \Omega^1(U_1)$ such that $d\alpha = \sigma$ on U_1. Since $\omega_t(p) = \omega_0(p)$ for all $t \in I$ and ω_0 is nondegenerate, so is ω_1 on a neighborhood U_0 of p, which we may suppose lies inside U_1. Use the nondegeneracy of ω_t on U_0 to obtain the vector field $Y_t := \omega_t^\sharp(\alpha)$, which means $i_{Y_t}\omega_t = \alpha$ on U_0. Then the flows F_t of the vector fields Y_t exist on a neighborhood $U \subseteq U_0$ of p, and satisfy $F_t(U) \subseteq U_0$ and $F_t(p) = p$. Moreover, by construction we have $\frac{d\omega_t}{dt} + d(i_{Y_t}\omega_t) = \frac{d\omega_t}{dt} + d\alpha = \frac{d\omega_t}{dt} + \sigma = 0$ and by the closedness of ω_t and Lemma 6 applied to the F_t and ω_t we see that

$$\frac{d}{dt}\big(F_t^*\omega_t\big) = F_t^*\left(\frac{d\omega_t}{dt} + i_{Y_t}(d\omega_t) + d(i_{Y_t}\omega_t)\right) = F_t^*0 = 0$$

Consequently, since $F_0 = \mathrm{id}_U$, we have that $F_t^*\omega_t = F_0^*\omega_0 = \omega_0$ for all t, including $t = 1$. ■

Corollary 9 *If (M, ω) is a symplectic manifold, then each point $p \in M$ has an open neighborhood U which is symplectomorphic to an open subset of $(\mathbb{R}^{2n}, \omega_0)$. This means that about every point $p \in M$ there is a chart (U, φ) with $\varphi = (x_1, \dots, x_n, y_1, \dots, y_n) : U \to \varphi(U) \subseteq \mathbb{R}^{2n}$ such that*

$$\omega = \sum_{i=1}^{n} dx^i \wedge dy^i$$

Proof: If (U, φ) is a chart about p, then define $\psi := \varphi^{-1}$ and define the symplectic form $\omega_1 := \psi^*\omega_p$ on $U_1 := \varphi(U) \subseteq \mathbb{R}^{2n}$. By applying a linear transformation if necessary we may assume that U_1 is a neighborhood of 0 in $\mathbb{R}^{2n}$ and $\omega_1(0) = \omega_0(0)$. Then use Darboux's theorem to get a neighborhood U_0 of $0 \in \mathbb{R}^{2n}$ and diffeomorphism $F_0 : U_0 \to F(U_0) \subseteq U_1$ such that $F_0^*\omega_1 = \omega_0$ and $F_0(0) = 0$. Finally, set $F := (\psi \circ F_0)^{-1} : V \to F(V) = U_0$, where $V = \psi(F_0(U_0))$, which is the desired symplectomorphism and chart (V, F). ■

2.4.4 The Poisson Bracket

Let (M, ω) be a symplectic manifold. Recall the isomorphisms $\omega^\flat : \mathfrak{X}(M) \to \Omega^1(M)$ and $\omega^\sharp = (\omega^\flat)^{-1} : \Omega^1(M) \to \mathfrak{X}(M)$, given, as in the linear case, by $\omega^\flat(X) := i_X\omega = \omega(X, \cdot)$, which we also denote by $X^\flat$, and we also write $\alpha^\sharp$ for $\omega^\sharp(\alpha)$. Using these, we define the **Poisson bracket of two 1-forms** $\alpha, \beta \in \Omega^1(M)$ by

$$\{\alpha, \beta\} := -[\alpha^\sharp, \beta^\sharp]^\flat = -i_{[\alpha^\sharp, \beta^\sharp]}\omega \tag{2.109}$$

where $[\cdot, \cdot]$ is the Lie bracket on $\mathfrak{X}(M)$. As a consequence, $\{\cdot, \cdot\}$ is a Lie bracket on $\Omega^1(M)$.

Theorem 18 *For all $\alpha, \beta \in \Omega^1(M)$ we have the following characterization of the Poisson bracket in terms of the Lie, exterior, and interior derivatives:*

$$\{\alpha, \beta\} = -\mathcal{L}_{\omega^\sharp(\alpha)}\beta + \mathcal{L}_{\omega^\sharp(\beta)}\alpha + d(i_{\omega^\sharp(\alpha)} i_{\omega^\sharp(\beta)}\omega) \tag{2.110}$$

Proof: This follows from the fact that ω is closed, so that for $X_1, X_2, X_3 \in \mathfrak{X}(M)$

$$0 = d\omega(X_1, X_2, X_3) = \sum_{i=1}^{3}(-1)^{i-1}X_i\big(\omega(\dots \hat{X}_i \dots)\big) + \sum_{i<j}(-1)^{i+j}\omega\big([X_i, X_j], X_k\big)$$

Taking $X_1 = \alpha^\sharp := \omega^\sharp(\alpha)$, $X_2 = \beta^\sharp := \omega^\sharp(\beta)$, we get that

$$\begin{aligned} 0 &= (-1)^{1-1}X_1\big(\omega(X_2, X_3) + (-1)^{2-1}X_2\big(\omega(X_1, X_3)\big) + (-1)^{3-1}X_3\big(\omega(X_1, X_2)\big) \\ &\quad +(-1)^{1+2}\omega\big([X_1, X_2], X_3\big) + (-1)^{1+3}\omega\big([X_1, X_3], X_2\big) + (-1)^{2+3}\omega\big([X_2, X_3], X_1\big) \\ &= \mathcal{L}_{\alpha^\sharp}\big(i_{\beta^\sharp}\omega(X_3)\big) - \mathcal{L}_{\beta^\sharp}\big(i_{\alpha^\sharp}\omega(X_3)\big) - \mathcal{L}_{X_3}\big(i_{\alpha^\sharp} i_{\beta^\sharp}\omega\big) \\ &\quad -\omega\big([\alpha^\sharp, \beta^\sharp], X_3\big) - \omega\big(\beta^\sharp, [\alpha^\sharp, X_3]\big) - \omega\big([X_1, X_2], X_3\big) - \omega\big([\beta^\sharp, X_3], \alpha^\sharp\big) \\ &= \mathcal{L}_{\alpha^\sharp}\big(\beta(X_3)\big) - \mathcal{L}_{\beta^\sharp}\big(\alpha(X_3)\big) - \mathcal{L}_{X_3}\big(i_{\alpha^\sharp} i_{\beta^\sharp}\omega\big) \\ &\quad +\{\alpha, \beta\}(X_3) - \beta\big(\mathcal{L}_{\alpha^\sharp}X_3\big) + \alpha\big(\mathcal{L}_{\beta^\sharp}X_3\big) \\ &= \mathcal{L}_{\alpha^\sharp}\big(\beta(X_3)\big) - \mathcal{L}_{\beta^\sharp}\big(\alpha(X_3)\big) + \{\alpha, \beta\}(X_3) - d(i_{\alpha^\sharp} i_{\beta^\sharp}\omega)(X_3) \end{aligned}$$

since, e.g. $i_{\alpha^\sharp}\omega = \omega^\flat(\alpha^\sharp) = \omega^\flat(\omega^\sharp(\alpha)) = \alpha$. This is true for all X_3. ■

2.4.4.1 The Poisson Bracket on Smooth Functions on $(\mathbb{R}^{2n}, \omega_0)$

Let us now look at the special case of $(\mathbb{R}^{2n}, \omega_0)$. Let $f \in C^\infty(\mathbb{R}^{2n}) \cong C^\infty(T^*\mathbb{R}^n)$, which in this context we call an **observable** (cf. Folland [34]). Then $df \in \Omega^1(\mathbb{R}^{2n})$, so we can use $\omega_0^\sharp$ to get a vector field $X_f \in \mathfrak{X}(\mathbb{R}^{2n})$,

$$X_f := \omega_0^\sharp(df), \qquad \text{so that} \quad df = -i_{X_f}\omega_0 \tag{2.111}$$

and therefore, for all $Y \in \mathfrak{X}(M)$, $df(Y) = -i_Y i_{X_f}\omega_0 = \omega_0^\flat(X_f)(Y) = \omega_0(Y, X_f)$. We can also define the **Poisson bracket** $\{f, g\}$ **of two functions/observables** $f, g \in C^\infty(\mathbb{R}^{2n}) \cong C^\infty(T^*\mathbb{R}^n)$, as follows,

$$\{f, g\} := -i_{X_f} i_{X_g]}\omega_0 = \omega_0(X_f, X_g) \tag{2.112}$$

Let us see what X_f and $\{f,g\}$ looks like in symplectic coordinates, $(x_1,\dots,x_n,y_1,\dots,x_n)$.

Proposition 52 *If $f\in C^\infty(\mathbb{R}^{2n})$ then in canonical symplectic coordinates we have*

$$X_f=\frac{\partial f}{\partial x^i}\frac{\partial}{\partial y^i}-\frac{\partial f}{\partial y^i}\frac{\partial}{\partial x^i} \tag{2.113}$$

where we used Einstein summation notation.

Proof: If we write $X_f=X^i\frac{\partial}{\partial x_i}+X^{n+i}\frac{\partial}{\partial y_i}$ and $Y=Y^i\frac{\partial}{\partial x_i}+Y^{n+i}\frac{\partial}{\partial y_i}$, then

$$df(Y)=\left(\frac{\partial f}{\partial x^i}dx^i+\frac{\partial f}{\partial y^i}dy^i\right)\left(Y^j\frac{\partial}{\partial x^j}+Y^{n+j}\frac{\partial}{\partial y^j}\right)=Y^i\frac{\partial f}{\partial x^i}+Y^{n+i}\frac{\partial f}{\partial y^i}$$

On the other hand,

$$\omega_0(Y,X_f)=(dx^i\wedge dy^i)\left(Y^j\frac{\partial}{\partial x^j}+Y^{n+j}\frac{\partial}{\partial y^j},\ X^k\frac{\partial}{\partial x^k}+X^{n+k}\frac{\partial}{\partial y^k}\right)=Y^iX^{n+i}-Y^{n+i}X^i$$

These two expressions are equal, since $df(Y)=\omega_0(Y,X_f)$, so comparing components we see that $X^i=-\frac{\partial f}{\partial y^i}$ and $X^{n+i}=\frac{\partial f}{\partial x^i}$. ■

Corollary 10 *For all $f,g\in C^\infty(\mathbb{R}^{2n})$ we have the coordinate expression in symplectic coordinates*

$$\{f,g\}=\omega_0(X_f,X_g)=\frac{\partial f}{\partial x^i}\frac{\partial g}{\partial y^i}-\frac{\partial f}{\partial y^i}\frac{\partial g}{\partial x^i} \tag{2.114}$$

Proof: This follows from the previous corollary, using the coordinate expressions for X_f and $Y=X_g$ and the formula $\omega_0(X_f,X_g)=Y^i\frac{\partial f}{\partial x^i}+Y^{n+i}\frac{\partial f}{\partial y^i}$. ■

Remark 34 Let $x^i,y^j\in(\mathbb{R}^{2n})^*$ be the dual vectors of the symplectic basis vectors of $\mathbb{R}^{2n}$. Then, $x^i,y^j\in C^\infty(\mathbb{R}^{2n})$ as well, so we may consider their Poisson bracket, $\{x^i,y^j\}=\omega_0(X_{x^i},Y_{y^j})=\omega_0(x_i,y_j)$. Then, clearly, the x^i and y^j satisfy the following relations:

$$\begin{aligned}\{x^i,x^j\}&=0\\ \{y^i,y^j\}&=0\\ \{x^i,y^j\}&=\delta_{ij}\end{aligned} \tag{2.115}$$

We shall return to these relations later. ■

2.4.4.2 The Poisson Bracket on Smooth Functions on (M, ω)

Let us now consider the general case of an arbitrary symplectic manifold (M, ω). We remark first that a smooth function $f \in C^\infty(M)$ is also here called an **observable** sometimes. The **Poisson bracket of two functions/observables** $f, g \in C^\infty(M)$ is then given as in the real case, namely by

$$\{f, g\} := -i_{X_f} i_{X_g} \omega = \omega(X_f, X_g) \tag{2.116}$$

where $X_f \in \mathfrak{X}(M)$ is the vector field associated to f given by

$$X_f := \omega_0^\sharp(df), \qquad \text{and so} \qquad df = -i_{X_f}\omega \tag{2.117}$$

Since by Darboux's theorem any symplectic manifold (M, ω) is locally symplectomorphic to $(\mathbb{R}^{2n}, \omega_0)$, the above local expressions (2.113) and (2.114) are also true for $f, g \in C^\infty(M)$.

We refer to Berndt [16] for proofs of the following facts, which are straightforward, if tedious, calculations.

Proposition 53 *Let (M, ω) be a symplectic manifold. If $f, g \in C^\infty(M)$, then*

$$\{f, g\} = -\mathcal{L}_{X_f} g = \mathcal{L}_{X_g} f \tag{2.118}$$

■

Corollary 11 *Let (M, ω) be a symplectic manifold. If $f, g \in C^\infty(M)$, then $\{f, g\} = 0$ iff f is constant on the integral curves of X_g iff g is constant on the integral curves of X_f.* ■

Proposition 54 *Let (M, ω) be a symplectic manifold. If $f, g \in C^\infty(M)$, then*

$$d\{f, g\} = \{df, dg\} \tag{2.119}$$

■

The Jacobi identity in the next proposition may be checked in local coordinates, or using the Lie derivative expressions introduced above (Proposition 53).

Proposition 55 *Let (M,ω) be a symplectic manifold. The Poisson bracket $\{\cdot,\cdot\}$ on $C^\infty(M)$ makes $C^\infty(M)$ into a Lie algebra. Moreover, if $f,g\in C^\infty(M)$, we have*

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

Theorem 19 *Let (M,ω_M) and (N,ω_N) be symplectic manifolds. If $F : M \to N$ is a diffeomeorphism, then F is symplectic, i.e. $F^*\omega_N = \omega_M$ iff the pullback $F^* : \Omega(N) \to \Omega(M)$, $F^*\alpha = \alpha\circ(TF\times\cdots\times TF)$, is a Lie algebra homomorphism, iff the pullback $F^* : C^\infty(N)\to C^\infty(M)$, $F^*f = f\circ F$, is a Lie algebra homomorphsm. Here the Lie brackets are the Poisson brackets.* $\blacksquare$

2.4.5 Hamiltonian Vector Fields

Let (M,ω) be a symplectic manifold. We saw in the last section that to each smooth function $f\in C^\infty(M)$ there is an associated vector field $X_f\in\mathfrak{X}(M)$, called the **Hamiltonian vector field** associated to f, given by $X_f = \omega^\sharp(df)$, so that $i_{X_f}\omega = df$, and in local canonical/symplectic coordinates $X_f = \frac{\partial f}{\partial x^i}\frac{\partial}{\partial y^i} - \frac{\partial f}{\partial y^i}\frac{\partial}{\partial x^i}$ (Theorem 52 and the corollary to Darboux's Theorem, 9).

Definition 17 Fix a function $H\in C^\infty(M)$ and call it the **energy function**, and consider the associated vector field $X_H\in\mathfrak{X}(M)$. Then the triple (M,ω,X_H) is called a **Hamiltonian system**, and X_H is called the **Hamiltonian vector field** of the system. The terminology comes from classical mechanics, and in that setting symplectic coordinates are denoted with p's and q's (as in Section 2.3.3, especially equation (2.37)), with the q_i being the position coordinates and the p_j being the momentum coordinates. With this terminology and notation, locally we know by Darboux's theorem that ω has coordinate expression

$$\omega = \sum_{i=1}^n dq^i\wedge dp^i \qquad \text{and} \qquad X_H = \sum_{i=1}^n \frac{\partial H}{\partial q^i}\frac{\partial}{\partial p^i} - \frac{\partial H}{\partial p^i}\frac{\partial}{\partial q^i}$$

The last expression can be summarized in what are called **Hamilton's** or the **Hamiltonian equations** for the system (M,ω,X_H). Let $\gamma : I\to M$ be an integral curve of X_H which has local coordinate expression $\hat{\gamma}(t) = (q(t), p(t))$ in canonical coordinates. Then, Hamilton's equations

$$\begin{aligned} \frac{\partial H}{\partial p^i} &= \quad \dot{q}_i = \frac{dq_i}{dt} \\ \frac{\partial H}{\partial q^i} &= -\dot{p}_i = -\frac{dp_i}{dt} \end{aligned} \tag{2.120}$$

are just a restatement of the coordinate expression of X_H given above. ■

Theorem 20 (Conservation of Energy) *If (M, ω, H) is a Hamiltonian system, then the energy function H is constant on any integral curve of its associated Hamiltonian vector field X_H.*

Proof: If γ is an integral curve of X_H, then $\gamma'(t) = X_H(\gamma(t))$ for all t in the domain of γ, and so Hamilton's equations (2.120) hold in local canonical coordinates. But $dH(\gamma'(t)) := \gamma'(t)(H) := (H \circ \gamma)'(t)$, and

$$\begin{aligned} (H \circ \gamma)'(t) = dH(\gamma'(t)) = dH\big((X_H)_{\gamma'(t)}\big) \\ = i_{(X_H)_{\gamma'(t)}}\omega\big((X_H)_{\gamma'(t)}\big) = \omega\big((X_H)_{\gamma'(t)}, (X_H)_{\gamma'(t)}\big) = 0 \end{aligned}$$

so $H(\gamma(t))$ is constant. ■

Theorem 21 (Liouville) *If (M, ω, H) is a Hamiltonian system, then the flow F of the Hamiltonian vector field X_H forms a one parameter family of symplecticomorphisms of M, that is $F_t \in \mathrm{Sp}(M)$ for all t in the flow domain.*

Proof: Since ω is closed and $\omega_t = \omega$ for all t, so that $d\omega/dt = 0$, we have by Lemma 6 that

$$\frac{d}{dt}\big(F_t^*\omega\big) = F^*\big(i_{X_H}d\omega + d(i_{X_H}\omega)\big) = F^*\big(d(dH)\big)\big) = 0$$

so $F_t^*\omega$ is independent of t. But $F_0 = \mathrm{id}_M$, because it is the flow of X_H, so we must have $F_t^*\omega = F_0^*\omega = \omega$ for all t, which shows that $F_t \in \mathrm{Sp}(M)$. ■

Corollary 12 *The symplectomorphisms F_t of the flow of X_H preserve the volume form $\tau_\omega := \frac{(-1)^{[n/2]}}{n!}\omega^n$.* ■

Definition 18 Let

$$\mathrm{Ham}(M)$$

denote the vector subspace of $\mathfrak{X}(M)$ consisting of Hamiltonian vector fields on (M, ω). We call a vector field X **locally Hamiltonian** if for each $p \in M$ there is a neighborhood U of p such that $X|_U \in \mathrm{Ham}(U)$. The space of all locally Hamiltonian vector fields is denoted

$$\mathrm{Ham}^0(M)$$

We say that a vector field $X \in \mathfrak{X}(M)$ is a **symplectic vector field** if its flow F gives a 1-parameter family of symplectomorphisms, i.e. $F_t \in \mathrm{Sp}(M)$ for all t. We thus have, by Liouville's theorem that all Hamiltonian vector fields are symplectic. ■

With this terminology, we can state the following:

Theorem 22 *Let (M, ω) be a symplectic manifold. Then, for any vector field $X \in \mathfrak{X}(M)$ the following are equivalent:*

(1) $X \in \mathrm{Ham}^0(M)$.

(2) $i_X\omega$ *is closed.*

(3) X *is symplectic.*

(4) $\mathcal{L}_X\omega = 0$.

Proof: (1) $\implies$ (2) is simply $d(i_X\omega) = d(dH) = 0$. (1) $\implies$ (3) is Liouville's theorem, and (1) $\implies$ (4) is found in the proof of Liouville's theorem, since $\frac{d}{dt}\big|_{t=0}(F_t^*\omega) = 0$. (3) $\implies$ (4) is the observation that $F_t^*\omega = \omega$ implies $\mathcal{L}_X\omega = \frac{d}{dt}\big|_{t=0}(F_t^*\omega) = \frac{d}{dt}\big|_{t=0}\omega = 0$, while conversely $\mathcal{L}_X\omega = 0$ means $F_t^*\omega$ is a constant, which, since $F_0 = \mathrm{id}_M$, implies that $F_t^*\omega = F_0^*\omega = \omega$. (2) $\implies$ (1) Poincaré's lemma says that $i_X\omega$ being closed implies that it is locally exact, which means that around each point $p \in M$ there is a neighborhood U and a function $H \in C^\infty(U)$ such that $i_X\omega|_U = dH|_U$, i.e. that $X \in \mathrm{Ham}^0(M)$. (4) $\implies$ (2) Suppose $\mathcal{L}_X\omega = 0$. Then, by Cartan's formula $0 = d(i_X\omega) + i_X(d\omega) = d(i_X\omega)$ since ω is closed by assumption. ■

Corollary 13 $\mathrm{Ham}^0(M)$ *is a Lie subalgebra of* $\mathfrak{X}(M)$*.*

Proof: This follows from $\mathcal{L}_X\omega = 0$ and $\mathcal{L}_{[X,Y]} = [\mathcal{L}_X, \mathcal{L}_Y]$. ■

Remark 35 $\mathrm{Ham}(M) = \mathrm{Ham}^0(M)$ iff $H^1(M,\mathbb{R}) = 0$, since that is exactly the statement that closed 1-forms are globally exact, which is needed to have (2) $\implies$ (1) in the proof of Theorem 22. ■

Remark 36 Let $i : \mathbb{R} \to C^\infty(M)$ be the inclusion, $i(c) = c$, where the right-hand-side is the constant function $c : M \to \mathbb{R}$ with value $c \in \mathbb{R}$. Let $j : C^\infty(M) \to \mathrm{Ham}(M)$ be given by $j(f) := X_f$, i.e. j associates to each observable f its Hamiltonian vector field X_f. Then we have an exact sequence,

$$0 \longrightarrow \mathbb{R} \xrightarrow{i} C^\infty(M) \xrightarrow{j} \mathrm{Ham}(M) \longrightarrow 0 \tag{2.121}$$

called the **fundamental exact sequence**. This follows from the fact that

$$\begin{aligned} \ker j &= \{X_f \mid X_f = 0\} \\ &= \{X_f \mid df = i_{X_f}\omega = \omega(0,\cdot) = 0\} \\ &= \{X_f \mid f = i(c) \text{ for some } c \in \mathbb{R}\} \\ &= \operatorname{im} i \\ &= C^\infty(M) \end{aligned}$$

Consequently, we have

$$\mathrm{Ham}(M) \cong C^\infty(M)/\mathbb{R} \tag{2.122}$$

which is a restatement in terms of the First Isomorphism Theorem. ■

2.4.6 Symplectic Connections

This section treats the basic object in the study of symplectic geometry, the symplectic connection. Though it is the natural analog to the Levi-Civitta connection on a Riemannian manifold, the symplectic connection is not unique, and in fact the space of symplectic connections is an affine space. To have a canonical choice of symplectic connection on (M, ω) one needs additional structure, such as a (pseudo-)Kähler structure or a symmetric space structure on M. The former case is equivalent to the condition that ∇ preserves an ω-calibrated almost complex structure $J \in \Gamma(\mathcal{J}_c(TM, \omega))$, $\nabla J = 0$. Our treatement follows those of Bieliavsky et al. [17], Vaisman [103], Habermann [55], [54], and Gel'fand, Ratekh and Shubin [37].

2.4.6.1 Definition and Basic Properties

Definition 19 Let (M, ω) be a symplectic manifold. A **symplectic connection** $\nabla : \Gamma(TM) \to \Gamma(TM \otimes T^*M)$ on M is one which *respects the symplectic form* or *keep the symplectic form parallel*, meaning that its extension to $\Omega^\bullet(M)$ satisfies

$$\nabla \omega = 0 \tag{2.123}$$

Equivalently, for vector fields $X, Y, Z \in \Gamma(TM)$, this means

$$X\big(\omega(Y, Z)\big) = \omega(\nabla_X Y, Z) + \omega(Y, \nabla_X Z) \tag{2.124}$$

Associated to each symplectic connection are some auxiliary tensors: First, the **torsion tensor** $T \in \Gamma(T^*M^{\otimes 2} \otimes TM) = \Omega^2(M; TM)$, the $(2, 1)$ tensor given by

$$T^\nabla(X, Y) := \nabla_X Y - \nabla_Y X - [X, Y]$$

If $(\mathsf{e}_1, \ldots, \mathsf{e}_n, \mathsf{f}_1, \ldots, \mathsf{f}_n) \in \Gamma(U, P_{\mathrm{Sp}(n,\mathbb{R})})$ is a local symplectic frame over $U \subseteq M$, then, following Habermann and Habermann [54], we can also define the associated **torsion vector field** $\mathfrak{T}$ of ∇,

$$\mathfrak{T} = \sum_{j=1}^{n} T^\nabla(\mathsf{e}_j, \mathsf{f}_j)$$ ■

Remark 37 Some authors, e.g. [17], also require a symplectic connection to be torsion-free, in which case the connection with torsion is called an *almost symplectic connection*. Following Habermann and Habermann [54] we prefer to use the term symplectic for connections only satisfying $\nabla \omega = 0$. ■

Remark 38 Recall Remark 31 on the trivialization of the vertical bundle of the symplectic frame bundle, $P_{\mathrm{Sp}(n,\mathbb{R})} \times \mathfrak{sp}(n,\mathbb{R}) \cong VP_{\mathrm{Sp}(n,\mathbb{R})}$. A principal connection in $P_{\mathrm{Sp}(n,\mathbb{R})}$ is an $\mathfrak{sp}(n,\mathbb{R})$-valued 1-form $\alpha \in \Omega^1(P_{\mathrm{Sp}(n,\mathbb{R})};\mathfrak{sp}(n,\mathbb{R}))$ which satisfies $\alpha(\tilde{X}) = X$ and $S^*\alpha = \mathrm{Ad}_S\,\alpha$ for all $X \in \mathfrak{sp}(n,\mathbb{R})$ and $S \in \mathrm{Sp}(n,\mathbb{R})$, and has $\ker\alpha = HP_{\mathrm{Sp}(n,\mathbb{R})}$, the complementary horizontal bundle to the vertical bundle $VP_{\mathrm{Sp}(n,\mathbb{R})}$, that is

$$P_{\mathrm{Sp}(n,\mathbb{R})} = HP_{\mathrm{Sp}(n,\mathbb{R})} \oplus VP_{\mathrm{Sp}(n,\mathbb{R})} = \ker\alpha \oplus \ker T\pi$$

where $\pi : P_{\mathrm{Sp}(n,\mathbb{R})} \to M$. Now, α determines and is determined by a symplectic connection ∇ as follows: Over a trivializing open subset U of M for $P_{\mathrm{Sp}(n,\mathbb{R})}$, take a section $s : U \to P_{\mathrm{Sp}(n,\mathbb{R})}$ (a local frame) and note that for any fixed $\mathbf{u} \in \mathbb{R}^n$ the vector $s\mathbf{u}$ is a vector field on U, $s\mathbf{u} \in \Gamma(U,TM)$, and $s^*\alpha \in \Omega^1(U;\mathfrak{sp}(n,\mathbb{R}))$, and these are related by

$$\nabla_X(s\mathbf{u}) = s\big((s^*\alpha)(X)\mathbf{u}\big), \quad \text{i.e.} \quad s^{-1}\nabla s = s^*\alpha$$

Moreover, the components of α and the soldering form θ are linearly independent and sapne the cotangent spaces of $P_{\mathrm{Sp}(n,\mathbb{R})}$ at each point, making $P_{\mathrm{Sp}(n,\mathbb{R})}$ parallelizable. This makes $P_{\mathrm{Sp}(n,\mathbb{R})}$ simpler to work with than M for many purposes. See Bieliavsky et al. [17] for applications of this observation. ■

To see the equivalence of the two conditions (2.123) and (2.124) in the definition above we need to unravel the definition of $\nabla\omega$.

Definition 20 As a map $\nabla : \Omega^\bullet(M) \to \Omega^\bullet(M)$, the extension of a connection $\nabla : \Gamma(TM) \to \Gamma(T^*M \otimes TM)$ is defined as follows:

(1) On functions, or 0-forms $f \in \Omega^0(M) = C^\infty(M)$, we define $\nabla : \Omega^0(M) \to \Omega^1(M)$ by

$$(\nabla f)(X) := \nabla_X f := df(X) = Xf$$

(2) On 1-forms $\alpha \in \Omega^1(M)$, we define $\nabla : \Omega^1(M) \to \Omega^2(M)$ by

$$\begin{aligned}(\nabla\alpha)(X,Y) &:= \nabla_X(\alpha(Y)) - \alpha(\nabla_X Y) \\ &= d(\alpha(Y))(X) - \alpha(\nabla_X Y) \\ &= X(\alpha(Y)) - \alpha(\nabla_X Y)\end{aligned}$$

(3) Then we extend ∇ to all other forms by requiring ∇ to be a derivation with respect to the wedge product:

$$\nabla(\alpha \wedge \beta) := (\nabla\alpha) \wedge \beta + \alpha \wedge (\nabla\beta)$$

Then we extend ∇ linearly over all $\Omega^\bullet(M)$. ■

For example, on simple 2-forms $\alpha \wedge \beta \in \Omega^2(M)$, where $\alpha, \beta \in \Omega^1(M)$, we have

$$\begin{aligned}
\nabla(\alpha \wedge \omega)(X, Y, Z) &= \big((\nabla\alpha) \wedge \beta + \alpha \wedge (\nabla\beta)\big)(X, Y, Z) \\
&= \big((\nabla_X(\alpha(Y)) - \alpha(\nabla_X Y)) \wedge \beta(Z) \\
&\qquad +\alpha(Y) \wedge ((\nabla_X(\beta(Z)) - \beta(\nabla_X Z))\big) \\
&= X(\alpha(Y)) \cdot \beta(Z) - \alpha(\nabla_X Y) \cdot \beta(Z) \\
&\qquad +\alpha(Y) \cdot (X\beta(Z)) - \alpha(Y) \cdot \beta(\nabla_X Z)
\end{aligned}$$

Remark 39 Let us explain the second of the conditions, the behavior of ∇ on 1-forms, for then we will see how to calculate the covariant derivative of a 2-form like ω as a function on vector fields. First, recall that for any connection ∇ on M the principled way to extend it to tensor fields $\Gamma(TM^{\otimes n})$ is by requiring it to behave as a derivation with respect to the tensor product,

$$\nabla_X(Y_1 \otimes Y_2 \otimes \cdots \otimes Y_n) := \sum_{j=1}^{n} Y_1 \otimes \cdots \otimes (\nabla_X Y_j) \otimes \cdots \otimes Y_n$$

To extend it to forms, or covectors, we additionally require it to commute with contraction. So let us explain contraction: **Contraction** or **trace** is a pairing function,

$$\begin{aligned}
&\mathrm{tr} : \Gamma(T^*M) \otimes \Gamma(TM) \to C^\infty(M) \\
&\mathrm{tr}(\alpha \otimes X) := \alpha(X)
\end{aligned}$$

This coordinate-free definition of contraction/trace is just the usual trace on square matrices, via the isomorphism $\operatorname{End} V \cong V^* \otimes V$, $\alpha \otimes v \mapsto \alpha(\cdot)v$. Given any $\alpha \in \Omega^1(M) = \Gamma(T^*M)$ and $X \in \Gamma(TM)$ we thus require of ∇ that

$$d(\alpha(Y)) = \nabla(\mathrm{tr}(\alpha \otimes Y)) = \mathrm{tr}(\nabla(\alpha \otimes Y) = \mathrm{tr}((\nabla\alpha) \otimes Y + \alpha \otimes (\nabla Y)) = (\nabla\alpha)(Y) + \alpha(\nabla Y)$$

from which we get an expression of $\nabla\alpha$ as a 2-form:

$$(\nabla\alpha)(X,Y) = d(\alpha(Y))(X) - \alpha(\nabla_X Y) = X(\alpha(Y)) - \alpha(\nabla_X Y)$$

which is precisely condition (2) in the definition of ∇ on 1-forms above.

Proof of the equivalence of (2.123) and (2.124): Consider $\omega \in \Omega^2(M)$, and let $X, Y \in \Gamma(TM)$. Then $\omega(X,\cdot),\ \omega(\cdot,Y) \in \Omega^1(M)$, and we write

$$\mathrm{tr}_1(\omega\otimes X) := \omega(X,\cdot)$$

$$\mathrm{tr}_2(\omega\otimes Y) := \omega(\cdot,Y)$$

Then,

$$\omega(Y,Z) = \mathrm{tr}_1\circ\mathrm{tr}_2(\omega\otimes Z\otimes Y) = \mathrm{tr}_2\circ\mathrm{tr}_1(\omega\otimes Y\otimes Z)$$

Therefore

$$\begin{aligned} d\big(\omega(Y,Z)\big) &= \nabla(\mathrm{tr}_1\circ\mathrm{tr}_2(\omega\otimes Z\otimes Y) \\ &= \mathrm{tr}_1\circ\mathrm{tr}_2\big(\nabla(\omega\otimes Z\otimes Y)\big) \\ &= \mathrm{tr}_1\circ\mathrm{tr}_2\big((\nabla\omega)\otimes Z\otimes Y + \omega\otimes(\nabla Z)\otimes Y + \omega\otimes Z\otimes(\nabla Y)\big) \\ &= (\nabla\omega)(Y,Z) + \omega(Y,\nabla Z) + \omega(\nabla Y, Z) \end{aligned}$$

That is, as a 3-form, $\nabla\omega$ is given by

$$\begin{aligned} (\nabla\omega)(X,Y,Z) &= (\nabla_X\omega)(Y,Z) \\ &= X\big(\omega(Y,Z)\big) - \omega(\nabla_X Y, Z) - \omega(Y,\nabla_X Z) \end{aligned}$$

The condition (2.123) on a symplectic for ω, that $\nabla\omega = 0$, therefore means

$$X\big(\omega(Y,Z)\big) = \omega(\nabla_X Y, Z) + \omega(Y,\nabla_X Z)$$

which completes the proof. ∎

Notation 3 *Let us introduce the following notation for the different spaces of connections we are likely to encounter. Let*

$$\mathcal{C}(M) = \{\textit{connections on } M\} \tag{2.125}$$

$$\mathcal{C}_0(M) = \{\textit{torsion-free connections on } M\} \tag{2.126}$$

denote the space of all connections on M*, an affine space modelled on the vector space* $\Omega^1(M; \operatorname{End} TM)$*, i.e.* $\mathcal{C}(M) = \nabla + \Omega^1(M; \operatorname{End} TM)$*. Let*

$$\mathcal{C}(M, g) = \{\nabla \in \mathcal{C}(M) \mid \nabla g = 0\} \tag{2.127}$$

$$\mathcal{C}_0(M, g) = \mathcal{C}(M, g) \cap \mathcal{C}_0(M) = \{\textit{Levi-Civita connection}\} \tag{2.128}$$

denote, respectively, the space of metric connections and torsion-free metric connections, that is those $\nabla \in \mathcal{C}(M)$ *preserving the metric* g *(equivalently, satisfying* $Xg(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$ *for all* $X, Y, Z \in \Gamma(TM)$*), the former an affine space modelled on the vector space* $\mathcal{E}^1(M, g) := \{A \in \operatorname{End} TM \mid g(A(X)Y, Z) = -g(A(X)Z, Y)\}$*,*

$$\mathcal{C}(M, g) = \nabla + \mathcal{E}^1(M, g)$$

The affine map

$$\Psi : \mathcal{C}(M) \to \Omega^2(M; TM), \qquad \Psi(\nabla) := T^\nabla$$

sending a connection to its torsion tensor has an underlying linear map

$$\Phi : \Omega^1(M; \operatorname{End} TM) \to \omega^2(M; TM), \qquad \Phi(A)(X, Y) := A(X)Y - A(Y)X$$

It is well-known (cf Lemma 2.1 and Corollary 2.2 of Habermann [55]) that the restriction of the linear map Φ *to* $\mathcal{E}^1(M, g)$ *is an isomorphism onto* $\Omega^2(M; TM)$*,*

$$\Phi \in \mathrm{GL}\big(\mathcal{E}^1(M, g), \omega^2(M; TM)\big)$$

and the associated affine map Ψ *is injective. In particular, the preimage of* 0 *under* Ψ *is the Levi-Civita connection,*

$$\Psi^{-1}(0) = \nabla^{LC} \quad (\textit{Levi-Civita connection})$$

Next, we introduce the following subspaces of $\mathcal{C}(M)$ needed in the symplectic category: Let (M,ω) be a symplectic space, and let

$$\mathcal{C}(M,\omega) = \{\nabla \in \mathcal{C}(M) \mid \nabla\omega = 0\} \tag{2.129}$$

denote the space of symplectic connections (we will see below that it is a nonempty affine space modelled on a certain vector space). Let

$$\mathcal{C}_0(M,\omega) := \{\nabla \in \mathcal{C}(M) \mid \nabla\omega = T^{\nabla} = 0\} \tag{2.130}$$

denote the space of torsion-free symplectic connections. If $J \in \Gamma(\mathcal{J}(TM,\omega))$ is an ω-compatible almost complex structure, let

$$\mathcal{C}(M,\omega,J) = \{\nabla \in \mathcal{C}(M,\omega) \mid \nabla J = 0\} \tag{2.131}$$

denote the space of symplectic connections preserving the almost complex structure J. The torsion-free analog of $\mathcal{C}(M,\omega,J)$ is denoted

$$\mathcal{C}_0(M,\omega,J) = \{\nabla \in \mathcal{C}_0(M,\omega) \mid \nabla J = 0\} \tag{2.132}$$

$\blacksquare$

Proposition 56 *Let (M,ω) be a symplectic manifold. The space $\mathcal{C}(M,\omega)$ of symplectic connections is a nonempty affine subspace of $\mathcal{C}(M)$ modelled on the vector space of $(1,2)$-tensors $\mathcal{E}^1(M,\omega) = \{S \in \Gamma(T^*M^{\otimes 2} \otimes TM) \mid \omega(S(X,Y),Z) = \omega(S(X,Z),Y)\}$,*

$$\mathcal{C}(M,\omega) = \nabla + \mathcal{E}^1(M,\omega) \tag{2.133}$$

for any given symplectic connection $\nabla \in \mathcal{C}(M,\omega)$. The space $\mathcal{C}_0(M,\omega)$ of torsion-free symplectic connections is a nonempty affine space modelled on the vector space of symmetric $(0,3)$-tensor $\Gamma(T^*M^{\odot 3})$,

$$\mathcal{C}_0(M,\omega) = \nabla + \Gamma(T^*M^{\odot 3}) \tag{2.134}$$

Proof: (1) Let us show existence first. Choose any connection $\nabla^0 \in \mathcal{C}(M)$, which we may suppose to be torsion-free so that we cover both cases, for instance the Levi-Civita connection, and use the nondegeneracy of ω to define the $(1,2)$-tensor N on M by

$$(\nabla^0_X\omega)(Y,Z) = \omega(N(X,Y),Z)$$

In other words $N(X,Y) = \omega^\sharp\big(\nabla^0_X\omega)(Y,\cdot)\big)$. Now, the skew-symmetry of ω implies

$$\begin{aligned}\omega(N(X,Z),Y) &= (\nabla^0_X)(Z,Y) = X\big(\omega(Z,Y)\big) - \omega(\nabla^0_X Z,Y) - \omega(Z,\nabla^0_X Y)\\ &= -X\big(\omega(Y,Z)\big) + \omega(\nabla^0_X Y,Z) + \omega(Y,\nabla^0_X Z) = -(\nabla^0_X\omega)(Y,Z) = -\omega(N(X,Y),Z)\end{aligned}$$

while the closedness of ω implies that

$$\begin{aligned}
&\omega(N(X,Y),Z) + \omega(N(Y,Z),X) + \omega(N(Z,X),Y)\\
&\quad= (\nabla^0_X\omega)(Y,Z) + (\nabla^0_Y\omega)(Z,X) + (\nabla^0_Z\omega)(X,Y)\\
&\quad= X(\omega(Y,Z)) - \omega(\nabla^0_X Y,Z) - \omega(Y,\nabla^0_X Z)\\
&\qquad + Y(\omega(Z,X)) - \omega(\nabla^0_Y Z,X) - \omega(Z,\nabla^0_Y X)\\
&\qquad\quad + Z(\omega(X,Y)) - \omega(\nabla^0_Z X,Y) - \omega(X,\nabla^0_Z Y)\\
&\quad= X(\omega(Y,Z)) - \omega(\nabla^0_X Y - \nabla^0_Y X,Z)\\
&\qquad + Y(\omega(Z,X)) - \omega(\nabla^0_Y Z - \nabla^0_Z Y,X)\\
&\qquad\quad + Z(\omega(X,Y)) - \omega(\nabla^0_Z X - \nabla^0_X Z,Y)\\
&\quad= X(\omega(Y,Z)) - \omega([X,Y],Z)\\
&\qquad - Y(\omega(X,Z)) - \omega([Z,Y],X)\\
&\qquad\quad + Z(\omega(X,Y)) + \omega([X,Z],Y)\\
&\quad= d\omega(X,Y,Z)\\
&\quad= 0
\end{aligned}$$

Now, define $\nabla \in \mathcal{C}(M)$ by

$$\nabla_X Y := \nabla^0_X Y + \frac{1}{3}N(X,Y) + \frac{1}{3}N(Y,X) \tag{2.135}$$

and notice that $T^{\nabla} = T^{\nabla^0} = 0$ since in the expression for T^{∇} the N terms cancel and ∇^0 is torsion-free. Moreover, ∇ is symplectic:

$$\begin{aligned}
(\nabla_X\omega)(Y,Z) &= X(\omega(Y,Z)) - \omega(\nabla_X Y, Z) - \omega(Y, \nabla_X Z) \\
&= \Big[(\nabla^0_X\omega)(Y,Z) + \omega(\nabla^0_X Y, Z) + \omega(Y, \nabla^0_X Z)\Big] \\
&\qquad -\Big[\omega(\nabla^0_X Y, Z) + \frac{1}{3}\omega(N(X,Y),Z) + \frac{1}{3}\omega(N(Y,X),Z)\Big] \\
&\qquad\quad -\Big[\omega(Y,\nabla^0_X Z) + \frac{1}{3}\omega(Y,N(X,Z)) + \frac{1}{3}\omega(Y,N(Z,X)\Big] \\
&= (\nabla^0_X\omega)(Y,Z) - \frac{1}{3}\omega(N(X,Y),Z) - \frac{1}{3}\omega(N(Y,X),Z) \\
&\qquad -\frac{1}{3}\omega(Y,N(X,Z)) - \frac{1}{3}\omega(Y,N(Z,X)) \\
&= 0
\end{aligned}$$

the last equality following from

(a) $(\nabla^0_X\omega)(Y,Z) = \omega(N(X,Y),Z)$

(b) $-\omega(Y,N(X,Z)) = \omega(N(X,Z),Y) = -\omega(N(X,Y),Z)$

(c) $-\omega(N(Y,X),Z) - \omega(Y,N(Z,X)) = \omega(N(Y,Z),X)) + \omega(N(Z,X),Y)$

$= -\omega(N(X,Y),Z)$

(2) $\nabla' = \nabla + S$ is symplectic iff $\nabla' - \nabla = S$ is symplectic, which means that for all $X, Y, Z \in \Gamma(TM)$ we have

$$\begin{aligned}
0 &= X\omega(Y,Z) - X\omega(Y,Z) \\
&= \Big(\omega(\nabla'_X Y, Z) + \omega(Y, \nabla'_X Z)\Big) - \Big(\omega(\nabla_X Y, Z) + \omega(Y, \nabla_X Z)\Big) \\
&= \omega\big(\nabla'_X Y - \nabla_X Y,\ Z\big) + \omega\big(Y,\ \nabla'_X Z - \nabla_X Z\big) \\
&= \omega(S(X,Y),Z) + \omega(Y,S(X,Z)) \\
&= \omega(S(X,Y),Z) - \omega(S(X,Z),Y)
\end{aligned}$$

(3) If additionally $T^{\nabla} = 0$, then the requirement that $T^{\nabla'} = 0$ means

$$0 = T^{\nabla'}(X,Y) = \nabla'_X Y - \nabla'_Y X - [X,Y]$$
$$= \nabla_X Y - \nabla_Y X + S(X,Y) - S(Y,X) - [X,Y] = S(X,Y) - S(Y,X)$$

or $S(X,Y) = S(Y,X)$, so that $S \in \Gamma(T^*M^{\odot 2} \otimes TM)$. Since also $\omega(S(X,Y),Z) = \omega(S(X,Z),Y)$, we have that $\omega(S(X,Y),Z)$ is totally symmetric. Thus, we can view $\omega(S(X,Y),Z)$, and hence S, as an element of $T^*M^{\odot 3}$. ■

Remark 40 For any symplectic connection $\nabla \in \mathcal{C}(M,\omega)$, the fact that $d\omega = 0$ implies

$$\begin{aligned}
0 &= d\omega(X,Y,Z) \\
&= X(\omega(Y,Z)) - Y(\omega(X,Z)) + Z(\omega(X,Y)) \\
&\qquad -\omega([X,Y],Z) + \omega([X,Z],Y) - \omega([Y,Z],X) \\
&= \omega(\nabla_X Y, Z) + \omega(Y, \nabla_X Z) - \omega(\nabla_Y X, Z) - \omega(X, \nabla_Y Z) + \omega(\nabla_Z X, Y) + \omega(X, \nabla_Z Y) \\
&\qquad -\omega([X,Y],Z) + \omega([X,Z],Y) - \omega([Y,Z],X) \\
&= \omega(\nabla_X Y - \nabla_Y X - [X,Y], Z) + \omega(\nabla_Y Z - \nabla_Z Y - [Y,Z], X) \\
&\qquad +\omega(\nabla_Z X - \nabla_X Z - [Z,X], Y) \\
&= \omega(T^{\nabla}(X,Y),Z) + \omega(T^{\nabla}(Y,Z),X) + \omega(T^{\nabla}(Z,X),Y)
\end{aligned}$$

That is, the sum over a cyclic permutation is zero:

$$\sum_{\sigma=(1\ 2\ 3)\in S_3} \omega(T^{\nabla}(X_{\sigma 1}, X_{\sigma 2}), X_{\sigma 3}) = 0$$

If we are given a torsion-free connection $\nabla^0 \in \mathcal{C}(M)$, we can alternatively define the $(1,2)$-tensor N by

$$\omega(N(X,Y),Z) = \frac{1}{3}\Big[(\nabla^0_X \omega)(Y,Z) + (\nabla^0_Y \omega)(X,Z)\Big]$$

then the connection $\nabla + N \in \mathcal{C}(M)$ is also torsion free (since the Ns cancel in the expressin for T^{∇}) and satisfies

$$X(\omega(Y,Z)) - \omega(\nabla_X Y, Z) - \omega(Y, \nabla_X Z) = \frac{1}{3} d\omega(X,Y,Z)$$

which is symplectic if ω is a symplectic form, since the right-hand side vanishes. This is an alternative construction of a symplectic connection, [55, Proposition 2.5]. ■

Corollary 14 *Let (M,ω) be a symplectic manifold. If we define the maps*

$$\Phi : \mathcal{C}(M,\omega) \to \Omega^2(M;TM), \quad \Phi(\nabla) := T^\nabla$$

and

$$\Psi : \Gamma(TM \otimes T^*M^{\otimes 2}) \to \Omega^2(M;TM), \quad \Psi(N)(X,Y) := N(X,Y) - N(Y,X)$$

then

(1) $\Psi(N) = 0$ *iff* $\omega(N(\cdot,\cdot),\cdot) \in \Gamma(T^*M^{\odot 3})$

(2) $\Psi(\mathcal{E}^1(M,\omega)) = \{A \in \Omega^2(M;TM) \mid \omega(A(X,Y),Z) + \omega(A(Y,Z),X) + \omega(A(Z,X),Y) = 0,\ \forall X,Y,Z \in \Gamma(TM)\}$

(3) Φ *is neither injective nor bijective, and the preimage $\Phi^{-1}(A)$ of any $A \in \Omega^2(M;TM)$ is either empty or infinite-dimensional.*

■

Proposition 57 *Let (M,ω) be a symplectic manifold, let $J \in \mathcal{J}(M,\omega)$, and let $g_J := \omega(\cdot, J\cdot)$ be the induced Riemannian metric. If $\nabla^0 \in \mathcal{C}(M,g_J) \cup \mathcal{C}(M,\omega)$ is either metric or symplectic, then the connection $\nabla \in \mathcal{C}(M)$ defined by*

$$\nabla_X Y := \nabla^0_X Y + \frac{1}{2}(\nabla^0_X J)(JY)$$

satisfies

$$\nabla\omega = \nabla g_J = \nabla J = 0$$

i.e. $\nabla \in \mathcal{C}(M,g_J) \cap \mathcal{C}(M,\omega)$ and $\nabla J = 0$. In this case, moreover, J commutes with ∇,

$$J(\nabla_X Y) = \nabla_X(JY)$$

Proof: (1) Let us write g for g_J in what follows, to avoid cumbersome notation. Recall that if $J \in \mathrm{End}(TM)$, then

$$(\nabla^0_X J)(Y) = \nabla^0_X(JY) - J(\nabla^0_X Y)$$

which follows from the requirement that ∇^0 commute with trace:

$$\begin{aligned}\nabla^0_X(J(Y)) = \nabla^0_X(\mathrm{tr}(J \otimes Y)) = \mathrm{tr}(\nabla^0_X(J \otimes Y))& \\ = \mathrm{tr}((\nabla^0_X J) \otimes Y + J \otimes (\nabla^0_X Y)) = (\nabla^0_X J)(Y) + J(\nabla^0_X Y)&\end{aligned}$$

Also, since $J \in \mathcal{J}(M,\omega)$ we must have $J^*g = g$, $J^*\omega = \omega$, and J skew-adjoint with respect to g (Proposition 28), so

$$\begin{aligned}\omega(\nabla_X Y, Z) &= \omega(\nabla^0_X Y, Z) + \frac{1}{2}\omega((\nabla^0_X J)(JY), Z) \\ &= \omega(\nabla^0_X Y, Z) + \frac{1}{2}\omega(\nabla^0_X(J^2 Y), Z) - \frac{1}{2}\omega(J(\nabla^0_X(JY)), Z) \\ &= \omega(\nabla^0_X Y, Z) - \frac{1}{2}\omega(\nabla^0_X Y, Z) + \frac{1}{2}g(Z, \nabla^0_X(JY)) \\ &= \frac{1}{2}\omega(\nabla^0_X Y, Z) + \frac{1}{2}g(Z, \nabla^0_X(JY)) \\ &= -\frac{1}{2}g(\nabla^0_X Y, JZ) + \frac{1}{2}g(Z, \nabla^0_X(JY))\end{aligned}$$

Thus also

$$\begin{aligned}\omega(Y, \nabla_X Z) &= -\omega(\nabla_X Z, Y) \\ &= \frac{1}{2}g(\nabla^0_X Z, JY) - \frac{1}{2}g(Y, \nabla^0_X(JZ)) \\ &= \frac{1}{2}g(JY, \nabla^0_X Z) - \frac{1}{2}g(\nabla^0_X(JZ), Y)\end{aligned}$$

(2) Suppose now that $\nabla^0 g = 0$. Then, by (1)

$$\begin{aligned}
\omega(\nabla_X Y, Z) + \omega(Y, \nabla_X Z) &= -\frac{1}{2}g(\nabla^0_X Y, JZ) + \frac{1}{2}g(Z, \nabla^0_X(JY)) \\
&\qquad +\frac{1}{2}g(JY, \nabla^0_X Z) - \frac{1}{2}g(\nabla^0_X(JZ), Y) \\
&= \frac{1}{2}X(g(JY, Z)) - \frac{1}{2}X(g(Y, JZ)) \\
&= -\frac{1}{2}X(g(Y, JZ)) - \frac{1}{2}X(g(Y, JZ)) \\
&= -X(g(Y, JZ)) \\
&= X(g(JY, Z)) \\
&= X(\omega(Y, Z))
\end{aligned}$$

which shows that $\nabla\omega = 0$. To see that $\nabla g = \nabla J = 0$ as well, note that

$$\begin{aligned}
\nabla_X(JY) &= \nabla^0_X(JY) + \frac{1}{2}(\nabla^0_X J)(J^2 Y) \\
&= \nabla^0_X(JY) - \frac{1}{2}\nabla^0_X(JY) + \frac{1}{2}J(\nabla^0_X Y) \\
&= \frac{1}{2}\nabla^0_X(JY) + \frac{1}{2}J(\nabla^0_X Y)
\end{aligned}$$

and

$$\begin{aligned}
J(\nabla_X Y) &= J(\nabla^0_X Y) + \frac{1}{2}J\big((\nabla^0_X J)(JY)\big) \\
&= J(\nabla^0_X Y) + \frac{1}{2}J\big(\nabla^0_X(J^2 Y) - J(\nabla^0_X(JY))\big) \\
&= J(\nabla^0_X Y) - \frac{1}{2}J(\nabla^0_X Y) + \frac{1}{2}\nabla^0_X(JY) \\
&= \frac{1}{2}J(\nabla^0_X Y) + \frac{1}{2}\nabla^0_X(JY)
\end{aligned}$$

Therefore,

$$\begin{aligned}
(\nabla_X J)(Y) &= \nabla_X(JY) - J(\nabla_X Y) \\
&= \Big[\frac{1}{2}\nabla^0_X(JY) + \frac{1}{2}J(\nabla^0_X Y)\Big] - \Big[\frac{1}{2}J(\nabla^0_X Y) + \frac{1}{2}\nabla^0_X(JY)\Big] \\
&= 0
\end{aligned}$$

Moreover, since J is ω-compatible and $\nabla J = 0$, we have $J(\nabla_X Y) = \nabla_X(JY) - (\nabla_X J)(Y) = \nabla_X(JY)$, showing that J commutes with ∇. As a result,

$$\begin{aligned} g(\nabla_X Y, Z) + g(Y, \nabla_X, Z) &= \omega(\nabla_X Y, JZ) + \omega(Y, J(\nabla_X Z)) \\ &= \omega(\nabla_X Y, JZ) + \omega(Y, \nabla_X(JZ)) \\ &= X(\omega(Y, JZ)) \\ &= X(g(Y, Z)) \end{aligned}$$

i.e. $\nabla g = 0$.

(3) If $\nabla^0 \omega = 0$, then by an analogous procedure we conclude that $\nabla g = \nabla \omega = \nabla J = 0$. ■

Corollary 15 *Let $\nabla \in \mathcal{C}_0(M, \omega)$ be a torsion-free symplectic connection and let $J \in \mathcal{J}(M, \omega)$. Then $\nabla J = 0$ iff ∇ is the unique Levi-Civita connection on (M, g_J). If $J \in \mathcal{J}(TM)$ only, then $\nabla J = 0$ iff ∇ is the unique Levi-Civita connection associated to the pseudo-Riemannian metric g_J.* ■

2.4.6.2 Divergence

Definition 21 If $\nabla \in \mathcal{C}(M)$ is any connection, we can define the **trace** of the endomorphism $\nabla Y \in \operatorname{End} TM$, $(\nabla Y)(X) := \nabla_X Y$. By means of this trace we can define the **divergence** of a vector field $X \in \Gamma(TM)$,

$$\operatorname{div}(X) := \operatorname{tr}(\nabla X)$$

In local coordinates ∂_j, the divergence takes the form $\operatorname{div}(X) = \sum_{j=1}^n dx^j(\nabla_{\partial_j} X)$. ■

Lemma 7 *If $X \in \Gamma(TM)$, then locally, with respect to a local symplectic frame* $(\mathsf{e}_1, \ldots, \mathsf{e}_n, \mathsf{f}_1, \ldots, \mathsf{f}_n) \in \Gamma(U, P_{\mathrm{Sp}(n,\mathbb{R})})$*, the divergence may be expressed as*

$$\operatorname{div}(X) = \sum_{j=1}^n \big(\omega(\nabla_{\mathsf{e}_j} X, \mathsf{f}_j) - \omega(\nabla_{\mathsf{f}_j} X, \mathsf{e}_j)\big)$$

Proof: This follows from the relations (2.35), which here appear as

$$\omega(\mathsf{e}_i,\mathsf{e}_j)=\omega(\mathsf{f}_i,\mathsf{f}_j)=0 \qquad \text{and} \qquad \omega(\mathsf{e}_i,\mathsf{f}_j)=\delta_{ij}$$

If we denote by e_j^* and f_j^* the duals basis vectors, then

$$\mathsf{e}_j^*(X)=\omega(X,\mathsf{f}_j) \qquad \text{and} \qquad \mathsf{f}_j^*(X)=\omega(\mathsf{e}_j,X)=-\omega(X,\mathsf{e}_j)$$

so $\mathrm{div}(X)=\mathrm{tr}(\nabla X)=\sum_{j=1}^n \mathsf{e}^*(\nabla_{\mathsf{e}_j}X)+\mathsf{f}^*(\nabla_{\mathsf{f}_j}X)=\sum_{j=1}^n\big(\omega(\nabla_{\mathsf{e}_j}X,\mathsf{f}_j)-\omega(\nabla_{\mathsf{f}_j}X,\mathsf{e}_j)\big)$. ■

Notation 4 *Let us write* $X^\flat:=\omega^\flat(X)=\omega(X,\cdot)\in\Omega^1(M)$ *for any* $X\in\Gamma(TM)$, *and denote the volume form on* M *induced by* ω *by*

$$dV:=\frac{1}{n!}\omega^n$$ ■

Proposition 58 *For any symplectic connection* $\nabla\in\mathcal{C}(M,\omega)$ *and any vector field* $X\in\Gamma(TM)$, *we have*

$$d(X^\flat\wedge\omega^{n-1})=2\cdot(n-1)!\Big(\mathrm{div}(X)+\omega(X,\mathfrak{T})\Big)dV$$

Proof: Let $\beta=(\mathsf{e}_1,\mathsf{f}_1,\dots,\mathsf{e}_n,\mathsf{f}_n)\in\Gamma(U,P_{\mathrm{Sp}(n,\mathbb{R})})$ be a local symplectic frame and note that $dV(\beta)=1$. Since ω is closed, $d(X^\flat\wedge\omega^{n-1})=dX^\flat\wedge\omega^{n-1}$, and since there are $(n-1)!$ permutations of $(n-1)$ of the n pairs $(\mathsf{e}_j,\mathsf{f}_j)$ and 2 (signed) permuations of each remaining pair, we have

$$d(X^\flat\wedge\omega^{n-1})(\beta)=\sum_{\sigma\in S_n}\mathrm{sgn}(\sigma)dX^\flat\wedge\omega^{n-1}(\beta)=2\cdot(n-1)!\sum_{j=1}^n dX^\flat(\mathsf{e}_j,\mathsf{f}_j)$$

Now, for any $Y,Z\in\Gamma(TM)$,

$$\begin{aligned}
dX^\flat(Y,Z) &= Y(X^\flat(Z))-Z(X^\flat(Y))-X^\flat([Y,Z])\\
&= Y(\omega(X,Z))-Z(\omega(X,Y))-\omega(X,[Y,Z])\\
&= \omega(\nabla_YX,Z)+\omega(X,\nabla_YZ)-\omega(\nabla_ZX,Y)-\omega(X,\nabla_ZY)-\omega(X,[Y,Z])\\
&= \omega(\nabla_YX,Z)-\omega(\nabla_ZX,Y)+\omega(X,\nabla_YZ-\nabla_ZY-[Y,Z])\\
&= \omega(\nabla_YX,Z)-\omega(\nabla_ZX,Y)+\omega(X,T^\nabla(Y,Z))
\end{aligned}$$

so

$$\begin{aligned}\sum_{j=1}^{n} dX^\flat(\mathsf{e}_j,\mathsf{f}_j) &= \sum_{j=1}^{n}\Big(\omega(\nabla_{\mathsf{e}_j}X,\mathsf{f}_j)-\omega(\nabla_{\mathsf{f}_j}X,\mathsf{e}_j)\Big)+\omega\Big(X,\sum_{j=1}^{n}T^\nabla(\mathsf{e}_j,\mathsf{f}_j)\Big)\\ &= \operatorname{div}(X)+\omega(X,\mathfrak{T})\end{aligned}$$

■

Corollary 16 *For any symplectic connection $\nabla\in\mathcal{C}(M,\omega)$ and any compactly supported vector field $X\in\Gamma_c(TM)$ we have*

$$\int_M\Big(\operatorname{div}(X)+\omega(X,\mathfrak{T})\Big)dV=0$$

Proof: By Stokes' Theorem, the previous proposition, and the compact support of X we have

$$\int_M\Big(\operatorname{div}(X)+\omega(X,\mathfrak{T})\Big)dV=\frac{1}{2(n-1)!}\int_M d(X^\flat\wedge\omega^{n-1})=\frac{1}{2(n-1)!}\int_{\partial M}X^\flat\wedge\omega^{n-1}=0$$ ■

2.4.6.3 Curvature Tensor of a Symplectic Connection

If (M,ω) is a symplectic manifold of dimension $2n$ and $\nabla\in\mathcal{C}(M,\omega)$ is a symplectic connection on M, then we have the usual **curvature tensor** of ∇,

$$R^\nabla\in\Omega^2(M;\operatorname{End}TM)$$

$$R^\nabla(X,Y)Z=([\nabla_X,\nabla_Y]-\nabla_{[X,Y]})Z$$

and the usual **Ricci tensor**

$$\mathsf{ric}^\nabla\ \text{ or }\ r^\nabla\in\Gamma(T^*M^{\odot 2})$$

$$\mathsf{ric}^\nabla(X,Y):=\operatorname{tr}R^\nabla(\cdot,X)Y$$

the symmetry of ric^∇ following from the Bianchi identities,

$$R^\nabla(X,Y)Z+R^\nabla(Y,Z)X+R^\nabla(Z,X)Y=0 \qquad \text{(Bianchi I)}$$

$$(\nabla_XR^\nabla)(Y,Z)+(\nabla_YR^\nabla)(Z,X)+(\nabla_ZR^\nabla)(X,Y)=0 \qquad \text{(Bianchi II)}$$

Since the Ricci tensor is symmetric and our only natural tensor on M is the skew-symmetric Poisson 2-tensor, we conclude that symplectic manifolds have no 'scalar curvature' (see Bieliavsky et al. [17]), though we will introduce something with that name which will nevertheless be useful for our purposes. We thus introduce the symplectic counterparts of the curvature and Ricci tensors.

Definition 22 The **symplectic curvature tensor** is the $(0,4)$-tensor

$$S^\nabla \in \Gamma(T^*M^{\otimes 4})$$
$$S^\nabla(X,Y,Z,W) := \omega(R^\nabla(X,Y)Z,W)$$

and the $(0,2)$-tensor **symplectic Ricci tensor** (see Habermann and Habermann [54])

$$\mathsf{sric}^\nabla \in \Gamma(T^*M^{\otimes 2})$$
$$\mathsf{sric}^\nabla(X,Y) := \sum_{j=1}^{n} S^\nabla(\mathsf{e}_j,\mathsf{f}_j,X,Y)$$

in a local symplectic frame $(\mathsf{e}_1,\dots,\mathsf{e}_n,\mathsf{f}_1,\dots,\mathsf{f}_n)$. Finally, the **symplectic scalar curvature** of ∇ with respect to an ω-compatible almost complex structure $J \in \mathcal{J}(M,\omega)$ is

$$\mathsf{sR}^\nabla := \sum_{j=1}^{2n} \mathsf{sric}^\nabla(\mathsf{e}_j,\mathsf{e}_j)$$

in a local unitary frame $(\mathsf{e}_1,\dots,\mathsf{e}_{2n}) \in P_{\mathrm{U}(n)}$ for M. ■

To see the first Bianchi identity for the symplectic curvature tensor, recall the **Koszul long exact sequence** for any vector space V (see Bieliavsky et al. [17]):

$$0 \to S^qV \xrightarrow{a} V\otimes S^{q-1}V \xrightarrow{a} \textstyle\bigwedge^2 V\otimes S^{q-2}V \xrightarrow{a} \cdots \xrightarrow{a} \textstyle\bigwedge^{q-1}V\otimes V \xrightarrow{a} \textstyle\bigwedge^q V \to 0$$

where a is the **skew-symmetrization operator**

$$a : \textstyle\bigwedge^r V\otimes S^sV \to \textstyle\bigwedge^{r+1} V\otimes S^{s-1}V$$
$$a\big((v^1\wedge\cdots\wedge v^r)\otimes(w^1\odot\cdots\odot w^s)\big) := \sum_{j=1}^{s}(v^1\wedge\cdots\wedge v^r\wedge w^j)\otimes(w^1\odot\cdots\odot\hat{w}^j\odot\cdots\odot w^s)$$

and satisfies $a^2 = 0$. The associated **symmetrization operator** s is given by

$$s : \bigwedge^r V \otimes S^s V \to \bigwedge^{r+1} V \otimes S^{s-1} V$$
$$s\big((v^1 \wedge \cdots \wedge v^r) \otimes (w^1 \odot \cdots \odot w^s)\big) := \sum_{j=1}^{s} (v^1 \wedge \cdots \wedge \hat{v}^j \wedge \cdots \wedge v^r \wedge w^j) \otimes (v^j \odot w^1 \odot \cdots \odot w^s)$$

Then $s^2 = 0$ and $(a \circ s + s \circ a)|_{\bigwedge^r V \otimes S^s V} = (r+s)\,\mathrm{id}$, and the **first Bianchi identity for** S^∇ at a point $p \in M$

$$S_p^\nabla(X, Y, Z, W) + S_p^\nabla(Y, Z, X, W) + S_p^\nabla(Z, X, Y, W) = 0 \tag{2.136}$$

becomes

$$S_p^\nabla \in \ker a \subseteq \bigwedge^2 T_p^* M \otimes S^2 T_p^* M$$

and the space of 4-tensors satisfying this relation is isomorphic to $(T_pM \otimes S^3 T_pM)/S^4 T_pM$, as can be seen by the action of $\mathrm{Sp}(T_pM, \omega)$ on $T_pM \otimes S^3 T_pM$. If $\dim T_pM = 2n \geq 4$, this action decomposes into three irreducible subspaces

$$S^4 T_pM \oplus S'^2 T_pM \oplus W$$

where $S'^2 T_pM = a(s(\omega_p \otimes S^2 T_pM)) \sim S^2 T_pM$, so that

$$S_p^\nabla = \mathcal{E}_p \oplus \mathcal{W}_p$$

The decomposition of the symplectic curvature tensor into its $\mathcal{E}_p$-component, denoted sE_p^∇ and its $\mathcal{W}_p$-component, denoted sW_p^∇,

$$S_p^\nabla = sE_p^\nabla + sW_p^\nabla$$

is given by

$$\begin{aligned} sE^\nabla(X, Y, Z, W) = \frac{1}{2(n+1)} \Big[& 2\omega_p(X, Y) \mathsf{ric}_p^\nabla(Z, W) + \omega_p(X, Z) \mathsf{ric}_p^\nabla(Y, W) \\ & + \omega_p(X, W) \mathsf{ric}^\nabla(Y, Z) - \omega_p(Y, Z) \mathsf{ric}^\nabla(X, W) - \omega_p(Y, W) \mathsf{ric}^\nabla(X, Z) \Big] \end{aligned}$$

The corresonding curvature tensor then has the form

$$R_p^\nabla = E_p^\nabla + W_p^\nabla$$

where

$$
\begin{aligned}
&E^{\nabla}(X,Y)Z \\
&\quad = \frac{1}{2(n+1)}\Big[2\omega(X,Y)\rho^{\nabla}Z + \omega(X,Z)\rho^{\nabla}Y - \omega(Y,Z)\rho^{\nabla}X + \omega(X,\rho^{\nabla}Z)Y - \omega(Y,\rho^{\nabla}Z)X\Big]
\end{aligned}
$$

with $\rho^{\nabla} \in \Gamma(\mathrm{End}\,TM)$ is given by

$$\omega(X,\rho^{\nabla}Y) = \mathsf{ric}^{\nabla}(X,Y)$$

For the next lemma we refer to Habermann and Habermann [54].

Lemma 8 *For any $X, Y, Z \in \Gamma(TM)$ and any connection $\nabla \in \mathcal{C}(M)$, the relation between curvature and torsion is:*

$$
\begin{aligned}
R^{\nabla}(X,Y)Z + & R^{\nabla}(Y,Z)X + R^{\nabla}(Z,X)Y \\
&= (\nabla_X T^{\nabla})(Y,Z) + (\nabla_Y T^{\nabla})(Z,X) + (\nabla_Z T^{\nabla})(X,Y) \\
&\qquad T^{\nabla}\big(T^{\nabla}(X,Y),Z\big) + T^{\nabla}\big(T^{\nabla}(Y,Z),X\big) + T^{\nabla}\big(T^{\nabla}(Z,X),Y\big)
\end{aligned}
$$

■

Lemma 9 *In a local symplectic frame $(\mathsf{e}_1,\ldots,\mathsf{e}_n,\mathsf{f}_1,\ldots,\mathsf{f}_n)$ we have*

$$\mathsf{ric}^{\nabla}(X,Y) = \sum_{j=1}^{n} S^{\nabla}(\mathsf{e}_j,X,Y,\mathsf{f}_j) - S^{\nabla}(\mathsf{f}_j,X,Y,\mathsf{e}_j)$$

That is, the Ricci tensor is obtained by contracting the symplectic curvature tensor S^{∇} with respect to ω.

Proof: We have

$$
\begin{aligned}
\mathsf{ric}^{\nabla}(X,Y) &= \mathrm{tr}(R^{\nabla}(\cdot,X)Y \\
&= \sum_{j=1}^{n} {}^{j}R^{\nabla}(\mathsf{e}_j,X)Y + {}^{j}R^{\nabla}(\mathsf{f}_j,X)Y \\
&= \sum_{j=1}^{n} \omega\big(R^{\nabla}(\mathsf{e}_j,X)Y,\mathsf{f}_j\big) - \omega\big(R^{\nabla}(\mathsf{f}_j,X)Y,\mathsf{e}_j\big) \\
&= \sum_{j=1}^{n} S^{\nabla}(\mathsf{e}_j,X,Y,\mathsf{f}_j) - S^{\nabla}(\mathsf{f}_j,X,Y,\mathsf{e}_j)
\end{aligned}
$$

which proves the claim. ■

For the next proposition, we follow the proofs of Vaisman [103] and Habermann and Habermann [54].

Proposition 59 *Let (M,ω) and let $\nabla \in \mathcal{C}(M)$ be any connection. Then,*

(1) *If $\nabla \in \mathcal{C}(M)$ is any connection, then S is skew-symmetric in its first two terms,*

$$S^{\nabla}(X,Y,Z,W) = -S^{\nabla}(Y,X,Z,W)$$

(2) *If $\nabla \in \mathcal{C}(M,\omega)$ is symplectic, then S is symmetric in its last two terms,*

$$S^{\nabla}(X,Y,Z,W) = S^{\nabla}(X,Y,W,Z)$$

(3) *If $\nabla \in \mathcal{C}_0(M,\omega)$ is symplectic and torsion-free, then the sum of S with values cyclically permuted is zero:*

$$S^{\nabla}(X,Y,Z,W) + S^{\nabla}(Y,Z,W,X) + S^{\nabla}(Z,W,X,Y) + S^{\nabla}(W,X,Y,Z) = 0$$

(4) *If $\nabla \in \mathcal{C}(M,\omega)$ and $\nabla J = 0$ for some $J \in \mathcal{J}(M,\omega)$, then*

$$S^{\nabla}(X,Y,JZ,JW) = S^{\nabla}(X,Y,Z,W) \quad \textit{and} \quad S^{\nabla}(X,Y,JZ,W) = -S^{\nabla}(X,Y,JW,Z)$$

(5) *If $\nabla \in \mathcal{C}(M,\omega)$, the sric^{∇} is symmetric.*

(6) *If $\nabla \in \mathcal{C}_0(M,\omega)$, then*

$$\mathsf{ric}^{\nabla} = \mathsf{sric}^{\nabla}$$

in which case sric^{∇} is symmetric.

∎

Proof: (1) This follows from the corresponding property of R^∇. (2) If $\nabla \in \mathcal{C}(M,\omega)$ is symplectic, then

$$\begin{aligned} 0 = XY\big(\omega(Z,W)\big) - YX\big(\omega(Z,W)\big) - [X,Y]\big(\omega(Z,W)\big) \\ &= X\big(\omega(\nabla_Y Z,W) + \omega(Z,\nabla_Y W)\big) - Y\big(\omega(\nabla_X Z,W) + \omega(Z,\nabla_X W)\big) \\ &\quad - Y\big(\omega(\nabla_{[X,Y]}Z,W) - \omega(Z,\nabla_{[X,Y]}W)\big) \\ &= \omega(\nabla_X\nabla_Y Z,W) + \omega(Z,\nabla_X\nabla_Y W) - \omega(\nabla_Y\nabla_X Z,W) \\ &\quad - \omega(X,\nabla_Y\nabla_X W) - \omega(\nabla_{[X,Y]}Z,W) - \omega(Z,\nabla_{[X,Y]}W) \\ &= \omega(R^\nabla(X,Y)Z,W) + \omega(Z,R^\nabla(X,Y)W) \\ &= S^\nabla(X,Y,Z,W) - S^\nabla(X,Y,W,Z) \end{aligned}$$

(3) If additionally $\nabla \in \mathcal{C}_0(M,\omega)$, then by Bianchi I

$$S^\nabla(X,Y,Z,W) + S^\nabla(Y,Z,X,W) + S^\nabla(Z,X,Y,W) = 0,$$
$$S^\nabla(X,Z,W,Y) + S^\nabla(Z,W,X,Y) + S^\nabla(W,X,Z,Y) = 0$$

Then, applying (2),

$$S^\nabla(X,Y,Z,W) + S^\nabla(Y,Z,W,X) = -S^\nabla(Z,X,Y,W),$$
$$S^\nabla(Z,W,X,Y) + S^\nabla(W,X,Y,Z) = -S^\nabla(X,Z,Y,W)$$

Together with (1) this yields (3). (4) $\nabla J = 0$ implies that J commutes with ∇, so we immediately have $R^\nabla(X,Y)\circ J = J \circ R^\nabla(X,Y)$ for all $X,Y \in \Gamma(TM)$, and therefore

$$\begin{aligned} S^\nabla(X,Y,JZ,JW) &= \omega\big(R^\nabla(X,Y)(JZ),JW\big) \\ &= \omega\big(J(R^\nabla(X,Y)Z),JW\big) \\ &= \omega\big(R^\nabla(X,Y)Z),W\big) \\ &= S^\nabla(X,Y,Z,W) \end{aligned}$$

Consequently, by (2),

$$S^\nabla(X,Y,JZ,W) = -S^\nabla(X,Y,Z,JW) = -S^\nabla(X,Y,JW,Z)$$

(5) This follows from (2) and the last lemma. (6) Again, by the last lemma and Bianchi I,

$$\mathsf{ric}^{\nabla} = -\sum_{j=1}^{n}\Big(S^{\nabla}(X,\mathsf{e}_j,\mathsf{f}_j,Y) - S^{\nabla}(\mathsf{f}_j,X,\mathsf{e}_j,Y)\Big) = \sum_{j=1}^{n} S^{\nabla}(\mathsf{e}_j,\mathsf{f}_j,X,Y) = \mathsf{sric}^{\nabla}(X,Y) \qquad \blacksquare$$

Thus the symplectic curvature tensor S^{∇} is an element of $\Gamma(\bigwedge^2 T^*M\otimes T^*M^{\odot 2}) = \Omega^2(T^*M^{\odot 2})$.

Chapter 3

Self-Adjoint and Elliptic Operators

3.1 Function Spaces and Differential Operators on Manifolds

We review, in this section, the definitions and basic properties of Sobolev spaces and distributions on manifolds, as well as (pseudo-)differential operators on these spaces and their symbols. We begin with a review of the situation in $\mathbb{R}^n$ before looking at the sheaf-theoretic extension to manifolds and vector bundles. The main sources for the standard theory of distributions and Sobolev spaces on Euclidean space were Leoni [74], Treves [102], Duistermaat and Kolk [29] and Adams [1]. For distributions and Sobolev spaces on manifolds, we used Ban and Crainic [104] and Shubin [95]. For differential operators and elliptic theory the major sources were Liviu Nicolaescu [85], Ban and Crainic [104] and Michael Taylor [100].

3.1.1 Function Spaces in the Euclidean Space Setting

We will give a brief review of the constructions and basic properties of distributions and Sobolev spaces in open sets Ω of $\mathbb{R}^n$, referring to the sources mentioned above for detailed proofs.

3.1.1.1 Distributions

Let $\Omega \subseteq \mathbb{R}^n$ be an open subset, and let us consider the spaces $C_c^\infty(\Omega)$ and $C^\infty(\Omega)$. $C_c^\infty(\Omega)$, for example, may be given various topologies, such as the uniform norm topology generated by $\|\cdot\|_\infty$, or the norm topology generated by the L^2-norm induced by the L^2-inner product $(f,g)_{L^2} := \int_\Omega f\overline{g}\,d\mu$. The first topology is not complete, because, for example, it is dense in $C_c(\Omega)$ with the uniform topology, while the second is not complete because it is dense in $L^2(\Omega)$. Similar problems arise with $C^\infty(\Omega)$. What is needed is a topology on these spaces that makes them complete. This is the motivation for introducing locally convex topologies and inductive limit topologies. After reviewing this basic theory, we consider the dual spaces of these spaces, the spaces of distributions.

Definition 23 Given an open set Ω in $\mathbb{R}^n$, let $\mathcal{K}_\Omega$ denote the set of all compact subsets K of Ω. For fixed $K \in \mathcal{K}_\Omega$, define

$$\mathcal{D}_K(\Omega) := \{f \in C_c^\infty(\Omega) \mid \operatorname{supp}(f) \subseteq K\} \tag{3.1}$$

and define a countable family of seminorms $\|\cdot\|_{K,j}$ on $\mathcal{D}_K(\Omega)$ indexed by $j \in \mathbb{N}_0$,

$$\|f\|_{K,j} := \sup_{x\in K, |\alpha|\leq j} |D^\alpha f(x)| \tag{3.2}$$

Equipped with the locally convex topology generated by these seminorms, $\mathcal{D}_K(\Omega)$ becomes a Fréchet space, a locally convex space that is a complete metric spaces, with metric

$$d_K(f,g) := \max_{j\in\mathbb{N}_0} \frac{1}{2^j}\left(\frac{\|f-g\|_{K,j}}{1+\|f-g\|_{K,j}}\right) \quad \text{or} \quad \sum_{j\in\mathbb{N}_0} \frac{1}{2^j}\left(\frac{\|f-g\|_{K,j}}{1+\|f-g\|_{K,j}}\right)$$

(see Reed and Simon [89, Theorem V.5], Leoni [74, p. 257]). Taking the union over all $K \in \mathcal{K}_\Omega$ we get $C_c^\infty(\Omega)$ as a set, whcih we denote by

$$\mathcal{D}(\Omega) := \bigcup_{K\in\mathcal{K}_\Omega} \mathcal{D}_K(\Omega) \tag{3.3}$$

because we put a different topology on it, the inductive limit topology on it (Reed and Simon [89, Theorem V.15, p. 146]). With this topology $\mathcal{D}(\Omega)$ becomes complete but not metrizable, hence not Fréchet, space, though it is a Montel space (Leoni [74, Theorem 9.8], Treves [102, Proposition 34.4]). ■

Definition 24 Putting the same seminorms $\|\cdot\|_{K,j}$ directly on $C^\infty(\Omega)$ we get a locally convex topology that makes $C^\infty(\Omega)$ a locally convex space which is also complete, hence metrizable (Treves [102, pp. 85-89]). The space $C^\infty(\Omega)$ equipped with this topology is denoted

$$\mathcal{E}(\Omega) \tag{3.4}$$

It is also a Montel space (Treves [102, Proposition 34.4]). We remark that, while $\mathcal{D}(\Omega)$ was constructed as a union of the spaces $\mathcal{D}_K(\Omega)$, $\mathcal{E}(\Omega)$ is not such a union, though it is the intersection of all $C^k(\Omega)$, where $k \in \mathbb{N}$. The spaces $C^k(\Omega)$ can also be given the locally convex topology generated by the seminorms $\|\cdot\|_{K,j}$ where $j \leq k$, which will make them Fréchet spaces (Treves [102, pp. 85-89]), and by letting $C^\infty(\Omega) = \bigcap_{k\in\mathbb{N}_0} C^k(\Omega)$ we get the completeness of $C^\infty(\Omega)$. ■

Definition 25 When $\Omega = \mathbb{R}^n$, there is a locally convex function space lying between the space $\mathcal{E}(\mathbb{R}^n)$ of smooth functions (whose derivatives may be unbounded at infinity) and the space $\mathcal{D}(\mathbb{R}^n)$ of smooth functions that are identically zero outside of a compact set. This space is the **Schwartz space** or **space of rapidly decreasing functions**, which is a subspace of $C^\infty(\mathbb{R}^n)$ defined as follows: for all $\alpha, \beta \in \mathbb{N}_0^n$, define the (extended-real valued) seminorms $\|\cdot\|_{\alpha,\beta}$ on $C^\infty(\mathbb{R}^n)$ by

$$\|f\|_{\alpha,\beta} := \sup_{\mathbf{x}\in\mathbb{R}^n} |\mathbf{x}^\alpha D^\beta f(\mathbf{x})| \tag{3.5}$$

Then the Schwartz space is defined to be the subset of smooth functions $f \in C^\infty(\mathbb{R}^n)$ for which $\|f\|_{\alpha,\beta}$ is finite for *all* $\alpha, \beta \in \mathbb{N}_0^n$,

$$\mathcal{S}(\mathbb{R}^n) := \{f \in C^\infty(\mathbb{R}) \mid \|f\|_{\alpha,\beta} < \infty,\ \forall \alpha, \beta \in \mathbb{N}_0^n\} \tag{3.6}$$

Thus, like the compactly supported smooth functions, Schwartz functions decay at infinity, though they are not necessarily identically zero there. They decay only polynomially at infinity. The family of seminorms $\|\cdot\|_{\alpha,\beta}$ makes $\mathcal{S}(\mathbb{R}^n)$ a locally convex space, and since this collection is countable, $\mathcal{S}(\mathbb{R}^n)$ is metrizable, and in fact complete with respect to the metric, making it a Fréchet space (Theorems V.5 and V.9, Reed and Simon [89]). ■

Remark 41 Clearly we have $C_c^\infty(\mathbb{R}) \subset \mathcal{S}(\mathbb{R}^n) \subset C^\infty(\mathbb{R}^n)$, so as sets

$$\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset \mathcal{E}(\mathbb{R}^n)$$

Now, $C_c(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for all $p \in [1, \infty)$ (Theorem 7.9, Folland [35]), and $C_c^\infty(\mathbb{R}^n)$ is dense in $C_c(\mathbb{R}^n)$ (Theorem 8.14 (b), Folland [34]). As a result $C_c^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$

(Proposition 8.17, Folland [34], though the denseness of $C_c^\infty(\mathbb{R}^n)$ can also be shown directly by means of convolution with bump function). Since $C_c^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$, we must have that $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ as well, once we know that $\mathcal{S}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$ for all $p \in [1, \infty)$ (in fact, $p \in [1, \infty]$). To see this, consider the cases $p = 1$, $p = \infty$ and $p \in (1, \infty)$, separately. For $p = \infty$, take $\alpha = \beta = (0, \ldots, 0)$, then $\|f\|_\infty = \|f\|_{\alpha,\beta} < \infty$ for any $f \in \mathcal{S}(\mathbb{R}^n)$. For $p = 1$, note that $1/(1 + x^2)$ is in $L^1(\mathbb{R})$ (being the derivative of $\tan^{-1}$, which is improper-Riemann integrable), so Fubini's theorem gives for any $f \in \mathcal{S}(\mathbb{R}^n)$ that

$$
\begin{aligned}
\|f\|_1 = \int_{\mathbb{R}^n} |f|\,d\mu = \int_{\mathbb{R}^n} \Big(\prod_{i=1}^n (1+x_i^2)|f|\Big)\Big(\prod_{i=1}^n \frac{1}{1+x_i^2}\Big)\,d\mu \\
\le \Big(\prod_{i=1}^n \|f\|_\infty + \|x_i^2 f\|_\infty\Big)\Big(\prod_{i=1}^n \int_{\mathbb{R}} \frac{1}{1+x_i^2}\,dx_i\Big) \\
= \Big(\prod_{i=1}^n \|f\|_\infty + \|x_i^2 f\|_\infty\Big)\pi^n < \infty
\end{aligned}
$$

which shows that $f \in L^1(\mathbb{R}^n)$. For $p \in (1, \infty)$ and $f \in \mathcal{S}(\mathbb{R}^n)$ we have

$$
\begin{aligned}
\|f\|_p = \Big(\int_{\mathbb{R}^n} |f|^p\,d\mu\Big)^{1/p} = \Big(\int_{\mathbb{R}^n} (|f|^{1/p}|f|^{1-1/p})^p\,d\mu\Big)^{1/p} \\
\le \Big(\int_{\mathbb{R}^n} (|f|^{1/p}\|f\|_\infty^{1-1/p})^p\,d\mu\Big)^{1/p} = \|f\|_\infty^{1-1/p}\|f\|_1^{1/p} < \infty
\end{aligned}
$$

so $f \in L^p(\mathbb{R}^n)$. ■

Remark 42 The importance of the denseness of $\mathcal{S}(\mathbb{R}^n)$ in $L^2(\mathbb{R}^n)$ lies in the fact that the Fourier transform, which is continuous and invertible on $\mathcal{S}(\mathbb{R}^n)$, extends to a unitary operator on $L^2(\mathbb{R}^n)$, which we will describe below. ■

Definition 26 Now that we have the spaces of test functions $\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset \mathcal{E}(\mathbb{R}^n)$, with $\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$ for all $p \in [1, \infty)$, so in particular for $p = 2$, with appropriate locally convex topologies making them complete spaces, we can consider their *continuous duals*, which are the different spaces of **distributions**: The continuous dual

$$\mathcal{D}'(\Omega) := \mathcal{B}(\mathcal{D}(\Omega), \mathbb{C}) \tag{3.7}$$

of the compactly supported smooth test functions is called the space of **distributions** or **generalized functions**, and we put the w^*-topology of pointwise convergence on $\mathcal{D}'(\Omega)$,

$$\sigma\big(\mathcal{D}(\Omega)^*, \mathcal{D}(\Omega)\big) \tag{3.8}$$

which is the weak topology on $\mathcal{D}(\Omega)^*$ generated by the double dual $\mathcal{D}(\Omega)^{**}$ of $\mathcal{D}(\Omega)$. This topology has the property that, if $(f_\alpha)_{\alpha\in A}$ is a net in $\mathcal{D}(\Omega)^*$, then $f_\alpha \xrightarrow{w^*} f$, i.e. f_α w^*-converges to f, iff $f_\alpha(\varphi) \to f(\varphi)$ for all $\varphi \in \mathcal{D}(\Omega)$. This w^*-convergence is called **convergence in the sense of distributions**.

Next, consider the continuous dual of the Schwartz space, called the space of **tempered distributions**,

$$\mathcal{S}'(\mathbb{R}^n) := \mathcal{B}(\mathcal{S}(\mathbb{R}^n), \mathbb{C}) \tag{3.9}$$

which is also given the w^*-topology. Finally, the continuous dual of $\mathcal{E}(\Omega)$,

$$\mathcal{E}'(\Omega) := \mathcal{B}(\mathcal{E}(\Omega), \mathbb{C}) \tag{3.10}$$

is called the space of **compactly supported distributions**. Since dualizing inverts the inclusions, we have

$$\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset \mathcal{E}(\mathbb{R}^n) \tag{3.11}$$

$$\mathcal{E}'(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n) \tag{3.12}$$

Moreover, there are inclusions $\mathcal{E}(\Omega) \hookrightarrow \mathcal{D}'(\Omega)$ and $\mathcal{D}(\Omega) \hookrightarrow \mathcal{E}'(\Omega)$, given by

$$f \mapsto T_f := \int_{\mathbb{R}^n} (\cdot) f \, d\mu$$

Combined with the inclusions $\mathcal{D}(\Omega) \hookrightarrow \mathcal{E}(\Omega)$ and $\mathcal{E}'(\Omega) \hookrightarrow \mathcal{D}'(\Omega)$, we have the following commutative diagram:

$$\begin{tikzcd}
\mathcal{D}(\Omega) \arrow[r, hook] \arrow[d, hook, "f \mapsto T_f"'] & \mathcal{E}(\Omega) \arrow[d, hook, "g \mapsto T_g"] \\
\mathcal{E}'(\Omega) \arrow[r, hook] & \mathcal{D}'(\Omega)
\end{tikzcd}$$

The arrows are continuous algebraic inclusions, and the spaces on the left are topologically the compactly supported version of the spaces on the right (see van den Ban and Crainic [104, p. 27]). ∎

The following theorem is proved in Leoni, [74, Theorems 9.8, 9.10].

Theorem 23 *The inductive limit topology on $C_c^\infty(\Omega)$ defining $\mathcal{D}(\Omega)$ has the following equivalent properties for any (not necessarily bounded/continuous) linear functional $u : \mathcal{D}(\Omega) \to \mathbb{C}$:*

(1) *u is continuous, i.e. $u \in \mathcal{D}(\Omega)^*$.*

(2) *u is bounded, i.e. u sends topologically bounded sets to topologically bounded sets (a subset E of a topological vector space V is topologically bounded if for every neighborhood U of 0 there is a $t > 0$ such that $E \subseteq tU$).*

(3) *If $\varphi_n \to \varphi$ in $\mathcal{D}(\Omega)$, then $u(\varphi_n) \to u(\varphi)$, i.e. u preserves limits.*

(4) *$u|_{\mathcal{D}_K(\Omega)}$ is continuous for all $K \in \mathcal{K}_\Omega$.*

(5) *For every $K \in \mathcal{K}_\Omega$ there is a $j \in \mathbb{N}_0$ and a constant $C_K > 0$ such that*

$$|u(\varphi)| \leq C_K \|\varphi\|_{K,j}, \quad \forall \varphi \in \mathcal{D}_K(\Omega)$$

That is, u is equivalently continuous, bounded, and convergence-preserving in a way compatible with restriction to each $\mathcal{D}_K(\Omega)$ for all $K \in \mathcal{K}_\Omega$. Furthermore, a sequence φ_n in $\mathcal{D}(\Omega)$ converges in the inductive limit topology to $\varphi \in \mathcal{D}(\Omega)$ iff the following two conditions hold:

(a) *There exists a $K \in \mathcal{K}_\Omega$ such that $\operatorname{supp}(\varphi_n) \subseteq K$ for all $n \in \mathbb{N}$.*

(b) *$D^\alpha \varphi_n \to D^\alpha \varphi$ uniformly for all $\alpha \in \mathcal{N}_0^n$.*

∎

Remark 43 The assignment $\Omega \mapsto \mathcal{D}'(\Omega)$ defines a *sheaf*, for if $\{\Omega_i\}_{i \in I}$ is an open cover of Ω and $u_i \in \mathcal{D}'(\Omega_i)$ are distributions on Ω_i satisfying $u_i|_{\Omega_i \cap \Omega_j} = u_j|_{\Omega_i \cap \Omega_j}$, then by using partition of unity $\{\rho_i\}_{i \in I}$ subordinate to the open cover there exists a unique distribution on Ω such that $u|_{\Omega_i} = u_i$, namely $u = \sum_{i \in I} u_i(\rho_i \cdot)$, by which we mean $u(\varphi) := \sum_{i \in I} u_i(\rho_i \varphi)$. Now, as with any sheaf, one can talk about sections with compact support. Note that $u \in \mathcal{D}'(\Omega)$ has **compact support** $\mathrm{supp}(u) \in \mathcal{K}_\Omega$ iff $u(\varphi) = 0$ for φ with support outside of the compact set $\mathrm{supp}(u)$. This is equivalent (see Knapp [67, Theorem 5.1]) to u being an element of $\mathcal{E}'(\Omega)$, or equivalently in the image of the inclusion $\mathcal{E}'(\Omega) \hookrightarrow \mathcal{D}'(\Omega)$, hence the name 'compactly supported distributions' for $\mathcal{E}'(\Omega)$. ■

Remark 44 $\mathcal{D}'(\Omega)$ is an $\mathcal{E}(\Omega)$-module,

$$\mathcal{E}(\Omega) \times \mathcal{D}'(\Omega) \to \mathcal{D}'(\Omega)$$

$$(f, u) \mapsto fu := u \circ f$$

that is, $(fu)(\varphi) := u(f\varphi)$ for all $\varphi \in \mathcal{D}(\Omega)$. ■

Definition 27 Let us now introduce the **derivative of a distribution**, also called the **distributional** or **weak derivative**. To motivate the definition, note that the partial derivative $D_j = \partial/\partial x_j$ operator is skew-symmetric on $L^2(\mathbb{R}^n)$, with domain $\mathcal{D}(D_j) := C_c^\infty(\mathbb{R}^n)$[1] , which follows from integration by parts,

$$(D_j f, g)_{L^2(\mathbb{R}^n)} = -(f, D_j g)_{L^2(\mathbb{R}^n)}$$

for all $f, g \in C_c^\infty(\mathbb{R}^n)$. Consequently, for all multi-indices $\alpha \in \mathbb{N}_0^n$ we have

$$(D^\alpha f, g)_{L^2(\mathbb{R}^n)} = (-1)^{|\alpha|}(f, D^\alpha g)_{L^2(\mathbb{R}^n)}$$

on $C_c^\infty(\mathbb{R}^n)$. Now, if $u \in \mathcal{D}'(\Omega)$ is a distribution, then we emulate this identity, taking it as the definition of $D^\alpha u$:

$$D^\alpha u := (-1)^{|\alpha|} T \circ D^\alpha \tag{3.13}$$

that is,

$$(D^\alpha u)(\varphi) := (-1)^{|\alpha|} u(D^\alpha \varphi), \quad \forall \varphi \in \mathcal{D}(\Omega) \tag{3.14}$$

[1] D_j is, in fact, sekw-adjoint, since iD_j is self-adjoint (see Theorem 37 below), on an appropriate enlargement of $C_c^\infty(\mathbb{R}^n)$.

Since $L^1_{\text{loc}}(\Omega)$ embeds into $\mathcal{D}'(\Omega)$ via $f \mapsto T_f := \int_\Omega (\cdot) f \, d\mu$, we can adapt this notion of a derivative to locally integrable functions. This time, however, we do not simply say that $D^\alpha f := T_f \circ D^\alpha$, but require the extra condition that there be a function $g \in L^1_{\text{loc}}(\Omega)$ such that

$$D^\alpha f = g \textbf{ weakly} \tag{3.15}$$

or g is the α**th weak derivative of** f, meaning, by definition,

$$D^\alpha f := (-1)^{|\alpha|} T_f \circ D^\alpha = T_g \in \mathcal{D}'(\Omega) \tag{3.16}$$

as distributions. In full, this means

$$(f, D^\alpha \varphi)_{L^2(\mathbb{R}^n)} = (-1)^{|\alpha|}(g, \varphi)_{L^2(\mathbb{R}^n)}, \quad \forall \varphi \in \mathcal{D}(\Omega) \tag{3.17}$$

That is, the weak derivative of f can be interpreted in terms of the *adjoint of* D^α, which is thought of as an operator on $L^2(\Omega)$ with core or perhaps initial domain $\mathcal{D}(\Omega)$. We thus say that $f \in L^1_{\text{loc}}(\Omega)$ is weakly differentiable if f lies in the domain $\mathcal{D}((D^\alpha)^*)$ of $(D^\alpha)^*$, in which case the weak αth derivative g of f is the image of f under the adjoint $(D^\alpha)^*$,

$$g := (D^\alpha)^* f$$

(see Section 3.2 for a fuller description of adjoints and self-adjointness). Since iD_j is self-adjoint (see Theorem 37 below), D_j is skew-adjoint, and D^α_j is either self-adjoint or skew-adjoint depending on the parity of $|\alpha|$, since by taking the Fourier transform of D^α_j we get a multiplication operator times $(-i)^{|\alpha|}$ (Proposition 60 below), and any multiplication operator by a real-valued map f is self-adjoint (Proposition 1, p. 259, Reed and Simon [89]), while if is skew-adjoint. Thus, on the domain of D^α, viewed as an unbounded operator acting on $L^2(\Omega)$, there always exist weak derivatives, namely all the functions in the extended domain of D^α. ■

The Schwartz representation theorem characterizes distributions in terms of weak derivatives. It ways that any distribution is the weak derivative of some continuous function. Its proof may be found in Leoni [74, Theorem 9.21].

Theorem 24 (Schwartz Representation Theorem) *Let $\Omega \subseteq \mathbb{R}^n$ be open. For any $u \in \mathcal{D}'(\Omega)$ and any $K \in \mathcal{K}_\Omega$ there is an $f \in C(\Omega)$ and a multi-index $\alpha \in \mathbb{N}_0^n$ such that*

$$u = D^\alpha T_f = (-1)^{|\alpha|}(D^\alpha(\cdot), \overline{f})_{L^2(\Omega)}$$

on $\mathcal{D}_K(\Omega)$. That is,

$$u(\varphi) = (-1)^{|\alpha|} \int_\Omega f D^\alpha \varphi \, d\mu$$

for all $\varphi \in \mathcal{D}_K(\Omega)$. ■

Let us now define the tensor product of distributions, and state the Schwartz kernel theorem, before we move on to Sobolev spaces. There are, of course, many other constructions with distributions, such as convolution and Fourier transform on distributions, but we do not need all of these for our purposes.

Definition 28 Let $X \subseteq \mathbb{R}^n$ and $Y \subseteq \mathbb{R}^m$ be open sets, and consider distributions $u \in \mathcal{D}'(X)$ and $v \in \mathcal{D}'(Y)$. If we had just continuous functions $u \in C(X)$ and $v \in C(Y)$, then we could define their tensor product $u \otimes v \in C(X \times Y)$ by $(u \otimes v)(x, y) := u(x)v(y)$. This idea works for distributions as well: by Theorem 5.5.1 in Hörmander [57], for each $u \in \mathcal{D}'(X)$ and $v \in \mathcal{D}'(Y)$ there is a unique distribution $w \in \mathcal{D}'(X \times Y)$ such that

$$w(f \otimes g) = u(f)v(g)$$

for all $f \in \mathcal{D}(X)$ and $g \in \mathcal{D}(Y)$. On an arbitrary $\varphi \in \mathcal{D}(X \times Y)$ we have

$$w(\varphi) = u \circ v \circ \varphi = v \circ u \circ \varphi$$

i.e. $u(v(\varphi(\cdot, \cdot))) = v(u(\varphi(\cdot, \cdot)))$. ■

Consider the open sets $X \subseteq \mathbb{R}^n$ and $Y \subseteq \mathbb{R}^m$ again. Given any $K \in C_c(X \times Y)$, we can define an **(integral) linear operator**

$$\mathcal{K} : C_c(Y) \to C_c(X)$$
$$(\mathcal{K}\varphi)(x) := \int_Y K(x, y)\varphi(y) \, dy$$

If we embed $\mathcal{K}\varphi$ into $\mathcal{D}(X)$, by $\mathcal{K}\varphi \mapsto T_{\mathcal{K}\varphi}$, then for any $\psi \in \mathcal{D}(X)$ we can write

$$(\mathcal{K}\varphi, \psi) = \int_X \int_Y K(x,y)\varphi(y)\psi(x)\,dy\,dx = (K, \varphi \otimes \psi)$$

by Fubini's theorem. Thus, given a K, we get the above pairing equality. There are two questions that arise: (1) Can we generalize this to distributions, that is can we let $K \in \mathcal{D}'(X \times Y)$? (2) If we start with $\mathcal{K} : \mathcal{D}(Y) \to \mathcal{D}'(Y)$, is there a $K \in \mathcal{D}'(X \times Y)$ for which the pairing equality holds? The answers to both these questions are 'Yes', and they are the content of the Schwartz Kernel Theorem (see Theorems 5.2.1, pp. 128-130, Hörmander [57], or Theorem 6.1, p. 345, Taylor [100]). The distribution K is called an integral kernel, hence the term 'kernel' in the name of the theorem.

Theorem 25 (Schwartz Kernel Theorem) *Let $X \subseteq \mathbb{R}^n$ and $Y \subseteq \mathbb{R}^m$ be open sets. Given a continuous map $\mathcal{K} : \mathcal{D}(Y) \to \mathcal{D}'(Y)$ there exists a unique $K \in \mathcal{D}'(X \times Y)$, called the **Schwartz kernel** of $\mathcal{K}$, such that*

$$(\mathcal{K}\varphi, \psi) = (K, \varphi \otimes \psi) \tag{3.18}$$

and conversely, given $K \in \mathcal{D}'(X \times Y)$, there is a unique continuous $\mathcal{K} : \mathcal{D}(Y) \to \mathcal{D}'(Y)$ giving the pairing equality above. ■

3.1.1.2 The Fourier Transform

Definition 29 We define the **Fourier transform** of a Schwartz function $f \in \mathcal{S}(\mathbb{R}^n)$ to be the function $\mathcal{F}f \equiv \hat{f} : \mathbb{R}^n \to \mathbb{C}$ given by

$$(\mathcal{F}f)(\xi) \equiv \hat{f}(\xi) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)\,dx \tag{3.19}$$

and we define the **inverse Fourier transform** of a Schwartz function $f \in \mathcal{S}(\mathbb{R}^n)$ to be the function $\mathcal{F}^{-1}f \equiv \check{f} : \mathbb{R}^n \to \mathbb{C}$ given by

$$(\mathcal{F}^{-1}f)(x) \equiv \check{f}(x) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ix\cdot\xi} f(\xi)\,d\xi \tag{3.20}$$

By the Fourier Inversion Theorem (Reed and Simon [89, Theorem IX.1, p. 320], $\mathcal{F}$ is a bicontinous bijection of $\mathcal{S}(\mathbb{R}^n)$ onto itself, and its inverse map is the inverse Fourier transform $\mathcal{F}^{-1}$, and by the

Parseval theorem (the Corollary to Theorem IX.1 in Reed and Simon [89]), for any $f \in \mathcal{S}(\mathbb{R}^n)$ we have $\|\mathcal{F}f\|_{L^2(\mathbb{R}^n)} = \|f\|_{L^2(\mathbb{R}^n)}$. By the Plancherel theorem (Theorem IX.6, p. 327, Reed and Simon [89]) $\mathcal{F}$ extends to a unitary map of $L^2(\mathbb{R}^n)$,

$$\mathcal{F} \in \mathrm{U}(L^2(\mathbb{R}^n)) \tag{3.21}$$

This is proved by first extending $\mathcal{F}$ to the space $\mathcal{S}'(\mathbb{R}^n)$ of tempered distributions by

$$\mathcal{F} : \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n) \tag{3.22}$$

$$\mathcal{F}u := u \circ \mathcal{F}, \quad \text{i.e. } (\mathcal{F}u)(\varphi) := u(\mathcal{F}\varphi),\ \forall \varphi \in \mathcal{S}(\mathbb{R}^n) \tag{3.23}$$

then showing that resulting map is a linear bijection of $\mathcal{S}'(\mathbb{R}^n)$ into itself (Theorem IX.2, p. 322, Reed and Simon [89]) and $\mathcal{S}(\mathbb{R}^n)$ is w^*-dense in $\mathcal{S}'(\mathbb{R}^n)$ (Corollary 1, p. 144, Reed and Simon [89]) and dense in $L^2(\mathbb{R}^n)$ in the L^2 Hilbert space topology (which follows from the denseness of $C_c^\infty(\mathbb{R}^n)$ and the fact that $C_c^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset L^2(\mathbb{R}^n)$), so since $\mathcal{F}$ is an isometry when restricted to the dense set $\mathcal{S}(\mathbb{R}^n)$ the result follows.

Proposition 60 *The Fourier transform is an intertwining operator between the position and momentum operators $Q_j f = x_j f$ and $P_j f = -i\partial f/\partial x_j$ on $\mathcal{S}(\mathbb{R}^n)$, and so on the intersection of the domains of Q_j and P_j in $L^2(\mathbb{R}^n)$ (see Section 3.2.2 below),*

$$P_j \circ \mathcal{F} = -\mathcal{F} \circ Q_j \quad \text{and} \quad Q_j \circ \mathcal{F} = -P_j \circ \mathcal{F}$$

that is

$$i\frac{\partial}{\partial \xi_j} \circ \mathcal{F} = \mathcal{F} \circ x_j \quad \text{and} \quad i\xi \circ \mathcal{F} = \mathcal{F} \circ \frac{\partial}{\partial x_j}$$

As a result, if $P(D)$ is a constant-coefficient differential operator, a polynomial in $D_j = -i\partial/\partial x_j$, $j = 1, \ldots, n$, then

$$P(D) \circ \mathcal{F} = \mathcal{F} \circ P(-x)$$

and

$$P(x) \circ \mathcal{F} = \mathcal{F} \circ P(-D)$$

Proof: If $f \in \mathcal{S}(\mathbb{R}^n)$, then

$$\begin{aligned}
-i\frac{\partial}{\partial \xi_j}(\mathcal{F}f)(\xi) &= -i\frac{1}{(2\pi)^{n/2}}\frac{\partial}{\partial \xi_j}\int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)\,dx \\
&= -\frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n}\frac{\partial}{\partial \xi_j} e^{-ix\cdot\xi} f(x)\,dx \\
&= -\frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} x_j e^{-ix\cdot\xi} f(x)\,dx \\
&= -\mathcal{F}(x_j f)(\xi)
\end{aligned}$$

Similarly,

$$\begin{aligned}
i\xi_j(\mathcal{F}f)(\xi) &= \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} i\xi_j e^{-ix\cdot\xi} f(x)\,dx \\
&= -\frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n}\frac{\partial}{\partial x_j}\big(e^{-ix\cdot\xi} f(x)\big)\,dx \\
&= -\mathcal{F}\left(\frac{\partial}{\partial x_j} f\right)(\xi)
\end{aligned}$$

which completes the proof. ■

3.1.1.3 Sobolev Spaces

Let $\Omega \subseteq \mathbb{R}^n$ be an open set, and let $p \in [1,\infty]$. Recall the distributional derivative, which can potentially work on $L^p(\Omega)$ just as well as $L^1_{\text{loc}}(\Omega)$: if $u \in L^p(\Omega)$, then we say that $D^\alpha u = g$ weakly, for some $g \in L^p(\Omega)$ if for all $\varphi \in \mathcal{D}(\Omega)$ we have

$$(u, D^\alpha \varphi)_{L^2(\Omega)} = (-1)^{|\alpha|}(g,\varphi)_{L^2(\Omega)}$$

i.e.

$$\int_\Omega u D^\alpha \varphi\,d\mu = (-1)^{|\alpha|}\int_\Omega g\varphi\,d\mu$$

Definition 30 The **Sobolev space** $W^{1,p}(\Omega)$ is defined to be the space of all $u \in L^p(\Omega)$ such that $D_i u \in L^p(\Omega)$ for all $i = 1,\ldots,n$, where $D_i = \partial/\partial x_i$ and $D_i u$ is the weak derivative,

$$W^{1,p}(\Omega) = \{u \in L^p(\Omega) \mid D_i u \in L^p(\Omega),\ \forall i = 1,\ldots,n\} \tag{3.24}$$

Given $u \in W^{1,p}(\Omega)$ we define its **distributional gradient**

$$\operatorname{grad} u := \left(\frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_n}\right)$$

and using it we may define recursively the other **Sobolev spaces** $W^{k,p}(\Omega)$. First, let

$$W^{1,p}(\Omega, \mathbb{R}^d) := \{u = (u_1, \dots, u_d) \in L^p(\Omega, \mathbb{R}^d) \mid u_i \in W^{1,p}(\Omega),\ \forall i = 1, \dots, d\} \tag{3.25}$$

then define, recursively,

$$W^{k,p}(\Omega) = \{u \in L^p(\Omega) \mid \operatorname{grad} u \in W^{k-1,p}(\Omega, \mathbb{R}^n)\} \tag{3.26}$$

Alternatively,

$$W^{k,p}(\Omega) = \{u \in L^p(\Omega) \mid D^\alpha u \in L^p(\Omega),\ \forall |\alpha| \leq k\} \tag{3.27}$$

When $p = 2$, we use the standard notation

$$H^k(\Omega) := W^{k,2}(\Omega) = \{u \in L^p(\Omega) \mid D^\alpha u \in L^2(\Omega),\ \forall |\alpha| \leq k\} \tag{3.28}$$

For $p \in [1, \infty)$ we have the following equivalent norms on $W^{1,p}(\Omega)$,

(1) $\|u\|_{W^{1,p}(\Omega)} = \|u\|_{L^p(\Omega)} + \|\operatorname{grad} u\|_{L^p(\Omega, \mathbb{R}^n)}$.

(2) $\|u\|_{W^{1,p}(\Omega)} = \|u\|_{L^p(\Omega)} + \sum_{i=1}^n \|D_i u\|_{L^p(\Omega)}$.

(3) $\|u\|_{W^{1,p}(\Omega)} = \left(\|u\|^p_{L^p(\Omega)} + \sum_{i=1}^n \|D_i u\|^p_{L^p(\Omega)}\right)^{1/p}$

while for $p = \infty$ we have

$$\|u\|_{W^{1,\infty}(\Omega)} = \max\{\|u\|_{L^\infty(\Omega)},\ \|D_1 u\|_{L^\infty(\Omega)}, \dots, \|D_n u\|_{L^\infty(\Omega)}\}$$

For $k > 1$ we have, analogously, the equivalent norms

(1) $\|u\|_{W^{k,p}(\Omega)} = \sum_{i=1}^n \|D^\alpha u\|_{L^p(\Omega)}$.

(2) $\|u\|_{W^{1,p}(\Omega)} = \left(\sum_{i=1}^n \|D^\alpha u\|_{L^p(\Omega)}\right)^{1/p}$

For $p \in [1, \infty)$ these norms make $W^{k,p}(\Omega)$ into Banach spaces (Theorem 10.5, Leoni [74], Theorem 3.2, Adams [1]) which are separable and reflexive (Theorem 3.5, Adams [1]), while for $p = 2$ they are even Hilbert spaces, with inner product

$$(u, v)_{H^k(\Omega)} := \sum_{|\alpha| \le k} \int_\Omega (D^\alpha u)(D^\alpha v)\, d\mu \tag{3.29}$$

■

The Sobolev spaces $W^{k,p}(\mathbb{R}^n)$ and $H^k(\mathbb{R}^n)$ were defined for only *positive integers* k. However, we can extend the definitions to $W^{s,p}(\mathbb{R}^n)$ and $H^s(\mathbb{R}^n)$ for all $s \in \mathbb{R}$.

Since the Fourier transform is a unitary operator on $L^2(\mathbb{R}^n)$, and $D^\alpha \circ \mathcal{F} f = \mathcal{F} \circ (-x)^\alpha f$, we have for any $f \in W^{k,2}(\mathbb{R}^n) = H^k(\mathbb{R}^n)$ and k an even integer

$$(1 + \|\xi\|^2)^{k/2} \circ \mathcal{F} f(\xi) = \Big(1 + \sum_{j=1}^n \xi_j^2\Big)^{k/2} \circ \mathcal{F} f(\xi) = -\mathcal{F} \circ \Big(1 + \sum_{j=1}^n D_j^2\Big)^{k/2} f(\xi)$$

since $(1 + \sum_{j=1}^n D_j^2)^{k/2} f \in L^2(\mathbb{R}^n)$ because the operator $(1 + \sum_{j=1}^n D_j^2)^{k/2}$ has order k and $f \in H^k(\mathbb{R}^n)$. Thus, we see that for all $|\alpha| \le k$

$$D_\xi^\alpha \big((1 + \|\xi\|^2)^{k/2} \circ \mathcal{F} f(\xi)\big) \in L^2(\mathbb{R}^n)$$

whence $(1+\|\xi\|^2)^{k/2} \circ \mathcal{F} f \in H^k(\mathbb{R}^n)$. Conversely, if $(1+\|\xi\|^2)^{k/2} \circ \mathcal{F} f \in H^k(\mathbb{R}^n)$, then applying the Fourier transform to $(1 + \sum_{j=1}^n D_j^2)^{k/2} f$ shows that $(1 + \sum_{j=1}^n D_j^2)^{k/2} f \in L^2(\mathbb{R}^n$, so in particular $\mathcal{F} \circ D_x^\alpha f = (-\xi)^\alpha \circ \mathcal{F} f$, whence

$$D^\alpha \big((1 + \|\xi\|^2)^{k/2} \circ \mathcal{F} f\big)(\xi) = D_\xi^\alpha \big((1 + \|\xi\|^2)^{k/2}\big) \circ \mathcal{F} f(\xi) + (1 + \|\xi\|^2)^{k/2} \circ D_\xi^\alpha \mathcal{F} f \in L^2(\mathbb{R}^n)$$

which implies that $f \in H^k(\mathbb{R}^n)$. Thus,

$$f \in H^k(\mathbb{R}^n) \iff (1 + \|\xi\|^2)^{k/2} \circ \mathcal{F} f \in L^2(\mathbb{R}^n \tag{3.30}$$

Recalling that the Laplacian $\Delta_\xi = \sum_{j=1}^n \frac{\partial^2}{\partial \xi_j^2} = -\sum_{j=1}^n \big(-i \frac{\partial}{\partial \xi_j}\big)^2 = -\sum_{j=1}^n D_j^2$, or $-\Delta_\xi = \sum_{j=1}^n D_j^2$, we can rephrase this condition as

$$f \in H^k(\mathbb{R}^n) \iff \mathcal{F} \circ (I - \Delta)^{k/2} f \text{ and } \mathcal{F}^{-1} \circ (1 + \|\xi\|^2)^{k/2} \circ \mathcal{F} f \in L^2(\mathbb{R}^n) \tag{3.31}$$

Since Δ is self-adjoint (its Fourier transform is a multiplication operator, which is self-adjoint), we can define

$$(I-\Delta)^{k/2}$$

for *any* real k. This is our definition of $H^s(\mathbb{R}^n)$, $s \in \mathbb{R}$ (see Shubin [95] for a fuller account):

Definition 31 For any $s \in \mathbb{R}$, define the **Sobolev space of real order** $s \in \mathbb{R}$ to be

$$H^s(\mathbb{R}) := \{f \in L^2(\mathbb{R}^n) \mid \mathcal{F} \circ (I-\Delta)^{s/2} f \in L^2(\mathbb{R}^n)\} \tag{3.32}$$

More generally, for $p \neq 2$ and $s \in \mathbb{R}$, we define

$$W^{s,p}(\mathbb{R}^n) := \{f \in L^p(\mathbb{R}^n) \mid \mathcal{F} \circ (I-\Delta)^{s/2} f \in L^p(\mathbb{R}^n)\} \tag{3.33}$$

These spaces are sometimes called **Bessel potential spaces**, since the operator $(I-\Delta)^{s/2}$ on $L^2(\mathbb{R}^n)$ is also called a **Bessel potential of order** s. These spaces are also Banach spaces, with norm

$$\|f\|_{W^{s,p}(\mathbb{R}^n)} := \|\mathcal{F} \circ (I-\Delta)^{s/2} f\|_{L^p(\mathbb{R})} \tag{3.34}$$

The spaces $H^s(\mathbb{R})$ are Hilbert spaces, with inner product

$$(u, v) := (J^s u, J^s v)_{L^2(\mathbb{R}^n)} \tag{3.35}$$

where $J^s = \mathcal{F} \circ (I-\Delta)^{s/2}$. We thus have

$$H^s(\mathbb{R}^n) = (I-\Delta)^{s/2} L^2(\mathbb{R}^n) \tag{3.36}$$

and

$$W^{s,p}(\mathbb{R}^n) = (I-\Delta)^{s/2} L^p(\mathbb{R}^n) \tag{3.37}$$

These are sometimes taken as the definitions of H^s and $W^{s,p}$. ■

Remark 45 The condition that $f \in L^2(\mathbb{R}^n)$ in the definition of $H^2(\mathbb{R}^n)$ can be weakened substantially. For exaple, Taylor [100, p. 316] takes $f \in \mathcal{S}'(\mathbb{R}^n)$, while van den Ban and Crainic [104, p. 39] take $f \in \mathcal{D}'(\mathbb{R}^n)$. ■

Remark 46 Clearly we have $D_j : H^s(\mathbb{R}^n) \to H^{s-1}(\mathbb{R}^n)$, in fact $f \in H^s(\mathbb{R}^n)$ iff $D_j f \in H^{s-1}(\mathbb{R}^n)$, so in general we have

$$D^\alpha : H^s(\mathbb{R}^n) \to H^{s-|\alpha|}(\mathbb{R}^n)$$

and therefore $f \in H^s(\mathbb{R}^n)$ iff $D^\alpha \in H^{s-|\alpha|}(\mathbb{R}^n)$. ■

We also have the following special Sobolev space, which is a type of local version of $W^{1,p}$.

Definition 32 Let $p \in [1, \infty)$, $\Omega \subseteq \mathbb{R}^n$ an open set, and define the **Sobolev space**

$$L^{1,p}(\Omega) := \{u \in L^1_{\mathrm{loc}}(\Omega) \mid \operatorname{grad} u \in L^p(\Omega, \mathbb{R}^n)\} \tag{3.38}$$

When Ω is connected, then fixing a nonempty open subset $\Omega' \Subset \Omega$ we can define a norm on $L^{1,p}(\Omega)$ by

$$\|u\|_{L^{1,p}(\Omega)} := \|u\|_{L^1(\Omega')} + \|\operatorname{grad} u\|_{L^p(\Omega,\mathbb{R}^n)} \tag{3.39}$$

which makes $L^{1,p}(\Omega)$ a Banach space, and clearly $W^{1,p}(\Omega) \subseteq L^{1,p}(\Omega)$. ■

Definition 33 For an open set Ω in $\mathbb{R}^n$, consider the space $C^k(\overline{\Omega})$ consisting, by definition, of bounded C^k functions $f \in C^k_b(\Omega)$ whose derivatives are bounded and uniformly continuous on Ω for all $|\alpha| \leq k$,

$$C^k(\overline{\Omega}) := \{f \in C^k_b(\Omega) \mid D^\alpha f \in UC_b(\Omega),\ \forall |\alpha| \leq k\} \tag{3.40}$$

It is a Banach space under the norm

$$\|f\|_{C^k(\overline{\Omega})} := \max_{|\alpha| \leq k} \sup_{x \in \Omega} |D^\alpha f(x)| \tag{3.41}$$

(Adams [1, p. 9]). By the recent results of Fefferman [33], which extends those of Whitney [109], we may also define $C^k(\overline{\Omega})$ to be those $f \in C^k(\Omega)$ which can be extended to $\tilde{f} \in C^k(\mathbb{R}^n)$,

$$C^k(\overline{\Omega}) := \{f \in C^k_b(\Omega) \mid f \text{ can be extended to } \tilde{f} \in C^k(\mathbb{R}^n)\} \tag{3.42}$$

The space $C^\infty(\overline{\Omega})$ is similarly defined to be the space

$$C^\infty(\overline{\Omega}) := \{f \in C^\infty(\Omega) \mid f \text{ can be extended to } \tilde{f} \in C^\infty(\mathbb{R}^n)\} \tag{3.43}$$

(see Leoni [74, pp. 560-564]). ■

To see the relationship between $C_c^\infty(\Omega)$, $C^k(\overline{\Omega})$, and $W^{k,p}(\Omega)$, consider the following Sobolev spaces:

Definition 34 Let Ω be an open subset of $\mathbb{R}^n$, and define the space

$$H^{k,p}(\Omega) = \overline{C^\infty(\Omega) \cap W^{k,p}(\Omega)} \tag{3.44}$$

where the closure is in $W^{k,p}(\Omega)$ with respect to the norm $\|\cdot\|_{W^{k,p}(\Omega)}$, or equivalently (see Adams [1]) to be the completion of the subset $\{f \in C^k(\Omega) \mid \|f\|_{W^{k,p}(\Omega)} < \infty\}$ with respect to the Sobolev norm $\|\cdot\|_{W^{k,p}(\Omega)}$. Define also

$$W_0^{k,p}(\Omega) := \overline{C_c^\infty(\Omega)} \tag{3.45}$$

the closure being in $W^{k,p}(\Omega)$ with respect to the norm $\|\cdot\|_{W^{k,p}(\Omega)}$. ■

By Corollary 3.4, Adams [1], we know that $H^{k,p}(\Omega) \subseteq W^{k,p}(\Omega)$. The Meyers and Serrin theorem [78] says that, for $p \in [1,\infty)$ the space $C^\infty(\Omega) \cap W^{k,p}(\Omega)$ is dense in $W^{k,p}(\Omega)$, and therefore

$$H^{k,p}(\Omega) = W^{k,p}(\Omega) \tag{3.46}$$

with no conditions on the boundary $\partial\Omega$! When dealing with $C_c^\infty(\Omega)$, however, we do need to worry about boundary conditions (called the *segment property* or the *class C property*, see Adams [1, p. 54], Leoni [74, p. 286-287]). Under these extra conditions, we also have (Adams [1, Theorem 3.18], Leoni [74, Theorem 10.29]) that, if $r_{\Omega,\mathbb{R}^n}$ is the restriction map to Ω, then $r_{\Omega,\mathbb{R}^n}(C^\infty(\mathbb{R}^n))$ is dense in $W^{k,p}(\Omega)$, i.e.

$$W^{k,p}(\Omega) = \overline{r_{\Omega,\mathbb{R}^n}(C^\infty(\mathbb{R}^n))} \tag{3.47}$$

the closure being in $W^{k,p}(\Omega)$ with respect to the norm $\|\cdot\|_{W^{k,p}(\Omega)}$. As a corollary, we have

$$W_0^{k,p}(\mathbb{R}^n) = W^{k,p}(\mathbb{R}^n) \tag{3.48}$$

Let us now describe the embedding theorems. There are, of course, many variations on the basic themes, but we do not need the full power of the theory. Our interest in the Sobolev and

Rellich embedding theorems is mainly as a tool for reaching elliptic regularity. The simplest version of a Sobolev embedding theorem is the following (found in Taylor [100, Prop. 1.3, Cor. 1.4]). It says that for sufficiently large s, the elements of $H^s(\mathbb{R}^n)$ are C^k in the classical sense:

Theorem 26 (Sobolev Embedding I) *If $s > n/2$, then all functions f in $H^s(\mathbb{R}^n)$ are bounded and continuous, i.e.*

$$H^s(\mathbb{R}^n) \subseteq C_b(\mathbb{R}^n)$$

If $s > n/2 + k$, then

$$H^s(\mathbb{R}^n) \subseteq C^k(\mathbb{R}^n)$$

and consequently

$$\bigcap_{s\in\mathbb{R}} H^s(\mathbb{R}^n) = C^\infty(\mathbb{R}^n) \qquad \blacksquare$$

More enerally (Nicolaescu [85, Theorem 10.2.21]),

Theorem 27 (Sobolev Embedding II) *If $r < s$ and $r - n/p = s - n/q < 0$, then $W^{r,p}(\mathbb{R}^n) \subset W^{s,q}(\mathbb{R}^n)$, and the inclusion is continuous, which follows from the inequality*

$$\|f\|_{W^{s,q}(\mathbb{R}^n)} \leq C\|f\|_{W^{s,q}(\mathbb{R}^n)}$$

for all $f \in W^{r,p}(\mathbb{R}^n)$, for some $C > 0$ depending on r, s, p, q and n. $\blacksquare$

Finally, we have the Rellich-Kondrachov compactness theorem,

Theorem 28 (Rellich-Kondrachov) *Let Ω be an open bounded Lipschitz domain in $\mathbb{R}^n$ (i.e. $\partial\Omega$ is of class C, see Leoni [74]). If $r < s$ and $s - n/p > r - n/q$, the inclusion*

$$W^{s,p}(\Omega) \hookrightarrow W^{r,q}(\Omega)$$

is compact, in the sense of operator theory. In particular,

$$H^s(\Omega) \hookrightarrow H^r(\Omega)$$

is compact. $\blacksquare$

3.1.2 Function Spaces in the Manifold Setting

We wish to define analogs of the locally convex spaces of test functions $\mathcal{D}(\Omega)$, $\mathcal{E}(\Omega)$, and their continuous duals, the space of distributions $\mathcal{D}'(\Omega)$ and the space of compactly supported distributions $\mathcal{E}'(\Omega)$, to spaces of sections $\Gamma(E)$ of complex vector bundles $E \to M$. Then we would like to define analogs Sobolev spaces of such sections, as well. We begin with the first task.

3.1.2.1 Test Functions and Distributions on Manifolds

Let E be a complex (or real) rank p smooth vector bundle over a smooth n-manifold M. We want to think of the space of smooth sections $\Gamma(E)$ in the same way as we do the space of smooth functions $C^\infty(\Omega)$, Ω an open subset of $\mathbb{R}^n$, namely equipped with a locally convex topology generated by a family of seminorms, giving the test functions $\mathcal{E}(\Omega)$. We can proceed locally, over chart domains (U_j, φ_j) trivializing E.

Definition 35 (Locally Convex Topology on $\boldsymbol{\Gamma(E)}$) Let $\{(U_j, \varphi_j)\}_{j\in J}$ be a vector bundle atlas for M, that is, an atlas which also gives local trivializations for E,

$$\tau_j : E|_{U_j} \to U_j \times \mathbb{C}^p$$

then let

$$\gamma := (j, l, K, r), \quad \text{where } j \in J,\ q \le \ell \le p,\ K \in \mathcal{K}_{\varphi_j(U_j)},\ r \in \mathbb{N}_0$$

Here the notation $\mathcal{K}_V$ denotes the set of compact subsets K of any open set V in $\mathbb{R}^n$. If $s \in \Gamma(E)$, let $s_j := s|_{U_j}$ denote its restriction to U_j, let $\pi_\ell : \mathbb{C}^p \to \mathbb{C}$ denote the ℓth projection, and then define the seminorm $\|\cdot\|_\gamma$ by

$$\|s\|_\gamma := \|\pi_\ell(s_j)\|_{K,r} \tag{3.49}$$

where $\|\cdot\|_{K,r}$ is the seminorm on $C^\infty(\Omega)$, $\Omega \subseteq \mathbb{R}^n$ open,

$$\|f\|_{K,j} := \sup_{x\in K, |\alpha|\le j} |D^\alpha f(x)|$$

Equipped with this locally convex topology, the space $\Gamma(E)$ becomes a locally convex space, which we denote by

$$\mathcal{E}(E) \tag{3.50}$$

It is an easy check that the locally convex topology does not depend on the choices involved in the construction of the seminorms, and basically boils down to taking a refinement of any two different trivializing covers of M. Also, as in the Euclidean case, $\mathcal{E}(E)$ becomes a Fréchet space, since the cover can always be chosen to be countable. ■

Remark 47 We observe that a sequence $(s_n)_{n\in\mathbb{N}}$ in $\mathcal{E}(E)$ converges to s in this topology iff, for any open set U in the domain of a chart (U_j, φ_j), and for any compact $K \subseteq U$, writing $s_n = (s_n^1, \ldots, s_n^p)$ and $s = (s^1, \ldots, s^p)$ in a local frame, we have that all the derivatives $D^\alpha s_n^i$ converge uniformly in K to $D^\alpha s^i$ in the local coordinates φ_j. This is just a local coordinate description of the Euclidean case, which is applied componentwise. ■

Definition 36 (Test Sections: Locally Convex Topology on $\boldsymbol{\Gamma_c(E)}$) For each compact subset K of M, define

$$\mathcal{E}_K(E) := \{s \in \Gamma_c(E) \mid \operatorname{supp}(s) \subseteq K\} \tag{3.51}$$

with the locally convex topology induced from $\mathcal{E}(E)$, i.e. generated by the same seminorms $\|\cdot\|_\gamma$ as in the noncompact case $\mathcal{E}(E)$. Letting $\mathcal{K}_M$ be the set of all compact subsets of M, we then define the **space of test sections**

$$\mathcal{D}(E) := \bigcup_{K\in\mathcal{K}_M} \mathcal{E}_K(E) \tag{3.52}$$

endowed with the inductive limit topology, as in the Euclidean case. This is just $\Gamma_c(E)$ endowed with a complete topology. ■

Definition 37 To define the duals of $\mathcal{E}(E)$ and $\mathcal{D}(E)$, we employ the integral over M. If M is orientable, we may suppose this integral to be with respect to some volume/top form (this will be the situation for us, on a symplectic manifold, since these are always orientable), otherwise we must define the integral with respect to some density of a density bundle. Let us suppose our manifold M orientable for simplicity. Then, the integral defines a linear functional on the space of compactly supported top forms $\Omega^n_c(M) = \Gamma_c(\bigwedge^n T^*M)$, taking a top form and integrating it:

$$\int_M : \Omega^n_c(M) \to \mathbb{C} \tag{3.53}$$

$$\alpha \mapsto \int_M \alpha \tag{3.54}$$

Let us denote by $E^\vee$ the "functional dual",

$$E^\vee := E^* \otimes \Omega^n_c(M) \cong \operatorname{Hom}_{C^\infty(M)}(E; \Omega^n_c(M)) \tag{3.55}$$

From $E^\vee$ we get a pairing map

$$\Gamma(E^\vee) \times \Gamma(E) \to \Omega^n_c(M) \tag{3.56}$$

$$(\sigma^* \otimes \alpha, \tau) \mapsto \sigma^*(\tau)\alpha \tag{3.57}$$

and from this we get the canonical pairing

$$[\cdot,\cdot] : \Gamma_c(E^\vee) \times \Gamma(E) \to \mathbb{C} \tag{3.58}$$

$$[\sigma^* \otimes \alpha, \tau] := \int_M (\sigma^* \otimes \alpha, \tau) = \int_M \sigma^*(\tau)\alpha \tag{3.59}$$

Using this pairing, we define the **space of distributional** or **generalized sections of** E to be the continuous dual of $\mathcal{D}(E^\vee)$,

$$\mathcal{D}'(E) := \mathcal{D}(E^\vee)^* \tag{3.60}$$

endowed with the strong topology. By the way $E^\vee$ was constructed we have the canonical inclusion $\mathcal{E}(E) \hookrightarrow \mathcal{D}'(E)$, namely via $s \mapsto T_s := [\cdot, s]$. And as in the Euclidean case, $\mathcal{D}'(E)$ is an $\mathcal{E}(E)$-module, via $(s\alpha)(f) = \alpha(sf)$. ■

Definition 38 The space of **compactly supported distributional** or **generalized sections of** E, is similarly defined bo be the continuous dual of $\mathcal{E}(E^\vee)$,

$$\mathcal{E}'(E) := \mathcal{E}(E^\vee)^* \tag{3.61}$$

and in the same way we get the inclusion $\mathcal{D}(E) \hookrightarrow \mathcal{E}'(E)$, $s \mapsto [\cdot, s]$. ■

3.1.2.2 Generalized Function Spaces on Manifolds

Before we introduce Sobolev spaces on manifolds, we want to explain the general principle. This principle is observed in most texts on analysis on manifolds, but is most clearly articulated in the book by van den Ban and Crainic [104], and it is this treatement that we follow here.

Definition 39 In the Euclidean case we have seen a variety of function spaces, from Schwartz functions to Sobolev spaces to distributions. Some are locally convex only, some are Fréchet, some are Banach spaces, and some are Hilbert spaces. In all these cases, the function space, call it $\mathcal{F}$, sits between the test functions and the distributions, $\mathcal{D}(\mathbb{R}^n) \subseteq \mathcal{F} \subseteq \mathcal{D}'(\mathbb{R}^n)$. Since all these spaces are in particular locally convex topological vector spaces, we begin with this as our notion of a **function space on** $\mathbb{R}^n$. We add only the mild requirement that, in addition to $\mathcal{F}$ being a locally convex TVS (topological vector space), the multiplication-by-test-function map, $m_\varphi : \mathcal{F} \to \mathcal{F}$, $m_\varphi(f) := \varphi f$, be continuous. Explicitly, then, a general function space $\mathcal{F}$ on $\mathbb{R}^n$ satisfies

(1) $\mathcal{F}$ is a locally convex TVS, with generating family Γ of seminorms.

(2) $\mathcal{D}(\mathbb{R}^n) \subseteq \mathcal{F} \subseteq \mathcal{D}'(\mathbb{R}^n)$

(3) $m_\varphi : \mathcal{F} \to \mathcal{F}$, $m_\varphi(f) := \varphi f$, is continuous for all $\varphi \in \mathcal{D}(\mathbb{R}^n)$.

Let us also introduce some variations and notation. Recall that for any topological space X the notation $\mathcal{K}_X$ means the set of all compact subsets of X.

(1) $\mathcal{F}_K := \{u \in \mathcal{F} \mid \operatorname{supp}(u) \subset K\}$ for any $K \in \mathcal{K}_{\mathbb{R}^n}$, with the topology induced from $\mathcal{F}$.

(2) $\mathcal{F}_\mathcal{K} := \bigcup_{K \in \mathcal{K}_{\mathbb{R}^n}} \mathcal{F}_K$, with the inductive limit topology.

(3) $\mathcal{F}_{\text{loc}} := \{u \in \mathcal{D}'(\mathbb{R}^n) \mid \varphi u \in \mathcal{F},\ \forall \varphi \in \mathcal{D}(\mathbb{R}^n)\}$, with the locally convex topology induced by the seminorms $q_{p,\varphi} := p(\varphi u)$, where $p \in \Gamma$. That is, the weakest topology making the multiplication map $m_\varphi : \mathcal{F}_{\text{loc}} \to \mathcal{F}$ continuous.

Example 19

(1) $\mathcal{D} := \mathcal{D}(\mathbb{R}^n) = \mathcal{D}_{\mathcal{K}}(\mathbb{R}^n) = \mathcal{E}_{\mathcal{K}}(\mathbb{R}^n)$.

(2) $\mathcal{E} := \mathcal{E}(\mathbb{R}^n) = \mathcal{D}_{\mathrm{loc}}(\mathbb{R}^n) = \mathcal{E}_{\mathrm{loc}}(\mathbb{R}^n)$.

(3) $\mathcal{S} := \mathcal{S}(\mathbb{R}^n) = \mathcal{S}_{\mathrm{loc}}(\mathbb{R}^n)$

(4) $\mathcal{D}' := \mathcal{D}'(\mathbb{R}^n) = \mathcal{D}'_{\mathrm{loc}}(\mathbb{R}^n) = \mathcal{E}'_{\mathrm{loc}}(\mathbb{R}^n)$.

(5) $\mathcal{E}' := \mathcal{E}'(\mathbb{R}^n) = \mathcal{E}'_{\mathcal{K}}(\mathbb{R}^n) = \mathcal{D}'_{\mathcal{K}}(\mathbb{R}^n)$

The Sobolev spaces satisfy only the strict inclusions

$$H^s_{\mathcal{K}}(\mathbb{R}^n) \subset H^s(\mathbb{R}^n) \subset H^s_{\mathrm{loc}}(\mathbb{R}^n)$$

Here, $H^s_{\mathrm{loc}} := H^s_{\mathrm{loc}}(\mathbb{R}^n) = \{u \in \mathcal{D}' \mid \varphi u \in H^s(\mathbb{R}^n),\ \forall \varphi \in \mathcal{D}\}$ is equipped with the seminorms $q_{\|\cdot\|_s,\psi}(f) := \|\psi f\|_s$, where $\|\cdot\|_s$ is the Sobolev norm induced by the innder product on H^s.

To put such a function space on a complex (or real) vector bundle E over a manifold M which is not necessarily compact, we require extra structure of our function spaces $\mathcal{F}$ on $\mathbb{R}^n$, which we take as *axioms*, in the sense that some or all of these axioms will be satisfied by the function spaces we are most interested in.

(1) (**Banach Axiom**) We say that $\mathcal{F}$ is **Banach** if its topology is a Banach topology. Otherwise, we say that $\mathcal{F}$ is **locally Banach** if for each $K \in \mathcal{K}_{\mathbb{R}^n}$ the topology of $\mathcal{F}_K$ is Banach.

(2) (**Fréchet Axiom**) We similarly say that $\mathcal{F}$ is **Fréchet** if its topology is a Fréchet topology, and we say $\mathcal{F}$ is **locally Fréchet** if each $\mathcal{F}_K$ is Fréchet.

(3) (**Hilbert Axiom**) We say that $\mathcal{F}$ is **Hilbert** if its topology is a Hilbert space topology. Otherwise, we say that $\mathcal{F}$ is **locally Hilbert** if for each $K \in \mathcal{K}_{\mathbb{R}^n}$ the topology of $\mathcal{F}_K$ is Hilbert.

(4) (**Diffeomorphism Invariance Axiom**) If $\chi \in \mathrm{Diff}(\mathbb{R}^n)$, we require that $\chi_* : \mathcal{D}' \to \mathcal{D}'$, $\chi_*(u)\varphi := u(\varphi \circ \chi^{-1})$, restrict to a topological isomorphism on $\mathcal{F}$. We say that $\mathcal{F}$ is **locally**

invariant if for all $\chi \in \mathrm{Diff}(\mathbb{R}^n)$ the map χ_* restricts to a topological isomorphism on $\mathcal{F}_K$ for all $K \in \mathcal{K}_{\mathbb{R}^n}$,

$$\chi_* : \mathcal{F}_K \xrightarrow{\sim} \mathcal{F}_{\chi(K)}$$

(5) (**Density Axiom**) We that $\mathcal{F}$ is **normal** if $\mathcal{D}$ is dense in $\mathcal{F}$, and we say that $\mathcal{F}$ is **locally normal** if $\mathcal{F}_K \subseteq \overline{\mathcal{D}} \subseteq \mathcal{F}$ for all $K \in \mathcal{K}_{\mathbb{R}^n}$, where the closure is in $\mathcal{F}$. This axiom becomes significant when we consider dualizing. Indeed, it follows that the duality between $\mathcal{F}_{\mathrm{loc}}$ and $\mathcal{F}_{\mathcal{K}}$ becomes

$$(\mathcal{F}_{\mathrm{loc}})^* = (\mathcal{F}^*)_{\mathcal{K}} \quad \text{and} \quad (\mathcal{F}_{\mathcal{K}})^* = (\mathcal{F}^*)_{\mathrm{loc}}$$

(6) (**Locality Axiom**) We say that $\mathcal{F}$ is **local** if, as a locally convex TVS we have $\mathcal{F} = \mathcal{F}_{\mathrm{loc}}$.

Remark 48 If $\mathcal{F}$ is locally Banach or locally Fréchet, the $\mathcal{F}_{\mathcal{K}}$ is a complete locally convex TVS which is not Fréchet. This follows from the properties of the inductive limit topology, which require it to restrict to the topologies on each $\mathcal{F}_K$. ■

Example 20 $\mathcal{E}$ is Fréchet, but not Banach or even locally Banach. $\mathcal{D}$ is not Fréchet, but it is locally Fréchet. The Sobolev spaces H^s are Hilbert, and so Banach, but their local versions H^s_{loc} are just Fréchet and locally Hilbert. $\mathcal{S}$ is Fréchet but not Banach. All the standard spaces $\mathcal{D}$, $\mathcal{E}$, $\mathcal{S}$, $\mathcal{D}'$, $\mathcal{E}'$, and $\mathcal{S}'$ are diffeomorphism invariant, but the Sobolev spaces H^s are not, though they are locally diffeomorphism invariant (a non-trivial result which requires pseud-differential operators to prove, Theorem 9.2.3, van den Ban and Crainic [104]). The spaces $\mathcal{D}$, $\mathcal{E}$, $\mathcal{S}$, $\mathcal{D}'$, $\mathcal{E}'$, and $\mathcal{S}'$ are normal, as is H^s, though $H^s(\Omega)$ not so for arbitrary open sets $\Omega \subseteq \mathbb{R}^n$, though these are locally normal. $\mathcal{E}$ and $\mathcal{D}'$ are local, but $\mathcal{D}$ and $\mathcal{E}'$ are not. Nor is H^s local, which is why we introduced H^s_{loc}.

The usefulness of these notions is nicely captured in the following Theorem, see van den Ban and Crainic [104, Theorem 3.5.4] for a proof.

Theorem 29 *If $\mathcal{F} = \mathcal{F}_{\mathrm{loc}}$, then $\mathcal{F}$ is locally Fréchet iff it is Fréchet, locally invariant iff it is invariant, and locally normal iff it is normal. On the other hand, if $\mathcal{F}$ is locally Fréchet, then it is local iff the following two test conditions are satisfied:*

(1) Let $u \in \mathcal{D}'$. Then $u \in \mathcal{F}$ iff $\varphi u \in \mathcal{F}$ for all $\varphi \in \mathcal{D}$.

(2) $u_n \to u$ in $\mathcal{F}$ iff $\varphi u_n \to \varphi u$ for all $\varphi \in \mathcal{D}$.

■

Now, in order that $\mathcal{F}$ be transferrable to a manifold, we need compatibility with restricitons to open sets and with local diffeomorphisms. The localization axiom implies that we can restrict distributions to opens, for the restriction map

$$r_{\Omega,\mathbb{R}^n} : C^\infty(\mathbb{R}^n) \to C^\infty(\Omega)$$

dualizes to the inclusion

$$i_{\Omega,\mathbb{R}^n} : \mathcal{E}'(\Omega) \hookrightarrow \mathcal{E}'(\mathbb{R}^n) \hookrightarrow \mathcal{D}'(\mathbb{R}^n)$$

which may be thought of as 'extension by zero'. Given a *local* function space $\mathcal{F} = \mathcal{F}_{\mathrm{loc}}$, therefore, we define, for each open $\Omega \subseteq \mathbb{R}^n$ the space

$$\mathcal{F}(\Omega) := \{u \in \mathcal{D}'(\Omega) \mid \varphi u \in \mathcal{F},\ \forall \varphi \in \mathcal{D}(\Omega)\} \tag{3.62}$$

endowed with the topology generated by the family of seminorms $\Gamma_\Omega = \{q_{p,\varphi} \mid p \in \Gamma_{\mathcal{F}},\ \varphi \in \mathcal{D}(\Omega)\}$. By Theorem 3.6.2, van den Ban and Crainic [104], we have that $\mathcal{F}(\Omega)$ is invariant, normal, or locally Banach/Hilbert/Fréchet if $\mathcal{F}$ is. Moreover, since $\mathcal{F}$ is local, so is $\mathcal{F}(\Omega)$, and for such a local function space the assignment

$$\Omega \mapsto \mathcal{F}(\Omega) \tag{3.63}$$

defines a sheaf, called the **sheaf of distributions**, which has the diffeomorphism invariance property

$$\chi_* : \mathcal{F}(\Omega_1) \xrightarrow{\sim} \mathcal{F}(\Omega_2)$$

for all $\chi \in \mathrm{Diff}(\Omega_1, \Omega_2)$.

Definition 40 Fix, now, a local invariant function space $\mathcal{F}$, and let $E \to M$ be a rank p smooth complex vector bundle over a smooth n-manifold M, and consider a trivializing atlas $\{(U_i, \varphi_i)\}_{i\in I}$ for E, with local trivializations $\tau_i : E|_{U_i} \to U_i \times \mathbb{C}^p$. From these trivializations we get isomorphisms

$$h_{\varphi_i,\tau_i} : \mathcal{D}'(U_i, E|_{U_i}) \xrightarrow{\sim} \mathcal{D}'(\varphi(U))^p$$

and similarly with $\mathcal{D}$, $\mathcal{E}$, $\mathcal{E}'$ and the others. Using these, we define a **function space of sections of a vector bundle** by

$$\mathcal{F}(E) := \{u \in \mathcal{D}'(E) \mid h_{\varphi,U}(u|_U) \in \mathcal{F}(\varphi(U)),\ \forall U \in \mathcal{U}\} \tag{3.64}$$

where $\mathcal{U} = \{U_i\}_{i\in I}$ is the open cover by the chart domains. Note that by locality $u \in \mathcal{F}(E)$ iff $u|_{U_i} \in \mathcal{F}(U_i, E|_{U_i})$ for all $i \in I$. By Theorem 3.7.4, van den Ban and Crainic [104], we have that $\mathcal{F}(E) = \mathcal{F}_{\mathrm{loc}}(E)$ is local, and is locally Banach/Fréchet/Hilbert if $\mathcal{F}$ is, and normal if $\mathcal{F}$ is. Moreover, any isomorphism $h : E \to F$ of vector bundles over M induces an isomorphism $h_*\mathcal{F}(E) \to \mathcal{F}(F)$ of function spaces. As a consequence, if $\mathcal{F}$ is locally Banach or Hilber, and M is compact, then $\mathcal{F}(E)$ is Banach or Hilbert, respectively, and $\mathcal{D}(E)$ is dense in $\mathcal{F}(E)$.

Remark 49 Consider now $H^s_{\mathrm{loc}} := H^s_{\mathrm{loc}}(\mathbb{R}^n) = \{u \in \mathcal{D}' \mid \varphi u \in H^s(\mathbb{R}^n),\ \forall \varphi \in \mathcal{D}\}$. It is locally Hilbert, invariant (as mentioned above), and normal, with

$$\mathcal{E} := \mathcal{E}(\mathbb{R}^n) = \bigcap_{s\in\mathbb{R}} H^s_{\mathrm{loc}}(\mathbb{R}^n)$$

Thus, $H^s_{\mathrm{loc}}(E)$ is Fréchet, locally Hilbert (and Hilbert if M is compact), $\mathcal{D}(E)$ is dense in $H^s_{\mathrm{loc}}(E)$, and

$$\bigcap_{s\in\mathbb{R}} H^s_{\mathrm{loc}}(E) = \mathcal{E}(E) \tag{3.65}$$

since any $u \in \mathcal{D}'(E)$ such that $u \in H^s_{\mathrm{loc}}(E)$ for all $s \geq 0$ lies in $\mathcal{E}$. Moreover, the Rellich-Kondrachov theorem extends to the compact setting: if $r < s$ and M is compact, then the inclusion

$$H^s_{\mathrm{loc}}(E) \hookrightarrow H^r_{\mathrm{loc}}(E) \tag{3.66}$$

is compact, which follows from a partition of unity argument. ■

3.1.3 Differential Operators on Manifolds

A linear differential operator in Euclidean space with variable coefficients is a polynomial in $D^\alpha = \frac{\partial^{|\alpha|}}{\partial x^\alpha}$ with smooth matrix-valued coefficients,

$$P(x,D):\ C^\infty(\mathbb{R}^n,\mathbb{R}^m)\to C^\infty(\mathbb{R}^n,\mathbb{R}^r)$$
$$P(x,D)f=\sum_{|\alpha|\le k}A^\alpha(x)D^\alpha f$$

where

$$D^\alpha\begin{pmatrix} f_1\\ \vdots\\ f_m\end{pmatrix}:=\begin{pmatrix} D^\alpha f_1\\ \vdots\\ D^\alpha f_m\end{pmatrix}$$

and the coefficient matrices

$$A^\alpha:\mathbb{R}^n\to M_{r,m}(\mathbb{R})\cong \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^m,\mathbb{R}^r)$$
$$A^\alpha(x)=\begin{pmatrix} a^\alpha_{11} & \cdots & a^\alpha_{1m}\\ \vdots & \ddots & \vdots\\ a^\alpha_{r1} & \cdots & a^\alpha_{rm}\end{pmatrix}$$

act on each $D^\alpha f$ by multiplication.

Remark 50 Of course, we could generalize this definition to complex-valued functions,

$$P(x,D):\ C^\infty(\mathbb{R}^n,\mathbb{C}^m)\to C^\infty(\mathbb{R}^n,\mathbb{C}^r)$$
$$P(x,D)f=\sum_{|\alpha|\le k}A^\alpha(x)D^\alpha f$$

where the D^α act on the real and imaginary parts of $f=f_1+if_2$,

$$D^\alpha f=D^\alpha f_1+iD^\alpha f_2$$ ■

The following well-known examples illustrate this definition, though in some sense they are simpler, because the coefficient matrices are constant.

Example 21 The **gradient** is a differential operator on $\mathbb{R}^n$,

$$\text{grad} : C^\infty(\mathbb{R}^n, \mathbb{R}) \to C^\infty(\mathbb{R}^n, \mathbb{R}^n)$$

which is given by

$$\text{grad}\, f := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \frac{\partial}{\partial x_1} f + \cdots + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \frac{\partial}{\partial x_n} f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

It is a first order operator. ■

Example 22 The **divergence** is a differential operator on $\mathbb{R}^n$,

$$\text{div} : C^\infty(\mathbb{R}^n, \mathbb{R}^n) \to C^\infty(\mathbb{R}^n, \mathbb{R})$$

given by

$$\begin{aligned} \text{div}\, f &= \text{div} \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} \\ &:= \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} \frac{\partial}{\partial x_1} \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} + \cdots + \begin{pmatrix} 0 & \cdots & 0 & 1 \end{pmatrix} \frac{\partial}{\partial x_n} \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} \\ &= \frac{\partial f_1}{\partial x_1} + \cdots + \frac{\partial f_n}{\partial x_n} \end{aligned}$$

It is a first order operator. ■

Example 23 The **Laplacian on real-valued functions** is a second order differential operator

$$\Delta : C^\infty(\mathbb{R}^n, \mathbb{R}) \to C^\infty(\mathbb{R}^n, \mathbb{R})$$

given by

$$\begin{aligned} \Delta f &:= (1) \frac{\partial^2}{\partial x_1^2} f + \cdots + (1) \frac{\partial^2}{\partial x_n^2} f \\ &= \frac{\partial^2 f}{\partial x_1^2} + \cdots + \frac{\partial^2 f}{\partial x_n^2} \end{aligned}$$

It is a second order operator. The coefficient matrices $A^\alpha = (1)$ are 1×1 identity matrices. Note that the Laplacian is the composition of the divergence and gradient operators,

$$\begin{tikzcd} C^\infty(\mathbb{R}^n, \mathbb{R}) \arrow[rr, "\Delta"] \arrow[dr, "\text{grad}"'] & & C^\infty(\mathbb{R}^n, \mathbb{R}) \\ & C^\infty(\mathbb{R}^n, \mathbb{R}^n) \arrow[ur, "\text{div}"'] & \end{tikzcd}$$

i.e. $\Delta f = \operatorname{div}(\operatorname{grad} f)$. ■

Example 24 The **Laplacian on vector-valued functions** is a second order differential operator,

$$\Delta^m := I_m \Delta : C^\infty(\mathbb{R}^n, \mathbb{R}^m) \to C^\infty(\mathbb{R}^n, \mathbb{R}^m)$$

where I_m is the identity matrix in $M_m(\mathbb{R})$, and so Δ^m is given by

$$\begin{aligned} \Delta^m f &= \Delta^m \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} \\ &:= I_m \frac{\partial^2}{\partial x_1^2} \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} + \cdots + I_m \frac{\partial^2}{\partial x_n^2} \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial^2 f_1}{\partial x_1^2} + \cdots + \frac{\partial^2 f_1}{\partial x_n^2} & 0 & \cdots & 0 \\ 0 & \frac{\partial^2 f_2}{\partial x_1^2} + \cdots + \frac{\partial^2 f_2}{\partial x_n^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\partial^2 f_m}{\partial x_1^2} + \cdots + \frac{\partial^2 f_m}{\partial x_n^2} \end{pmatrix} \\ &= \begin{pmatrix} \Delta f_1 & \cdots & 0 \\ \vdots & \ddots & 0 \\ 0 & \cdots & \Delta f_m \end{pmatrix} \end{aligned}$$

We will usually omit the superscript m and simply write Δ. ■

Let us try to reformulate the definition of $P(x, D)$ in terms of manifolds and vector bundles. First, we may view $\mathbb{R}^n$ as a manifold equipped with the trivial vector bundle $E = \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$

and similarly with $F = \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n$ (or $E = \mathbb{R}^n \times \mathbb{C}^m \to \mathbb{R}^n$ and $E = \mathbb{R}^n \times \mathbb{C}^r \to \mathbb{R}^n$ in the complex case). The space of smooth functions $C^\infty(\mathbb{R}^n, \mathbb{R}^m)$ is then identically the space of sections of E, and similarly with $C^\infty(\mathbb{R}^n, \mathbb{R}^r)$,

$$\Gamma(F) = C^\infty(\mathbb{R}^n, \mathbb{R}^m),$$

$$\Gamma(F) = C^\infty(\mathbb{R}^n, \mathbb{R}^r)$$

Then, a linear differential operator $P(x, D)$ is a map of sections,

$$P(x, D) : \Gamma(E) \to \Gamma(F)$$

where $x \in \mathbb{R}^n$ is an element of the base manifold and the derivatives D^α act on sections of the vector bundle E componentwise, the differentiation being with respect to the base coordinates x. The coefficient matrices $A^\alpha : \mathbb{R}^n \to M_{r,m}(\mathbb{R}) \cong \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^m, \mathbb{R}^r)$ act on the result of D^α by multiplication. Thus, in full, if $f \in \Gamma(E)$,

$$P(x, D)f = \sum_{|\alpha| \leq k} A^\alpha(x) D^\alpha f = \sum_{|\alpha| \leq k} \begin{pmatrix} a_{11}^\alpha & \cdots & a_{1m}^\alpha \\ \vdots & \ddots & \vdots \\ a_{r1}^\alpha & \cdots & a_{rm}^\alpha \end{pmatrix} \begin{pmatrix} \frac{\partial^{|\alpha|} f_1}{\partial x^\alpha} \\ \vdots \\ \frac{\partial^{|\alpha|} f_m}{\partial x^\alpha} \end{pmatrix}$$

Abstracting the above construction to real or complex vector bundles $E \to M$ and $F \to M$ involves two changes: (1) The base manifold $\mathbb{R}^n$ is changed to a smooth manifold M which is only *locally* like $\mathbb{R}^n$, though it may possess different local and global topological properties, and (2) The vector bundles E and F are no longer globally trivial, but only locally trivial. Thus, the generalization to manifolds involves adding additional topological structure to M, which manifests itself also in the twisting of the vector bundles attached to M. As it turns out, however, this extra structure doesn't affect the definition of a differential operator, though, as we will see, certain differential operators (the elliptic self-adjoint operators particularly) are affected by the topology of M. This is the content of the index theorems.

Definition 41 (Local Definition) The simplest generalization of a linear differential operator to manifolds is the local coordinate definition, and may be found in Hörmander's book [57, p. 151].

Let E and F be real or complex smooth vector bundles over a smooth manifold M, of rank m and r, respectively. A **linear differential operator** is an $\mathbb{R}$- or $\mathbb{C}$-linear operator (depending on whether the vector bundles are real or complex) map of sections

$$P : \Gamma(E) \to \Gamma(F)$$

which is *locally* a differential operator from $C^\infty(\Omega, \mathbb{R}^m)$ to $C^\infty(\Omega, \mathbb{R}^k)$, where $\Omega \subseteq \mathbb{R}^n$ is an open subset. That is, in any trivializing charts (U, φ) for E and (V, ψ) for F about any point $s(p) \in E_p$ and $Ps(p) \in F_{s(p)}$ we have that $\psi \circ P \circ \varphi^{-1}$ is of the form $\sum_{|\alpha| \leq k} A^\alpha(x) D^\alpha$. ■

Definition 42 (Sheaf Morphism Definition) Closely connected to the local definition is the sheaf-theoretic definition. Let E and F be real or complex smooth vector bundles over a smooth manifold M, of rank m and r, respectively, and define a **linear differential operator** to be an $\mathbb{R}$- or $\mathbb{C}$-linear map of sections

$$P : \Gamma(E) \to \Gamma(F)$$

which is **support decreasing**,

$$\operatorname{supp}(Ps) \subseteq \operatorname{supp}(s)$$

for all compactly supported sections $s \in \Gamma_c(E)$. If we consider the restriction map $r_{V,U} : U \to V$, where $V \subseteq U$ are open sets of M, then we can rephrase this condition as commutation with restriction $r_{V,U}$: a linear differential operator is an $\mathbb{R}$- or $\mathbb{C}$-linear map of sections commutig with restriction,

$$P \circ r_{V,U} = r_{V,U} \circ P$$

i.e. $(P|_V)s = (Ps)|_V$ for all local sections $s \in \Gamma(U, E)$. ■

Remark 51 The equivalence of these two definitions is known as Peetre's Theorem [88]. The order of the operator P is the integer k in its local description, $P(x, D)s = \sum_{|\alpha| \leq k} A^\alpha(x) D^\alpha s$. It must be checked that this order is invariant under chart transitions, though this is intuitively clear, since chart transitions are elements of $\mathrm{GL}(n, \mathbb{R})$ or $\mathrm{GL}(n, \mathbb{C})$. ■

A more invariant, global definition is the following, found in, for example, Guillemin and Sternberg [48]. Its form is similar to that of the universal enveloping algebra of a Lie algebra, and likely has that as a model for its construction.

Definition 43 We define the *space* of linear differential operators simultaneously with the linear differential operators themselves, and we do this inductively over k, the order of the operator. This has the advantage of telling us how to view differential operators globally and invariantly, and also provides a mechanism for defining and calculating the symbol of an operator. Towards this end, let E and F be real or complex smooth vector bundles over a smooth manifold M. A **zeroth order linear differential operator** is defined to be a bundle map, $P : E \to F$,

$$\begin{tikzcd} E \arrow[rr, "P"] \arrow[dr, "\pi_E"'] & & F \arrow[dl, "\pi_E"] \\ & M & \end{tikzcd}$$

Now, a bundle morphism is by definition a $C^\infty(M)$-linear smooth map from E to F which is fiberwise $\mathbb{R}$- or $\mathbb{C}$-linear, as the case may be. In other words, P is a section of the bundle $\mathrm{Hom}(E, F) \to M$, or, what is equivalent since they are isomorphic, a $C^\infty(M)$-linear homomorphism from the sections $\Gamma(E)$ to the sections $\Gamma(F)$,

$$P \in \Gamma\big(\mathrm{Hom}(E, F)\big) \cong \mathrm{Hom}_{C^\infty(M)}(\Gamma(E), \Gamma(F)) \tag{3.67}$$

(recall that $\Gamma(\mathrm{Hom}(E, F))$ is a $C^\infty(M)$-module). The $C^\infty(M)$-linearity can be expressed in terms of commutators: If $s \in \Gamma(E)$ and $f \in C^\infty(M)$, then

$$P(fs) = fP(s) \quad \Longleftrightarrow \quad [P, f](s) := (P \circ f - f \circ P)(s) = 0 \tag{3.68}$$

Here f means *multiplication by* f. Then define the **space of zeroth order linear differential operators** to be the space of bundle maps,

$$\mathcal{D}^0(E, F) := \mathrm{Hom}_{C^\infty(M)}(\Gamma(E), \Gamma(F)) = \{P \in \mathrm{Hom}(\Gamma(E), \Gamma(F)) \,|\, [P, f] = 0,\ \forall f \in C^\infty(M)\} \tag{3.69}$$

Analogously, we define the **space of differential operators of order at most one** to be

$$\mathcal{D}^1(E, F) := \{\mathrm{Hom}(\Gamma(E), \Gamma(F)) \,|\, [P, f] \in \mathcal{D}^0(E, F),\ \forall f \in C^\infty(M)\} \tag{3.70}$$

and, continuing inductively, we define the **space of linear differential operators of order at most** $k+1$ by

$$\mathcal{D}^{k+1}(E,F) := \{\mathrm{Hom}(\Gamma(E),\Gamma(F)) \mid [P,f] \in \mathcal{D}^k(E,F),\ \forall f \in C^\infty(M)\} \tag{3.71}$$

Then we define the **space of all linear differential operators** from E to F to be the union of these spaces,

$$\mathcal{D}(E,F) := \bigcup_{k=0}^{\infty} \mathcal{D}^k(E,F) \tag{3.72}$$

■

We follow the discussion in Nicolaescu [85] for the rest of this section.

Remark 52 The appearance of the commutator here is analogous to the situation of the adjoint representation of a Lie algebra, $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$, which satisfies $\mathrm{ad}_u(v) = [u,v]$. Let us use this notation in this context, as

$$\begin{aligned}&\mathrm{ad} : C^\infty(M) \to \mathrm{End}_{\mathbb{C}}\big(\mathrm{Hom}_{\mathbb{C}}(\Gamma(E),\Gamma(F))\big)\\ &\mathrm{ad}_f = [\cdot, f]\end{aligned}$$

where we insert a map $P \in \mathrm{Hom}_{\mathbb{C}}(\Gamma(E),\Gamma(F))$ into the open slot

$$\mathrm{ad}_f(P) = [P,f]$$

Here $\mathbb{C}$ may be replaced with $\mathbb{R}$, if E and F are real vector bundles. Then we can see that the zeroth order linear differential operators are the kernel of this adjoint map,

$$\mathcal{D}^0(E,F) = \ker \mathrm{ad} \tag{3.73}$$

Similarly, if $\mathrm{ad}_f(P) = [P,f] \in \mathcal{D}^0(E,F) = \ker \mathrm{ad}$ for all $f \in C^\infty(M)$, then $P \in \mathcal{D}^1(E,F)$. But this is equivalent to saying

$$P \in \ker \mathrm{ad}^2$$

where $\mathrm{ad}^2 : C^\infty(M)^2 \to \mathrm{End}_{\mathbb{C}}\big(\mathrm{Hom}_{\mathbb{C}}(\Gamma(E), \Gamma(F))\big)$ is given by

$$\mathrm{ad}^2(f,g)(P) := \mathrm{ad}_f \circ \mathrm{ad}_g(P) = \mathrm{ad}_f[P,g] = [[P,g],f]$$

Thus,

$$\mathcal{D}^1(E,F) = \ker \mathrm{ad}^2 \tag{3.74}$$

More generally, if we define, for all $k \in \mathbb{N}$,

$$\mathrm{ad}^k : C^\infty(M)^k \to \mathrm{End}_{\mathbb{C}}\big(\mathrm{Hom}_{\mathbb{C}}(\Gamma(E), \Gamma(F))\big) \tag{3.75}$$

$$\mathrm{ad}^k(f_1, \ldots, f_k)(P) := \mathrm{ad}_{f_1} \circ \cdots \circ \mathrm{ad}_{f_k}(P) = [\cdots [[P, f_k], f_{k-1}] \cdots, f_1] \tag{3.76}$$

Then we have

$$\mathcal{D}^k(E,F) := \ker \mathrm{ad}^{k+1} \tag{3.77}$$

for all $k \in \mathbb{N}$, and we no longer require recursion to define $\mathcal{D}^k(E,F)$. ■

By the following proposition, the local and global definitions 42 and 43 are equivalent. We follow the proof in Nicolaescu [85, Lemma 10.1.3].

Proposition 61 *Any $P \in \mathcal{D}(E,F)$ is support decreasing,* $\mathrm{supp}(Ps) \subseteq \mathrm{supp}(s)$ *for all $s \in \Gamma(E)$, and conversely.*

Proof: The proof is by induction over k, the order of P. If $k = 0$, this is clear since P is a bundle map. Now suppose that $P \in \mathcal{D}^{k+1}(E,F)$ and that the result holds for all $L \in \mathcal{D}^r(E,F)$, $0 \leq r \leq k$. Then, for all $f \in C^\infty(M)$ and all $s \in \Gamma_c(U,E)$

$$P(fs) = [P,f](s) + fP(s)$$

so since $\mathrm{supp}(P(fs)) \subseteq \mathrm{supp}(f) \cup \mathrm{supp}(s)$, by choosing f to be a bump function with support containing $\mathrm{supp}(s)$ and zero outside any open set V containing $\mathrm{supp}(s)$, with $fs \equiv s$, the result follows from the induction hypothesis, since $[P,f] \in \mathcal{D}^k(E,F)$. Conversely, if $\mathrm{supp}(Ps) \subseteq \mathrm{supp}(s)$, the by Peetre's theorem we know that P is a local linear differential operator of some order k. Any such operator satisfies $P \in \ker \mathrm{ad}^{k+1}$ locally, and hence, by use of a partition of unity subordinate to an open cover by trivializing charts, globally on M. ■

Proposition 62 ad^k *is symmetric in the entries* $f_j \in C^\infty(M)$,

$$\mathrm{ad}^k(f_{\sigma(1)}, \ldots, f_{\sigma(k)})(P) = \mathrm{ad}^k(f_1, \ldots, f_k)(P)$$

for all permutations $\sigma \in S_k$ *on* k *letters.*

Proof: Since any permutation $\sigma \in S_k$ decomposes into transpositions, it is enough to prove it for transpositions $\tau(i\,j)$. First, note that on $C^\infty(M)$, viewed as embedded in $\mathcal{D}^0(E, F)$ as multiplication operators, we have $[f, g] = 0$, so by the Jacobi identity

$$[[L, f], g] = [L, [f, g]] + [[L, g], f] = [[L, g], f]$$

whence, for all $f, g, h \in C^\infty(M)$,

$$[[[L, f], g], h] = [[[L, g], f], h] = [[[L, g], h], f] = [[[L, h], g], f] = [[[L, h], g], f] = [[[L, h], f], g]$$

and so on. The general result follows by induction and the application of transpositions. ■

Proposition 63 *Let* E*,* F*, and* G *be complex (or real) smooth vector bundles over* M*. If* $P \in \mathcal{D}^m(F, G)$ *and* $Q \in \mathcal{D}^n(E, F)$*, then* $P \circ Q \in \mathcal{D}^{m+n}(E, G)$*.*

Proof: The proof is by induction over $m + n$. For $m + n = 0$ we have $m = n = 0$, so the result follows from the fact that bundle maps compose to form bundle maps. Suppose the result true for some $m + n \geq 0$, then, and consider the case $m + n + 1$. For all $f \in C^\infty(M)$ we have

$$\begin{aligned}
[P \circ Q, f] &= P \circ Q \circ f - f \circ P \circ Q + P \circ f \circ Q - P \circ f \circ Q \\
&= (P \circ f - f \circ P) \circ Q + P \circ (Q \circ f - f \circ Q) \\
&= [P, f] \circ Q + P \circ [Q, f]
\end{aligned}$$

But the operators $[P, f] \circ Q$ and $P \circ [Q, f]$ have orders $\leq m + n$, so by the induction hypothesis $[P \circ Q, f] \in \mathcal{D}^{m+n}(E, F)$. But since this is true for all f, we must have $P \circ Q \in \mathcal{D}^{m+n+1}(E, F)$, and by induction we conclude that the statement holds for all $m, n \in \mathbb{N}_0$. ■

The discussion so far has centered on linear differential operators defined as maps of *smooth* sections. However, we would like to think of differential operators as densely defined operators on a Hilbert space generally, and we really need this generality in order to be able to use the techniques of functional analysis and PDE theory. Indeed, we need the full framework of distributions and Sobolev spaces to be able to do this. Let us describe this generalization now.

Definition 44 Let E and F be complex (or real) smooth vector bundles over a smooth manifold M, and let $\mathcal{D}(E)$ and $\mathcal{D}(F)$ be the spaces of smooth sections $\Gamma(E)$ and $\Gamma(F)$, respectively, considered as locally convex spaces endowed with the inductive limiti topology induced from the locally convex spaces $\mathcal{D}_K(E)$ and $\mathcal{D}_K(F)$, as explained in the previous section. A **generalized linear differential operator** is then a continuous $\mathbb{C}$-linear map

$$P : \mathcal{D}(E) \to \mathcal{D}'(F) \tag{3.78}$$

Letting $\mathcal{F}_1, \mathcal{F}_2 \in \{\mathcal{D}, \mathcal{E}, \mathcal{D}', \mathcal{E}', H^s_{\text{loc}}\}$ be among the local invariant function spaces on $\mathbb{R}^n$, which we transfer to the bundles E and F, the inclusion of $\mathcal{D}$ in each $\mathcal{F}_i$ implies that P extends to a $\mathbb{C}$-linear operator

$$P : \mathcal{F}_1(E) \to \mathcal{F}_2(F) \tag{3.79}$$

and conversely this operator induces the former by the denseness of $\mathcal{D}$ in the other spaces. By Proposition 3.8.4, van den Ban and Crainic [104]) any operator $P \in \mathcal{D}^k(E, F)$, thought of as $P : \mathcal{D}(E) \to \mathcal{D}(F)$, extends uniquely to an operator

$$P : H^s_{\text{loc}}(E) \to H^{s-k}_{\text{loc}}(F)$$

for any $s \geq k \geq 0$. But by viewing $\mathcal{D}(F)$ as contained in $\mathcal{D}'(F)$, we can extend P to any local invariant function space, in fact. The idea here is that a generalized operator P takes *smooth* sections and gives generalized sections, possibly non-smooth. Our main interest will by in Sobolev spaces and L^2, for our symplectic Dirac operators.

A **smoothing operator**, by contrast, should take generalized sections to smooth sections, and so we define it as a continuous $\mathbb{C}$-linear map

$$P : \mathcal{E}'(E) \to \mathcal{E}(F) \tag{3.80}$$

A smoothing operator, then, is generalized operator taking values in $\mathcal{E}(F)$ which extends to a continuous/bounded operator on $\mathcal{E}'(E)$. The space of all smoothing operators is denoted

$$\begin{aligned}
\Psi^{-\infty}(E,F) &:= \{P : \mathcal{D}(E) \to \mathcal{D}'(F) \mid P \text{ extends to } \mathcal{E}'(E) \to \mathcal{E}(F)\} && (3.81)\\
&= \mathcal{B}\big(\mathcal{E}'(E), \mathcal{E}(F)\big) && (3.82)\\
&\subseteq \mathcal{B}\big(\mathcal{D}(E), \mathcal{D}'(F)\big) && (3.83)
\end{aligned}$$

■

Remark 53 Recalling the Schwartz Kernel Theorem 25, we can begin to see the significance of these definitions, and their relationship. In the manifold setting, this takes the form of an isomorphism of locally convex topological vector spaces. Let $E \to M$ and $F \to N$ be complex (or real) smooth vector bundles, and consider the bundle : $E^\vee = E^* \otimes \Omega^n_c(M) \to M$. If $\pi_1 : N \times M \to N$ and $\pi_2 : N \times M \to M$ are the projections onto each component, define

$$F \boxtimes E^\vee := \pi_1^*(F) \otimes \pi_2^*(E^\vee)$$

Then, the Schwartz Kernel Theorem then takes the form

$$\mathcal{D}'(N \times M, F \boxtimes E^\vee) \cong \mathcal{B}\big(\mathcal{D}(E), \mathcal{D}'(F)\big) \tag{3.84}$$

as in the Euclidean case (Theorem 2.4.5, van den Ban and Crainic [104]). The subspace $\Psi^{-\infty}(E,F)$ of $\mathcal{B}\big(\mathcal{D}(E), \mathcal{D}'(F)\big)$ therefore has a corresponding subspace in $\mathcal{D}'(N \times M, F \boxtimes E^\vee)$ to which it is isomorphic, called the **space of generalized sections of** $F \boxtimes E^\vee$, namely $\mathcal{E}(N \times M, F \boxtimes E^\vee)$,

$$\mathcal{E}(N \times M, F \boxtimes E^\vee) \cong \Psi^{-\infty}(E,F) \tag{3.85}$$

■

3.1.3.1 The Symbol of a Differential Operator

We now describe the principal symbol of a differential operator, which gives, roughly, a coarse estimate of the behavior of the operator in by isolating its highest order terms. To get an invariant, non-local analog of this quantity, we observe that in the Euclidean setting, the highest order terms $\sum_{|\alpha|=k} A^\alpha(x)D^\alpha$ of an operator $P(x,D) = \sum_{|\alpha|\leq k} A^\alpha(x)D^\alpha$ are captured by the adjoint operator ad given above,

$$\frac{1}{k!}\,\mathrm{ad}^k(f,\dots,f)(P) = \sum_{|\alpha|=k} A^\alpha(x)D^\alpha f$$

As we can see, this quantity depends only on the value of $df(x_0) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}\,dx^j$, and therefore describes the fiberwise behavior of P, encoded in terms of an element of $(\mathbb{R}^n)^*$. This gives us an invariant, coordinate-free way to define the symbol of any operator $P \in \mathcal{D}^k(E,F)$. When this quantity is a linear isomorphism of the fibers $E^*_{x_0}$, the operator is called elliptic. Elliptic operators are such nicely behaved operators that they occupy a special place in PDE and operator theory. For example, an elliptic operator is 'regular', in the sense that, even when it satisfies an equality $Pu = v$ weakly, if we know that v has a certain amount of smoothness, then we can conclude that so, too, does u. A *symmetric* elliptic operator also has very good spectral properties, with discrete spectrum consisting only of real eigenvalues, while elliptic regularity guarantees that the corresponding eigenspaces, which are finite-dimensional, consist of *smooth* functions. We study elliptic operators more closely in the following section. In this section we follow the treatement in Nicolaescu [85].

Our first task is to show that the symbol, whose definition we have only sketched, only depends on the value of df at a point p_0, so that if we take two functions f and g whose differentials agree at p_0, then the symbol will agree if evaluated at f and g.

Lemma 10 *Let $p_0 \in M$ and define the ideals*

$$\mathfrak{m}_{p_0} := \{f \in C^\infty(M) \mid f(p_0) = 0\}$$

$$\mathfrak{I}_{p_0} := \{f \in C^\infty(M) \mid f(p_0) = df(p_0) = 0\}$$

in $C^\infty(M)$. *Then,*

$$\mathfrak{I}_{p_0} = \mathfrak{m}_{p_0}^2$$

which means that each $f \in \mathfrak{I}_{p_0}$ *can be written as a finite sum* $f = \sum_{i=1}^r f_i g_i$ *for* $f_i, g_i \in \mathfrak{m}_{p_0}$.

Proof: Since $df = \sum_{j=1}^n \frac{\partial f}{\partial x_j}\, dx^j$, we see that $df(p_0) = 0$ iff $\frac{\partial f}{\partial x_j}|_{p_0} = 0$ for all i, in which case we can write $f(p) = \int_\gamma df$ for any path $\gamma : I \to M$ from p_0 to p. In particular,

$$\begin{aligned} f(p) = \sum_{i=1}^n \int_\gamma \frac{\partial f}{\partial x_j}\, dx^i = \sum_{i=1}^n \int_0^1 \frac{\partial f}{\partial x_j}(\gamma(t))\gamma'(t)\, dt \\ = \sum_{i=1}^n \int_0^1 \frac{\partial (f \circ \varphi^{-1})}{\partial x_j}(\varphi \circ \gamma(t)) \frac{d}{dt}(\varphi \circ \gamma)(t)\, dt \end{aligned}$$

for any chart (U, φ) about $\gamma(t)$. Now, if we choose γ carefully, for example so that $(\varphi \circ \gamma)(t) = t(\varphi(p) - \varphi(p_0)) + \varphi(p_0)$, i.e.

$$\gamma(t) = \varphi^{-1}\big(t(\varphi(p) - \varphi(p_0)) + \varphi(p_0)\big)$$

then

$$\gamma'(t) := \frac{d}{dt}(\varphi \circ \gamma)(t) = \varphi(p) - \varphi(p_0) = \big(x^1(p) - x^1(p_0), \ldots, x^n(p) - x^n(p_0)\big)$$

and therefore

$$\begin{aligned} f(p) &= f(p) - f(p_0) \\ &= f(\gamma(1)) - f(\gamma(0)) \\ &= \int_0^1 (f \circ \gamma)'(t)\, dt \\ &= \int_0^1 df(\gamma(t))\, \gamma'(t)\, dt \\ &= \int_0^1 \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\gamma(t)) dx^i \circ \big(x^1(p) - x^1(p_0), \ldots, x^n(p) - x^n(p_0)\big)\big)\, dt \\ &= \sum_{i=1}^n (x^i(p) - x^i(p_0)) \int_0^1 \frac{\partial f}{\partial x_i}(\gamma(t))\, dt \\ &= \sum_{i=1} f_i(p) g_i(p) \end{aligned}$$

where $f_i(p) = x^i(p) - x^i(p_0)$ and $g_i(p) = g_i(\gamma(t)) = \int_0^1 \frac{\partial f}{\partial x_i}(\gamma(t))\, dt$, which clearly vanish at p_0. ■

Proposition 64 *Let $P \in \mathcal{D}^k(E,F)$ be a kth order linear differential operator. If there are functions $f_i, g_i \in C^\infty(M)$, $i = 1,\dots,k$, such that $df_i(p) = dg_i(p)$ for all i at a point $p_0 \in M$, then*

$$\mathrm{ad}^k(f_1,\dots,f_k)(P)|_{p_0} = \mathrm{ad}^k(g_1,\dots,g_k)(P)|_{p_0}$$

as zeroth order operators at p_0, i.e. as $\mathbb{C}$-linear maps of vector spaces E_{p_0} to F_{p_0}.

Proof: Since $\mathrm{ad}_c(P) = [P,c] = 0$ for any constant $c \in \mathbb{C}$, by the $\mathcal{C}$-linearity of P, we may assme, by adding constants to the f_i and g_i, that $f_i(p_0) = g_i(p_0)$ for all i. Then, let $\psi := f_1 - g_1$ and let $Q = \mathrm{ad}^{k-1}(f_2,\dots,f_k)(P) \in \mathcal{D}^1(E,F)$ and note that $\mathrm{ad}_\psi(Q)|_{p_0} = 0$ in $\mathrm{Hom}(E_{p_0}, F_{p_0})$: since $\psi \in \mathfrak{I})p_0$, we must have $\psi = \sum_{j=1}^n a_j b_j$ for $a_j, b_j \in \mathfrak{m}_{p_0}$, which means

$$\mathrm{ad}_\psi(Q)|_{p_0} \quad = \quad \sum_{j=1}^n \mathrm{ad}_{a_j b_j}(Q)|_{p_0} \quad = \quad \sum_{j=1}^n [Q, a_j b_j] \quad = \quad \sum_{j=1}^n \Big(a_j[Q,b_j] \ + \ [Q,a_j]b_j \Big)\Big|_{p_0} \quad = \quad 0$$

since the a_j and b_j belong to $\mathfrak{m}_{p_0}$. Letting $\psi_i = f_i - g_i$ and repeating the above procedure for all i shows that $\mathrm{ad}^k(\psi_1,\dots,\psi_n)(P)|_{p_0} = 0$. ∎

Definition 45 If $P \in \mathcal{D}^k(E,F)$ and $f_1,\dots,f_k \in C^\infty(M)$, then at any point $p_0 \in M$ we have a linear map

$$\frac{1}{k!}\,\mathrm{ad}^k(f_1,\dots,f_k)(P)|_{p_0} = \frac{1}{k!}\,\mathrm{ad}_{f_1} \circ \cdots \circ \mathrm{ad}_{f_k}(P)|_{p_0} \in \mathrm{Hom}_{\mathbb{C}}(E_{p_0}, F_{p_0})$$

which dependes on the values of the differentials $df_i(p_0) \in T^*_{p_0}M$ at p_0. If $\xi_i \in T^*_{p_0}M$ are 1-forms evaluated at p_0, then locally, in a neighborhood about p_0, the Poincaré lemma says $\xi_i = df_i$ near p_0, and by the above proposition the choice of such f_i is immaterial so long as two choices agree at p_0. Thus, for any $P \in \mathcal{D}^k(E,F)$ we get a unique map,

$$\sigma(P)(\xi_1,\dots,\xi_k) \in \mathrm{Hom}_{\mathbb{C}}(E_{p_0}, F_{p_0}) \tag{3.86}$$

which is, moreover, symmetric in the ξ_i by Proposition 62. Since for any vector space V we have the isomorphisms $\mathrm{Hom}^k_{\mathbb{R},\mathrm{Sym}}(V;\mathbb{R}) \cong (V^*)^{\odot k} = \mathbb{R}_k[x_1,\dots,x_n]$, and for any such symmetric form f we have the **polarization formula** (Nicolaescu [85, p. 329])

$$f(v_1\dots,v_k) = \frac{1}{k!}\frac{\partial^k}{\partial t_1 \cdots \partial t_k} f(t_1v_1 + \cdots + t_kv_k, \dots, t_1v_1 + \cdots + t_kv_k)$$

because the right-hand-side equals $\frac{1}{k!}\sum_{i=1}^n \frac{\partial f}{\partial v_i} v_i$ and f is a polynomial. By means of this polarization formula we see that $\sigma(P)$ may just as well be defined, not in terms of k different f_i, but in terms of a single f, which gives a single $\xi = df(p_0) \in T^*_{p_0}M$,

$$\sigma(P)(\xi, \ldots, \xi) \in \mathrm{Hom}_{\mathbb{C}}(E_{p_0}, F_{p_0}) \tag{3.87}$$

This is the **principal symbol of** P. Clearly, then $\sigma(P)$ induces a map

$$\sigma(P) : (T^*M)^{\odot k} \to \mathrm{Hom}(E, F) \tag{3.88}$$

which we call the **symbol map**. We can describe this in terms of a short exact sequence,

$$0 \to \mathcal{D}^{k-1}(E, F) \to \mathcal{D}^k(E, F) \xrightarrow{\sigma} \big(\pi^*(\mathrm{Hom}(E, F)) \to S^*M^{\odot k}\big) \to 0 \tag{3.89}$$

where the space on the right is the space of all sections of $\pi^*(\mathrm{Hom}(E, F))$ restricted to the cosphere bundle (meaning the nonzero, unit-length ξ in T^*M), the pullback being via the projection $\pi : T^*M \to M$. For since $\pi^*(E) \subseteq T^*M \times E$, we can view the symbol map as taking nonzero (hence, why not unit-length) elements $\xi \odot \cdots \odot \xi$ of $(T^*M)^{\odot k}$ into $\mathrm{Hom}(E, F)$, viewed as pulled back over T^*M by π. ■

Definition 46 If the symbol $\sigma(P)$ of an operator $P \in \mathcal{D}^k(E, F)$ is invertible for all nonzero ξ, then we say that P is **elliptic**. ■

3.2 Essentially Self-Adjoint Operators

3.2.1 The General Theory of Essential Self-Adjointness on a Hilbert Space

Given a symmetric unbounded operator $(T, D(T))$ on a Hilbert space $\mathcal{H}$, where $D(T) \subseteq \mathcal{H}$ is the dense domain of T and $T \subset T^*$, that is $D(T) \subseteq D(T^*)$, we are interested to know the conditions under which T is self-adjoint, $T = T^*$, or at least under which T admits a unique self-adjoint extension $\overline{T} = \overline{T}^* \supset T$; in the language of functional analysis, such a T is called *essentially self-adjoint*. A major reason for this interest is the spectral theorem (Theorem VIII.6 Reed and Simon [89]), which says that if T is self-adjoint, then it has an integral representation $T = \int_{\mathbb{R}} \lambda \, dP_\lambda$, where P_λ is a spectral measure, or projection-valued measure, on $\mathcal{H}$, and this allows us to define $g(T)$ for any real-valued Borel measurable function g on $\mathbb{R}$, by letting $g(T) = \int_{\mathbb{R}} g(\lambda) dP_\lambda$. In particular, applying this to the real and imaginary parts of $g(x) = e^{ix}$ we can define the exponential e^{iT}. Indeed, we can do more: for any $t \in \mathbb{R}$ we can define e^{itT}, and it is rather immediate that this operator behaves much like the usual exponential: (1) $e^{i(s+t)T} = e^{isT} \circ e^{itT}$ for all $s, t \in \mathbb{R}$, (2) e^{itT} is continuous and even differentiable in t. Moreover, e^{itT} is a unitary operator, an element of $\mathrm{U}(\mathcal{H})$, and by Stone's Theorem (Theorems VIII.7-VIII.8 Reed and Simon Vol.1 [89]) any such unitary operator-valued map $U : \mathbb{R} \to \mathrm{U}(\mathcal{H})$ satisfying $U(s+t) = U(s) \circ U(t)$ and $\lim_{t \to t_0} U(t) = U(t_0)$ is generated by a self-adjoint operator T, so $U(t) = e^{itT}$. Thus, self-adjoint operators, spectral measures, and exponentials are intimately connected. And the distinction between merely symmetric operators, which only satisfy $T \subset T^*$, and self-adjoint operators becomes apparent here. The spectrum of a self-adjoint operator is a subset of the reals, and it is from this spectrum that we can construct our spectral measures and the functional calculus which gives us the one-parameter group of unitary operators e^{itT}. Symmetric operators, by contrast, generally have spectra lying outside of the reals, and it is impossible to construct their integral representations and the functional calculus. However, in certain cases it is possible to extend a given symmetric operator to a unique self-adjoint operator, and we will look for means of achieving this in our case, for our symmetric symplectic Dirac operators D and $\tilde{D}$.

From our point of view, and from the point of view of index theory generally, the relevance of self-adjointness lies in the connection between the index of a certain self-adjoint first order differential operator (the Dirac operator acting on the L^2-closure of the space of spinors) and its exponential, which is used to construct an integral kernel called the heat kernel. The first few terms in the asymptotic expansion this heat kernel yield highly interesting topological information about the underlying spin manifold over which the spinors lie, and this information is related back to the purely analytic index of the Dirac operator by the famous Atiyah-Singer Index Theorem. This, at least, is the case for the Dirac operator on a Riemannian spin manifold. Our interest is in a similar idea, which concerns a slightly different, though neighboring, area, that of symplectic geometry. Over any symplectic manifold (M, ω) we have two symplectic Dirac operators D and $\tilde{D}$ acting on the L^2-closure of the space of *symplectic* spinors sitting over M. If we wish to proceed along similar lines to the Riemannian case, we need to first assure ourselves of the self-adjointness of our Dirac operators, and that is the purpose of these notes.

In this section we review the theory of unbounded symmetric operators and essentially self-adjoint operators on a Hilbert space, culminating in the theorems of von Neumann characterizing the domain of an essentially self-adjoint symmetric operator. We will use these theorems below in our proof of the essential self-adjointness of the symplectic Dirac operators D and $\tilde{D}$.

Notation 5 *Throughout this section we denote an abstract Hilbert space by $\mathcal{H}$ and the innner product on $\mathcal{H}$ by $(\cdot, \cdot)$, unless there is a need to be specific about the Hilbert space to which $(\cdot, \cdot)$ belongs, in which case we will indicate the dependence by $(\cdot, \cdot)_{\mathcal{H}}$.* ■

Definition 47 Let T be a densely defined unbounded linear operator on a Hilbert space $\mathcal{H}$, with dense domain $\mathcal{D}(T)$. Now, $\mathcal{H} \oplus \mathcal{H}$ is also a Hilbert space, with inner product

$$\big((x_1, y_1), (x_2, y_2)\big)_{\mathcal{H}\oplus\mathcal{H}} := (x_1, x_2)_{\mathcal{H}} + (y_1, y_2)_{\mathcal{H}}$$

and norm

$$\|(x, y)\|_{\mathcal{H}\oplus\mathcal{H}} = \left(\|x\|^2 + \|y\|^2\right)^{1/2}$$

The **graph** of T is the subspace of $\mathcal{H} \oplus \mathcal{H}$

$$\Gamma(T) := \{(x, Tx) \mid x \in \mathcal{D}(T)\}$$

and we say that T is **closed** if $\Gamma(T)$ is closed in $\mathcal{H} \oplus \mathcal{H}$. By the Closed Graph Theorem (Reed and Simon [89, Theorem III.2, p. 83]) we know that when $\mathcal{D}(T) = \mathcal{H}$, the operator T is closed iff it is bounded. Therefore a closed operator T is undbounded iff $\mathcal{D}(T) \subsetneq \mathcal{H}$. An **extension** of T is an operator T_1 with domain $\mathcal{D}(T_1) \supseteq \mathcal{D}(T)$ and agreeing with T on $\mathcal{D}(T)$, which makes T the **restriction** of T_1. We denote this relationship by

$$T \subset T_1$$

An example of an extension of T is its **closure** $\overline{T}$, which, when it exists, is by definition the operator obtained from T by taking the closure of its graph. Since closedness can be interpreted in terms of limits of sequences, if we take a Cauchy sequence $(x_n, Tx_n)_{n\in\mathbb{N}}$ in $\Gamma(T)$, then its limit lies in $\overline{\Gamma(T)}$. By the definition of the norm on $\mathcal{H} \oplus \mathcal{H}$, this means

$$\begin{aligned} \lim_{n\to\infty}(x_n, T(x_n)) =: (x, y) \in \overline{\Gamma(T)} \quad &\iff \quad \lim_{n\to\infty}\left(\|x_n - x\|^2 + \|T(x_n) - y\|^2\right)^{1/2} = 0 \\ &\iff \quad \lim_{n\to\infty}\|x_n - x\| = 0 \;\text{ and }\; \lim_{n\to\infty}\|T(x_n) - y\| = 0 \end{aligned}$$

so if we define $\overline{T}(x) := \lim_{n\to\infty} T(x_n)$, then we see that

$$\overline{T}(\lim_{n\to\infty} x_n) = \overline{T}(x) = \lim_{n\to\infty} \overline{T}(x_n)$$

That is, $\overline{T}$ is the operator whose domain is precisely the set of limits $x = \lim_{n\to\infty} x_n$ which are preserved by $\overline{T}$,

$$\begin{aligned} \mathcal{D}(\overline{T}) &= \{\lim_{n\to\infty} x_n \in \mathcal{H} \mid (x_n)_{n\in\mathbb{N}} \text{ is Cauchy in } \mathcal{D}(T) \text{ and } (Tx_n)_{n\in\mathbb{N}} \text{ is Cauchy in } \mathcal{H}\} \\ &= \{x \in \mathcal{H} \mid x = \lim_{n\to\infty} x_n \text{ for some } (x_n)_{n\in\mathbb{N}} \text{ in } \mathcal{D}(T) \text{ and } \lim_{n\to\infty} Tx_n \text{ exists in } \mathcal{H}\} \end{aligned}$$

It is one of the basic facts of life that unbounded operators do not preserve all limits, else they would be continuous and bounded. Closed operators at least preserve *some* limits, the ones which *can* exist, namely those for which $(T(x_n))_{n\in\mathbb{N}}$ converges whenever $(x_n)_{n\in\mathbb{N}}$ does. Of course, the closure of $\Gamma(T)$ need not produce an *operator* whose graph is $\overline{\Gamma(T)}$, since, as we will see, this can only happen of the adjoint of T has a dense domain, and there are cases where this fails (see Reed and Simon [89, Ex. 4, p. 252]). In the case that $\overline{\Gamma(T)}$ is the graph of an operator $\overline{T}$, we say that T is **closable**. A **core** for a closable operator T is a subset C of $\mathcal{D}(T)$ such that $\overline{T|_C} = \overline{T}$. This is a useful notion, since we may know something about C though not much about $\mathcal{D}(T)$, and this is sometimes enough to begin studying $\overline{T}$. ■

To understand closed and closable operators, then, we need to understand the other object to which these notions are attached, the adjoint operator T^* of an operator T.

Definition 48 Let T be a densely defined unbounded operator on a Hilber space $\mathcal{H}$ with dense domain $\mathcal{D}(T)$. The **adjoint** T^* of T is the (not necessarily densely defined) linear operator on $\mathcal{H}$ with domain

$$\mathcal{D}(T^*) = \{y \in \mathcal{H} \mid \exists z \in \mathcal{H} \text{ s.t. } (Tx, y) = (x, z)\ \forall x \in \mathcal{D}(T)\}$$

It is not immediately obvious that we can define T^*y to be this $z \in \mathcal{H}$, because z may not be unique. However, the following proposition shows that the denseness of $\mathcal{D}(T)$ in $\mathcal{H}$ guarantees the uniqueness of z, as a simple consequence of the the Riesz Representation Theorem. We remark that $\mathcal{D}(T^*)$ may not be dense even if $\mathcal{D}(T)$ is, as mentioned above. ■

Proposition 65 *The denseness of $\mathcal{D}(T)$ in $\mathcal{H}$ implies the uniqueness of the element $z \in \mathcal{H}$ for each $y \in \mathcal{D}(T^*)$, making T^* a well-defined operator. As a byproduct, we have the following characterization of $\mathcal{D}(T^*)$: An element $y \in \mathcal{H}$ belongs to $\mathcal{D}(T^*)$ iff the associated linear functional $(T\cdot, y)$ is bounded/continuous on $\mathcal{D}(T)$, meaning that there exists a constant $C > 0$ such that $|(Tx, y)| \leq C\|x\|$ for all $x \in \mathcal{D}(T)$.*

Proof: First, we observe that $y \in \mathcal{D}(T^*)$ iff the linear functional $(T\cdot, y)$ is bounded on $\mathcal{D}(T)$. For suppose that for some $y \in \mathcal{H}$ we know that $(T\cdot, y) \in \mathcal{D}(T)^*$, i.e. is bounded on $\mathcal{D}(T)$. Then the denseness of $\mathcal{D}(T)$ in $\mathcal{H}$ means we can extend $(T\boldsymbol{\cdot}, y)$ to a unique bounded linear extension $\eta \in \mathcal{H}^*$ on all of $\mathcal{H}$ (without denseness we would have to use the Hahn-Banach Extension Theorem, Reed and Simon [89, Corollary 1, p. 77], Kadison and Ringrose [62, p. 10], which would give the existence but not the uniqueness of η). By the Riesz Representation Theorem (Reed and Simon [89, Lemma II.2, p. 42]) there is a unique element $z \in \mathcal{H}$ such that

$$\eta(x) = (x, z), \quad \forall x \in \mathcal{H}$$

Restricting η to $\mathcal{D}(T)$ then gives

$$(Tx, y) = \eta(x) = (x, z), \quad \forall x \in \mathcal{H}$$

But then $y \in \mathcal{D}(T^*)$ by the definition of $\mathcal{D}(T^*)$. Conversely, suppose that $y \in \mathcal{D}(T^*)$. Then there is a $z \in \mathcal{H}$ such that for all $x \in \mathcal{D}(T)$ we have $(Tx, y) = (x, z)$. The Cauchy-Schwarz inequality gives the boundedness of $(T\boldsymbol{\cdot}, y)$,

$$|(Tx, y)| = |(x, z)| \leq \|z\|\|x\| = C\|x\|, \quad \text{where } C = \|z\|$$

This proves our claim that $y \in \mathcal{D}(T^*)$ iff $(T\boldsymbol{\cdot}, y) \in \mathcal{D}(T)^*$.

Now we can prove that z is unique. If z and z' are two such elements of $\mathcal{H}$ for a given y, i.e. satisfying

$$(\boldsymbol{\cdot}, z) = (T\boldsymbol{\cdot}, y) = (\boldsymbol{\cdot}, z') \in \mathcal{D}(T)^*$$

then by this assumption their associated bounded functionals agree on the dense space $\mathcal{D}(T)$, and this bounded functional extends by continuity to a unique bounded linear functional η on $\mathcal{H}$, so using the Riesz Representation Theorem we know that η can be represented by a unique vector $w \in \mathcal{H}$, i.e. $\eta(x) = (x, w)$ for all $x \in \mathcal{H}$. Restricting η to $\mathcal{D}(T)$ we then conclude that $z = w = z'$ by the uniqueness of w. ∎

The following theorem characterizes closability, and our proof follows those in Reed and Simon [89, Theorem VIII.1, p. 252] and Kadison and Ringrose [62, 2.7.8].

Theorem 30 *Let T be a densely defined unbounded linear operator on a Hilbert space $\mathcal{H}$. Then the following hold:*

(1) *T^* is closed.*

(2) *If T is closable, then $(\overline{T})^* = T^*$.*

(3) *T is closable iff $\mathcal{D}(T^*)$ is dense. In this case, $\overline{T} = T^{**}$.*

(4) *Let $V : \mathcal{H} \oplus \mathcal{H} \to \mathcal{H} \oplus \mathcal{H}$ be given by $V(x, y) = (-y, x)$. Then*

$$\Gamma(T^*) = V(\Gamma(T))^\perp$$

Proof: We start by proving (4), since we will use the map V to prove the other assertions. First note that the map V is unitary: $\|V(x, y)\|^2 = \|(-y, x)\|^2 = \| - y\|^2 + \|x\|^2 = \|(x, y)\|^2$, for all $x, y \in \mathcal{H}$. To prove the first assertion, note first that if E is a subspace of $\mathcal{H} \oplus \mathcal{H}$, then $V(E^\perp) = V(E)^\perp$, because V is unitary. Taking $E = \Gamma(T)$, we see that $V(\Gamma(T)^\perp) = V(\Gamma(T))^\perp$, so for any $(x, Tx) \in \Gamma(T)$, i.e. for all $x \in \mathcal{D}(T)$, and any $(k, h) \in \mathcal{H} \oplus \mathcal{H}$ we have that

$$\begin{aligned}
(k, h) \perp V(x, Tx) = (-Tx, x) \quad &\iff \quad \big((k, h), (-Tx, x)\big)_{\mathcal{H}\oplus\mathcal{H}} = 0 \\
&\iff \quad -(k, Tx) + (h, x) = 0 \\
&\iff \quad (x, h) = (Tx, k) \\
&\iff \quad (k, h) = (k, T^*k) \in \Gamma(T^*)
\end{aligned}$$

which shows that $(k, h) \in \Gamma(T^*)$ iff $(k, h) \in V(\Gamma(T))^\perp$, and thus

$$\Gamma(T^*) = V(\Gamma(T))^\perp$$

(1) Now, $V(\Gamma(T))^\perp$ is closed in $\mathcal{H} \oplus \mathcal{H}$ (the orthogonal complement of any subset S of $\mathcal{H}$ is always closed, since the inner product $(\cdot, \cdot)$ is continuous in each variable by the Cauchy-Schwarz inequality and $S^\perp = \bigcap_{s\in S}(s, \cdot)^{-1}(0)$ is the intersection of the closed spaces $(s, \cdot)^{-1}(0)$ for each $s \in S$), so we conclude that T^* is closed.

(2) Suppose T is closable, and observe that $T \subseteq \overline{T}$ implies $(\overline{T})^* \subseteq T^*$, so we need only prove the reverse inclusion. Let $y \in \mathcal{D}(T^*)$, so that $(Tx, y) = (x, z)$ for some $z =: T^*y \in \mathcal{H}$

and all $x \in \mathcal{D}(T)$. To show that $y \in \mathcal{D}(\overline{T}^*)$ we must show that for all $x \in \mathcal{D}(\overline{T})$ we have $(\overline{T}x, y) = (x, z)$ for some unique $z \in \mathcal{H}$ depending on y. By definition of $\overline{T}$ as the closure of $\Gamma(T)$, any $x \in \mathcal{D}(\overline{T})$ is the limit of a sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathcal{D}(T)$ for which the sequence $(T(x_n))_{n\in\mathbb{N}}$ converges to $\overline{T}x := \lim_{n\to\infty} T(x_n)$. Moreover, since $y \in \mathcal{D}(T^*)$, Proposition 65 tells us that $(T\cdot, y) = (\cdot, T^*y)$ is bounded/continuous on $\mathcal{D}(T)$, and extends to a bounded linear functional on $\mathcal{H}$, so it preserves limits of sequences in $\mathcal{D}(T)$. Thus, we have for our $x = \lim_{n\to\infty} x_n$ that

$$(\overline{T}x, y) = (\lim_{n\to\infty} Tx_n, y) = \lim_{n\to\infty} (Tx_n, y) = \lim_{n\to\infty} (x_n, T^*y) = (\lim_{n\to\infty} x_n, T^*y) = (x, T^*y)$$

which shows that $y \in \mathcal{D}(\overline{T})$.

Alternatively, (2) follows from (3) by noting that if T is closable, then $\overline{T} = T^{**}$, so $T^* = \overline{T^*} = T^{***} = (T^{**})^* = (\overline{T})^*$.

(3) Since $\Gamma(T)$ is a subspace of $\mathcal{H} \oplus \mathcal{H}$, we have by (4) that

$$\overline{\Gamma(T)} = \big(\Gamma(T)^\perp\big)^\perp = V^2\big(\Gamma(T)^\perp\big)^\perp = V\big(\Gamma(T^*)\big)^\perp$$

If $\mathcal{D}(T^*)$ is dense, then (4) applies again to give $\overline{\Gamma(T)} = V\big(\Gamma(T^*)\big)^\perp = \Gamma(T^{**})$, which shows that $\overline{T} = T^{**}$, and since T^{**} is an operator, we see that T is closable with closure T^{**}. We will prove the converse, T is closable implies $\mathcal{D}(T^*)$ is dense, by proving its contrapositive: Suppose that $\mathcal{D}(T^*)$ is not dense. Then $\mathcal{D}(T^*)^\perp$ is nontrivial and $\mathcal{H} = \mathcal{D}(T^*) \oplus \mathcal{D}(T^*)^\perp$, so we may take a nonzero $y \in \mathcal{D}(T^*)^\perp$. Then $(y, 0) \in \Gamma(T^*)^\perp$, since for all $x \in \mathcal{D}(T^*)$ we have $x \perp y$, which means

$$\big((x, T^*x),\ (y, 0)\big) = (x, y) + (T^*x, 0) = 0$$

But then $(0, y) \in V\big(\Gamma(T^*)\big)^\perp$, yet $(0, y)$ clearly cannot be in the graph of any linear operator since $y \neq 0$. In particular $(0, y) \notin \Gamma(\overline{T})$, and since $V\big(\Gamma(T^*)\big)^\perp = \overline{\Gamma(T)}$, this shows that $\overline{T}$ does not exist as a linear operator. ∎

Definition 49 A densely defined operator T on a Hilbert space $\mathcal{H}$ with domain $\mathcal{D}(T)$ is called **symmetric** or **formally self-adjoint** if $T \subseteq T^*$, that is if $\mathcal{D}(T) \subseteq \mathcal{D}(T^*)$ and $T^*|_{\mathcal{D}(T)} = T$. By the previous proposition we see that all symmetric operators are closable, with one closed extension being T^*, since the denseness of $\mathcal{D}(T)$ implies the denseness of $\mathcal{D}(T^*)$. Since $\overline{T}$ is the smallest closed extension of T, we have

$$T \subseteq \overline{T} = T^{**} \subseteq T^*$$

If additionally $T \supseteq T^*$, so that $T = T^*$, then we say that T is **self-adjoint**. In this case, the above inclusions become equalities,

$$T = \overline{T} = T^{**} = T^*$$

Thus, self-adjoint operators are necessarily closed. Finally, a symmetric operator $T \subseteq T^*$ is called **essentially self-adjoint** if its closure $\overline{T}$ is self-adjoint, $\overline{T} = (\overline{T})^*$. The importance of this idea lies in the fact that we do not always know the domain $\mathcal{D}(T)$ of T, whereas sometimes we can at least say something about a core C for T on which T is symmetric, i.e. $T|_C \subseteq T|_C^*$. ■

Lemma 11 *T is symmetric iff (Tx, x) is real for all $x \in \mathcal{D}(T)$.*

Proof: If T is symmetric, then $(Tx, x) = (x, Tx) = \overline{(Tx, x)}$, so $(Tx, x) \in \mathbb{R}$ for all $x \in \mathcal{D}(T)$. Conversely, if $(Tx, x) \in \mathbb{R}$ for all $x \in \mathcal{D}(T)$, then by the polarization identity in the form

$$4(Tx, y) = (T(x+y), x+y) - (T(x-y), x-y) + i(T(x+iy), x+iy) - i(T(x-iy), x-iy)$$

(which is proved by computing the right-hand-side, see Schmüdgen [94, p. 5]), we have for all $x, y \in \mathcal{D}(T)$ that

$$\begin{aligned} 4(Tx, y) &= (T(x+y), x+y) - (T(x-y), x-y) \\ &\qquad + i(T(x+iy), x+iy) - i(T(x-iy), x-iy) \\ &= (x+y, T(x+y)) - (x-y, T(x-y)) \\ &\qquad + i(x+iy, T(x+iy)) - i(x-iy, T(x-iy)) \\ &= 4(x, Ty) \end{aligned}$$

and T is symmetric. ■

Lemma 12 *Let $T \subseteq T^*$ be a symmetric operator on a Hilbert space $\mathcal{H}$. Then* $\operatorname{im}(T \pm iI)$ *is closed in $\mathcal{H}$.*

Proof: Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{D}(T)$ such that $\big((T \pm iI)(x_n)\big)_{n\in\mathbb{N}}$ converges to $y \in \mathcal{H}$, so by the previous lemma we have that (Tx, x) is real for all $x \in \mathcal{D}(T)$, and therefore

$$\begin{aligned}\|x\|^2 = \big((x,x)^2\big)^{1/2} &\leq \big((Tx,x)^2 + (x,x)^2\big)^{1/2} \\ &= |(Tx,x) \pm i(x,x)| = |((T \pm iI)x, x)| \leq \|(T \pm iI)x\|\|x\|\end{aligned}$$

which implies $\|x\| \leq \|(T \pm iI)x\|$. Consequently, $\|x_n - x_m\| \leq \|(T \pm iI)(x_n - x_m)\| \to 0$ which shows that $(x_n)_{n\in\mathbb{N}}$ is Cauchy and therefore convergent, say with limit $x \in \mathcal{H}$. Since by assumption $(T(x_n))_{n\in\mathbb{N}}$ converges to $y \mp ix$ because $(T \pm iI)x_n \to y$, and by assumption T is closed, we must have $x \in \mathcal{D}(T)$ and $y \mp ix = Tx$, so $(T \pm iI)x = y$, and $T \pm iI$ have closed ranges. ■

The following theorem and its corollary give the basic criterion for essential self-adjointness of a symmetric operator. It was first proved by von Neumann in the foundational 1929-1930 paper [107], and may now be found in every standard textbook on functional analysis, for example Reed and Simon [89, Theorem VIII.3, p. 256], or Kadison and Ringrose [62, 2.7.10].

Theorem 31 *Let T be a closed symmetric operator on a Hilbert space $\mathcal{H}$. Then the following are equivalent:*

(1) *T is self-adjoint, that is $T = T^*$.*

(2) $\operatorname{im}(T \pm iI) = \mathcal{H}$.

(3) $\overline{\operatorname{im}(T \pm iI)} = \mathcal{H}$.

(4) $\ker(T^* \pm iI) = \{0\}$

Proof: (1) $\implies$ (4): Suppose $T = T^*$. If $x \in \mathcal{D}(T) = \mathcal{D}(T^*)$ also lies in $\ker(T^* \pm iI)$, then $Tx = T^*x = \mp ix$, so

$$\mp i\|x\|^2 = \mp i(x,x) = (\mp ix, x) = (Tx, x) = (x, Tx) = (x, \mp ix) = \pm i(x,x) = \pm i\|x\|^2$$

which shows that $\|x\| = 0$ and therefore $x = 0$.

(4) $\implies$ (2) $\iff$ (3): Suppose $\ker(T^* \pm iI) = \{0\}$. By Lemma 12, $T \subseteq T^*$ implies that $\mathrm{im}(T \pm iI)$ is closed, hence it suffices to show that it is dense in $\mathcal{H}$, or equivalently that the orthogonal complement of $\mathrm{im}(T \pm iI)$ is trivial, since in general $M^\perp = \overline{M}^\perp$ for any subspace M of $\mathcal{H}$. Now, if $y \in \mathrm{im}(T \pm iI)^\perp$, then $((T \pm iI)x, y) = 0$ for all $x \in \mathcal{D}(T)$, which means that $(Tx, y) = \mp i(x, y) = (x, \pm iy)$, or $y \in \mathcal{D}(T^*)$ and therefore $T^*y = \pm iy$, i.e. $y \in \ker(T \pm iI) = \{0\}$, or $y = 0$. Thus (4) implies (2).

(2) $\implies$ (1): Suppose $\mathrm{im}(T \pm iI) = \mathcal{H}$. Since $T \subseteq T^*$ and T is closed, $\Gamma(T)$ is closed and contained in $\Gamma(T^*)$. Thus, let V be the orthogonal complement of $\Gamma(T)$ in $\Gamma(T^*)$,

$$\Gamma(T^*) = \Gamma(T) \bigoplus^{\perp} V$$

Then, for all $(y, T^*y) \in V$ we will have $(y, T^*y) \perp (x, Tx)$ for all $x \in \mathcal{D}(T)$, so

$$0 = \big((y, T^*y),\ (x, Tx)\big) = (y, x) + (T^*y, Tx)$$

But by assumption $\mathrm{im}(T \pm iI) = \mathcal{H}$, so there is some $x \in \mathcal{D}(T)$ for which

$$(T^2 + I)x = (T + iI) \circ (T - iI)x = y$$

so from this and the previous equality we have

$$\|y\|^2 = (y, y) = (y, (T^2 + I)x) = (y, T^2x) + (y, x) = (T^*y, Tx) + (y, x) = -(y, x) + (y, x) = 0$$

Therefore, $y = 0$ and $T^*y = T^*0 = 0$, so $V = \{0\}$ and $\Gamma(T) = \Gamma(T^*)$, or $T = T^*$. ■

Corollary 17 *If T is a symmetric operator on a Hilbert space $\mathcal{H}$, then the following are equivalent:*

(1) *T is essentially self-adjoint.*

(2) $\ker(T^* \pm iI) = \{0\}$.

(3) $\overline{\mathrm{im}(T \pm iI)} = \mathcal{H}$. ■

Proof: If T is essentially self-adjoint, then it is closable, so by Proposition 30 we have $\overline{T} = (\overline{T})^* = T^*$, and since the previous theorem applies to $\overline{T}$, we get

$$\begin{aligned} \overline{T} = (\overline{T})^* \quad &\Longleftrightarrow \quad \ker(T^* \pm iI) = \ker(\overline{T}^* \pm iI) = \{0\} \\ &\Longleftrightarrow \quad \overline{\operatorname{im}(T \pm iI)} = \ker(T^* \mp iI)^\perp = \{0\}^\perp = \mathcal{H} \end{aligned}$$

which completes the proof. ■

To determine whether *any* (not-necessarily unique) self-adjoint extensions of a given symmetric operator T exist, we need a further criterion (which has the additional benefit of giving us a second criterion for essential self-adjointness). The **deficiency subspaces** of $\mathcal{H}$ for T are defined to be

$$\begin{aligned} \ker(T^* - iI) &= \operatorname{im}(T + iI)^\perp \\ \ker(T^* + iI) &= \operatorname{im}(T - iI)^\perp \end{aligned}$$

and their dimensions

$$\begin{aligned} n_+ &= \dim \ker(T^* - iI) \\ n_- &= \ker(T^* + iI) \end{aligned}$$

are called the **deficiency indices** of T. Then, assuming that T is closed (which is not a severe restricition as symmetric operators are always closable), we have:

Theorem 32 *If T is a closed symmetric operator on a Hilbert space $\mathcal{H}$ with deficiency indices $n_\pm$, then T has self-adjoint extensions iff $n_+ = n_-$, and T is self-adjoint iff $n_+ = n_- = 0$.* ■

This theorem was first proved by von Neumann in the same 1929-1930 paper [107] cited above. The proof we follow is found in Reed and Simon Vol.2 [90, p. 141], who credit it to Dunford and Schwartz [30], and requires the following proposition (Reed and Simon Vol.2 [90, p. 138]), which we need to preface with some terminology.

Definition 50 The **graph inner product** on $\mathcal{D}(T^*)$, for a given closed symmetric operator T, is defined by

$$(x,y)_T := (x,y)_{\mathcal{H}} + (T^*x, T^*y)_{\mathcal{H}}$$

where $(\cdot,\cdot)_{\mathcal{H}}$ is the inner product on $\mathcal{H}$. We say that vectors $u, v \in \mathcal{D}(T^*)$ are T-**orthogonal** if they are so with respect to $(\cdot,\cdot)_T$, and likewise subspaces U and V of $\mathcal{D}(T^*)$ are T-**orthogonal**, denoted $U \oplus_T V$, if they are so with respect to $(\cdot,\cdot)_T$. Analogously, we say that a subspace U of $\mathcal{D}(T^*)$ is T-**closed** in $\mathcal{D}(T^*)$ if it is so with respect to the metric topology derived from the graph inner product. We also introduce the sesquilinear form

$$[x,y]_T := (T^*x, y)_{\mathcal{H}} - (x, T^*y)_{\mathcal{H}}$$

and we say that a subspace X of $\mathcal{D}(T^*)$ is T-**symmetric** if $[x,y]_T = 0$ for all $x, y \in X$. ■

With this terminology in hand, we can state the preliminary theorem:

Proposition 66 *If T is a closed symmetric operator with dense domain $\mathcal{D}(T)$ in a Hilbert space $\mathcal{H}$, then*

$$\mathcal{D}(T^*) = \mathcal{D}(T) \oplus_T \ker(T^* - iI) \oplus_T \ker(T^* + iI)$$ ■

We begin the proofs of Theorems 66 and 32 with some brief observations. From their statements and the statement of Theorem 31 we can see the importance of the kernles $\ker(T^* \pm iI)$. Let us suppose for a minute that one of these was nontrivial, say $\ker(T^* - iI)$. Then a nonzero element x of $\ker(T^* - iI)$ satisfies $T^*x = ix$, which shows that i is an eigenvalue of T^*. It is, of course, possible that i is also an eigenvalue of T, or, if not an eigenvalue, then perhaps an element of its continuous or residual spectrum. Were T self-adjoint, this would be impossible, as we are about to show. In any case, it appears that there is a connection between the spectrum of a symmetric operator T, the kernels $\ker(T^* \pm iI)$, and the possible self-adjoint extensions of T. The precise connection will be the content of the proofs of the theorems stated above. Let us first lay down the necessary terminology concerning spectra and eigenvalues for unbounded operators, as we will need these later anyway, before we employ them to prove our main theorems.

Definition 51 Let T be a closed unbounded operator, and define the **resolvent set** of T to be the subset of $\mathbb{C}$

$$\rho(T) := \{\lambda \in \mathbb{C} \mid T - \lambda I : \mathcal{D}(T) \to \mathcal{H} \text{ is bijective and } (T - \lambda I)^{-1} \in \mathcal{B}(\mathcal{H}, \mathcal{D}(T))\}$$

and define the **spectrum** of T to be the complement of $\rho(T)$ in $\mathbb{C}$,

$$\sigma(T) := \mathbb{C}\backslash\rho(T) = \{\lambda \in \mathbb{C} \mid T - \lambda I \text{ is not a bijection between } \mathcal{D}(T) \text{ and } \mathcal{H}\}$$

Thus,

$$\mathbb{C} = \sigma(T) \sqcup \rho(T)$$

By Theorem VIII.2, p. 254, in Reed and Simon Vol. I [89] we know that $\rho(T)$ is an open subset of $\mathbb{C}$, and therefore $\sigma(T)$ is a closed subset of $\mathbb{C}$. Let us consider the ways in which a complex number λ can lie in $\sigma(T)$. There are three possibilities: (1) $T - \lambda I$ is not injective, (2) $T - \lambda I$ is injective, not surjective, but has dense image, and (3) $T - \lambda I$ is injective, not surjective nor has dense image. These three possibilities correspond to the three classes of spectral values, the **point spectrum** $\sigma_p(T)$, which consists of eigenvalues of T, the **continuous spectrum** $\sigma_c(T)$, which consists of spectral values whose associated "resolvent operators" $R_\lambda(T) : \operatorname{im}(T - \lambda I)^{-1} \to \mathcal{D}(T)$ (see below) are densely defined *un*bounded operators with domains $\operatorname{im}(T - \lambda I)^{-1}$, and the **residual spectrum** $\sigma_r(T)$ whose spectral values do not give densely defined resolvent operators at all. Thus,

$$\begin{aligned} \mathbb{C} &= \sigma(T) \sqcup \rho(T) \\ &= \sigma_p(T) \sqcup \sigma_c(T) \sqcup \sigma_r(T) \sqcup \rho(T) \end{aligned}$$

gives the set of all possible spectral and resolvent values for T. ∎

Remark 54 The two definitions of $\sigma(T)$ may not at first sight seem to be equivalent, but they are. To see this, take the second definition of $\sigma(T)$, and consider some $\lambda \in \mathbb{C}\backslash\sigma(T)$. Then $T - \lambda I$ is a bijection between $\mathcal{D}(T)$ and $\mathcal{H}$, and $(T - \lambda I)^{-1} : \mathcal{H} \to \mathcal{D}(T)$ is also a bijection. Since $\Gamma(T)$ is closed, $\Gamma(T - \lambda I)$ is closed, too, for if we take a sequence (x_n, Tx_n) converging to (x, Tx), then

$(x_n, (T-\lambda I)x_n)$ converges to $(x, (T-\lambda I)x)$. Clearly $\Gamma((T-\lambda I)^{-1})$ is also closed, because if $F \in \mathrm{End}(\mathcal{H} \oplus \mathcal{H})$ is the map $F(x,y) := (y,x)$, then F is unitary and $\Gamma((T-\lambda I)^{-1}) = F(\Gamma(T-\lambda I))$ is the image of a closed set under a unitary map. The Closed Graph Theorem then applies to give the boundedness of the everywhere-defined closed operator $(T-\lambda I)^{-1}$. Thus, we have $\lambda \in \rho(T)$, and as a corollary, we always have $(T-\lambda I)^{-1} \in \mathcal{B}(\mathcal{H}, \mathcal{D}(T))$. ■

Definition 52 From the observation above, we see that to each $\lambda \in \rho(T)$ in the resolvent set of a closed unbounded operator T we may assign the bounded operator $(T-\lambda I)^{-1} \in \mathcal{B}(\mathcal{H}, \mathcal{D}(T))$, and this assigment is called the **resolvent operator** of T:

$$R : \rho(T) \subseteq \mathbb{C} \to \mathcal{B}(\mathcal{H}, \mathcal{D}(T))$$

$$R(\lambda) \equiv R_\lambda(T) := (T-\lambda I)^{-1}$$

This operator-valued map is readily seen to be analyitic in λ on the open set $\rho(T)$ (Theorem VIII.2, p. 254, in Reed and Simon Vol. I [89]). ■

Notation 6 *Let us denote the open upper and lower half planes of the complex plane, respectively, by*

$$\mathbb{H}^+ = \{z \in \mathbb{C} \mid \mathrm{Im}\, z > 0\}, \qquad \mathbb{H}^- = \{z \in \mathbb{C} \mid \mathrm{Im}\, z < 0\}$$

and the closed upper and lower half planes, respectively, by

$$\overline{\mathbb{H}^+} = \{z \in \mathbb{C} \mid \mathrm{Im}\, z \geq 0\}, \qquad \overline{\mathbb{H}^-} = \{z \in \mathbb{C} \mid \mathrm{Im}\, z \leq 0\}$$

■

The proof of the next theorem follows that in Reed and Simon [90, Theorem X.I, p. 136].

Theorem 33 *Let $T \subseteq T^*$ be a closed symmetric unbounded operator on a Hilbert space $\mathcal{H}$. Then,*

(1) (a) $\dim\ker(T^* - \lambda I)$ *is constant on the open upper half plane* $\mathbb{H}^+$.

(b) $\dim\ker(T^* - \lambda I)$ *is constant on the open lower half plane* $\mathbb{H}^-$.

(2) $\sigma(T) \in \{\overline{\mathbb{H}^+},\ \overline{\mathbb{H}^-},\ \mathbb{C},$ *some subset* $X \subseteq \mathbb{R}\}$

(3) $T = T^*$ *iff* $\sigma(T) \subseteq \mathbb{R}$ *iff* $\dim\ker(T^* - \lambda I) = 0$ *on* $\mathbb{C}\backslash\mathbb{R} = \mathbb{H}^+ \sqcup \mathbb{H}^-$.

Proof: (1) Let $\lambda = a + ib$, with $b \neq 0$, and note that for all $x \in \mathcal{D}(T)$ the symmetry of T implies

$$\begin{aligned}
\|(T - \lambda I)x\|^2 &= \big((T - \lambda I)x,\ (T - \lambda I)x\big) \\
&= (Tx, Tx) - \overline{\lambda}(Tx, x) - \lambda(x, Tx) + \lambda\overline{\lambda}(x, x) \\
&= (Tx, Tx) - (\lambda + \overline{\lambda})(Tx, x) + |\lambda|^2(x, x) \\
&= (Tx, Tx) - 2a(Tx, x) + (a^2 + b^2)(x, x) \\
&= \big((T - aI)x,\ (T - aI)x)\big) + b^2(x, x) \\
&= \|(T - aI)x\|^2 + b^2\|x\|^2 \\
&\geq b^2\|x\|^2
\end{aligned}$$

From this inequality and the closedness of T we get, as in Lemma 12, that $\mathrm{im}(T - \lambda I)$ is closed in $\mathcal{H}$. Since $y \in \ker(T^* - \lambda I)$ iff $T^*y = \lambda y$ iff for all $x \in \mathcal{D}(T)$ we have $(Tx, y) = (x, T^*y) = (x, \lambda y) = \overline{\lambda}(x, y)$ iff $((T - \overline{\lambda}I)x, y) = 0$, we see that

$$\ker(T^* - \lambda I) = \mathrm{im}(T - \overline{\lambda}I)^\perp \tag{3.90}$$

With this in mind, we will show that for small enough $\eta \in \mathbb{C}$, namely $|\eta| < |b|$, we have

$$\dim\ker(T^* - (\lambda + \eta)I) = \dim\ker(T - \lambda I) \tag{3.91}$$

We will prove this in two stages, showing $\leq$, then $\geq$. Towards this end, suppose first that $u \in \mathcal{D}(T^*) \cap \ker(T - (\lambda + \eta)I)$, $\|u\| = 1$, and note that if $u \perp \ker(T^* - \lambda I)$, then $u \in \mathrm{im}(T - \overline{\lambda}I)$, so there is an $x \in \mathcal{D}(T)$ for which $u = (T - \overline{\lambda}I)x$, and consequently

$$\begin{aligned}
0 = (0, x) = \big((T^* - (\lambda + \eta)I)u, x\big) &= \big(u, (T - (\overline{\lambda} + \overline{\eta})I)x\big) \\
&= \big(u, (T - \overline{\lambda}I)x\big) + \eta(u, x) = (u, u) + \eta(u, x)
\end{aligned}$$

or $\|u\|^2 = -\eta(u,x)$. Now, since $b^2\|x\|^2 \le \|(T-\overline{\lambda}I)x\|^2 = \|u\|^2$, we have $\|x\| \le \|u\|/|b|$, so if $|\eta| < |b|$, then we reach a contradiction, because

$$\|u\|^2 = -\eta(u,x) = -\eta\big((T-\overline{\lambda}I)x, x\big) \le |\eta|\big|\big((T-\overline{\lambda}I)x, x\big)\big|$$

$$|\eta|\|(T-\overline{\lambda}I)x\|\|x\| = |\eta|\|u\|\|x\| < |b|\|u\|\|x\|$$

which means $\|u\| < |b|\|x\|$, though by supposing $u \perp \ker(T^* - \lambda I)$ we had concluded $\|x\| \le \|u\|/|b|$. Thus, for $u \in \mathcal{D}(T^*) \cap \ker(T-(\lambda+\eta)I)$, $\|u\| = 1$, if $|\eta| < |b|$, we cannot have $u \perp \ker(T^* - \lambda I)$, i.e. the space $\ker(T^* - (\lambda+\eta)I)$ cannot "stick out" of $\ker(T^* - \lambda I)$. We conclude, then, that

$$\dim\ker(T^* - (\lambda+\eta)I) \le \dim\ker(T^* - \lambda I)$$

This is a general sort of conclusion whose proof is as follows: if M and N are closed subspaces of a separable Hilbert space $\mathcal{H}$, then $\dim M > \dim N$ implies that there is a unit vector $u \in M \cap N^\perp$. For since M and N are closed, so is $M \cap N$, and $M = (M\cap N) \bigoplus (M \cap N)^\perp$, where the $\perp$ is with respect to the restriction of the inner product of $\mathcal{H}$ to M. But $(M\cap N)^\perp$ is nontrivial, because $\dim M > \dim N \ge \dim(M \cap N)$, which completes the proof of this basic fact. The contrapositive of this statement is, if there is no unit vector $u \in N^\perp$, then $\dim M \le \dim N$, which is what we concluded above for $M = \ker(T^* - (\lambda+\eta)I)$ and $N = \ker(T^* - \lambda I)$. The other inequality follows as above by supposing $|\eta| < |b|/2$, and this shows the equality (3.91). This shows that the dimension of $\ker(T^* - \lambda I)$ is locally constant, and therefore constant in $\mathbb{H}^+$. The same proof applies to $\mathbb{H}^-$, though the constants $\dim\ker(T-\lambda I)$ may differ for $\mathbb{H}^+$ and $\mathbb{H}^-$.

(2) From the inequality $|b|\|x\| \le \|(T-\lambda I)x\|$, $b \ne 0$, proved above we must have that $T - \lambda I$ is left-invertible with bounded inverse, and from the equality $\ker(T^* - \lambda I) = \operatorname{im}(T - \overline{\lambda}I)^\perp$ we can see that this inverse is everywhere defined iff $\dim\ker(T^* - \overline{\lambda}I) = 0$. We therefore have that

$$\lambda \in \rho(T),\ \operatorname{Im}\lambda \ne 0 \quad\Longleftrightarrow\quad \dim\ker(T^* - \overline{\lambda}I) = 0$$

and we see that therefore $\mathbb{H}^\pm \subseteq \sigma(T)$ or $\mathbb{H}^\pm \subseteq \rho(T)$. Since $\rho(T)$ is open and $\sigma(T)$ is closed, we must have $\sigma(T) = \overline{\mathbb{H}^+}$, $\overline{\mathbb{H}^-}$, or $\overline{\mathbb{H}^+} \sqcup \overline{\mathbb{H}^-} = \mathbb{C}$, or the only remaining alternative, a closed subset of $\mathbb{R}$.

In view of (1) and (2), (3) and (4) are restatements of Theorem 31. ■

Corollary 18 *Let $T \subseteq T^*$. If $\rho(T) \cap \mathbb{R} = \varnothing$, then $T = T^*$.*

Proof: Since $\rho(T)$ is open in $\mathbb{C}$, if there is an $x \in \rho(T) \cap \mathbb{R}$ then we know that $\rho(T) \cap \mathbb{H}^{\pm} \neq \varnothing$, so $\sigma(T) \neq \mathbb{H}^{\pm}$, and therefore $\sigma(T) \subseteq \mathbb{R}$, and $T = T^*$. ■

The following theorem includes Proposition 66 as a special case. We follow the proof in Reed and Simon [90, p. 138].

Theorem 34 *Let $T \subseteq T^*$ be a closed symmetric operator on a Hilbert space $\mathcal{H}$. Then,*

(1) *The closed symmetric extensions of T,*

$$T \subseteq S \subseteq S^* \subseteq T^*$$

are the restrictions of T^* to T-closed, T-symmetric subspaces of $\mathcal{D}(T^*)$,

$$\begin{tikzcd}
T \arrow[d, leftrightarrow] & \subseteq \cdots \subseteq & S \arrow[d, leftrightarrow] & \subseteq \cdots \subseteq & S^* \arrow[d, leftrightarrow] & \subseteq \cdots \subseteq & T^* \arrow[d, leftrightarrow] \\
\mathcal{D}(T) & \subseteq \cdots \subseteq & \mathcal{D}(S) & \subseteq \cdots \subseteq & \mathcal{D}(S^*) & \subseteq \cdots \subseteq & \mathcal{D}(T^*)
\end{tikzcd}$$

(2) *$\mathcal{D}(T)$, $\ker(T^* - iI)$ and $\ker(T^* + iI)$ are T-closed, mutually T-orthogonal subspaces of $\mathcal{D}(T^*)$, and*

$$\mathcal{D}(T^*) = \mathcal{D}(T) \oplus_T \ker(T^* - iI) \oplus_T \ker(T^* + iI)$$

(3) *There is a one-to-one correspondence between T-closed, T-symmetric subspaces Y of $\mathcal{D}(T^*)$ containing $\mathcal{D}(T)$ and the T-closed, T-symmetric subspaces Y_1 of the direct sum of the deficiency subspaces $\ker(T^* - iI) \oplus_T \ker(T^* + iI)$, namely*

$$\mathcal{D}(T) \subseteq Y \subseteq \mathcal{D}(T^*) \quad \longleftrightarrow \quad Y = \mathcal{D}(T) \oplus_T Y_1$$

Proof: (1) Recall Definition 50 for the terms T-closed and T-symmetric. Because $T \subseteq S \subseteq S^* \subseteq T^*$ (which is a basic property of the adjoint: $T \subseteq S$ iff $S^* \subseteq T^*$, every symmetric extension of T is contained in T^*, and this extension is closed iff its domain is T-closed and this extension is symmetric iff its domain is T-symmetric. For S is closed iff $\Gamma(S)$ is closed in $\mathcal{H} \oplus \mathcal{H}$ iff $\mathcal{D}(S)$ is closed under the graph inner product $(\cdot,\cdot)_T := (\cdot,\cdot) + (T^*\cdot, T^*\cdot)$, since $S \subseteq T^*$, and similarly S is symmetric iff $S \subseteq S^*$ iff for all $x, y \in \mathcal{D}(S)$ we have $(Sx, y) = (x, Sy)$ iff $[x, y]_T = (T^*x, y) - (x, T^*y) = (Sx, y) - (x, Sy) = 0$.

(2) Since T is closed, $\mathcal{D}(T)$ is T-closed, and $\ker(T^* \pm iI)$ are closed since they are closed in the weaker topology of $\mathcal{H}$. The fact that the three subspaces are T-orthogonal is a straightforward calculation: if $x \in \mathcal{D}(T)$ and $y \in \ker(T \pm iI)$, then

$$(x, y)_T = (x, y) + (T^*x, T^*y) = (x, y) + (Tx, \pm iy) = (x, y) + (x, \pm iT^*y) = (x, y) - (x, y) = 0$$

while if $x \in \ker(T - iI)$ and $y \in \ker(T + iI)$, then

$$(x, y)_T = (x, y) + (T^*x, T^*y) = (x, y) + (ix, -iy) = (x, y) - (x, y) = 0$$

To show that $\mathcal{D}(T^*) = \mathcal{D}(T) \oplus_T \ker(T^* - iI) \oplus_T \ker(T^* + iI)$ we need only show the inclusion $\subseteq$, since the other inclusion is obvious. Suppose, then, that $x \in \mathcal{D}(T^*)$. If $x \perp_T \mathcal{D}(T) \oplus_T \ker(T^* - iI) \oplus_T \ker(T^* + iI)$, then in particular for each $y \in \mathcal{D}(T)$ we have $0 = (x, y)_T = (x, y) + (T^*x, T^*y) = (x, y) + (T^*x, Ty)$, so that

$$(y, x) = -(Ty, T^*x) = (Ty, -T^*x)$$

and therefore $T^*x \in \mathcal{D}(T^*)$. Moreover, $T^*T^*x = -x$, and we find that

$$(T^* + iI)(T^* - iI)x = (T^*T^* + I)x = -x + x = 0$$

so we conclude that $(T^* - iI)x \in \ker(T^* + iI)$. Suppose now that $y \in \ker(T^* + iI)$. Then, $T^*y = -iy$, so that

$$i\big((T^* - iI)x, y\big) = (T^*x, -iy) + (x, y) = i(T^*x, T^*y) + (x, y) = (x, y)_T = 0$$

Thus, $(T^* - iI)x \in \ker(T^* + iI)$ and $(T^* - iI)x \perp \ker(T^* + iI)$, which means $(T^* - iI)x = 0$, or $x \in \ker(T^* - iI)$. Since $x \perp \ker(T^* - iI)$, we must have $x = 0$. Thus, $\mathcal{D}(T^*)$ is not contained in the T-orthogonal complement of $\mathcal{D}(T) \oplus_T \ker(T^* - iI) \oplus_T \ker(T^* + iI)$, so must be contained inside it.

(3) Let Y_1 be a T-closed, T-symmetric subspace of $\ker(T^* - iI) \oplus_T \ker(T^* + iI)$, and let $x, y \in \mathcal{D}(T) \oplus Y_1$. Writing $x = x_0 + x_1$ and $y = y_0 + y_1$, where $x_0, y_0 \in \mathcal{D}(T)$ and $x_1, y_1 \in Y_1$,

we have $[x_0, y_0]_T = 0$ since T is symmetric, and likewise $[x_0, y_1]_T = (T^* x_0, y_1) - (x_0, T^* y_1) = 0$ and $[x_1, y_0]_T = 0$. Thus, $[x, y]_T = 0$, so $Y = \mathcal{D}(T) \oplus_T Y_1$ is T-symmetric and T-closed since $\mathcal{D}(T)$ and Y_1 are T-closed and T-orthogonal. Conversely, let Y be a T-closed, T-symmetric subspace of $\mathcal{D}(T^*)$, and define $Y_1 := Y \cap (\ker(T^* - iI) \oplus_T \ker(T^* + iI))$. Then clearly $Y = \mathcal{D}(T) \oplus_T Y_1$. ■

We are now nearing our main theorem for this section. In fact, that theorem is a corollary of the following theorem, which contains the key ingredients. Recall, however, the following definition.

Definition 53 An **isometry** of a Hilbert space $\mathcal{H}$ is a (necessarily bounded) linear operator U on $\mathcal{H}$ satisfying $\|Ux\| = \|x\|$ for all $x \in \mathcal{H}$. An isometry is clearly bijective, and thereofre unitary. A **partial isometry** is an operator U which is not necessarily bijective, but is an isometry when restricted to $(\ker U)^\perp$. If U is a partial isometry, then writing $\mathcal{H} = \ker U \bigoplus (\ker U)^\perp$, we see that U is a unitary operator from $(\ker U)^\perp$ to $\operatorname{im} U$, which we shall denote by

$$U \in \mathrm{U}\big((\ker U)^\perp, \operatorname{im} U\big)$$

or

$$U \in \mathcal{PU}(\mathcal{H})$$

Here, $(\ker U)^\perp$ is called the **initial subspace** of U and $\operatorname{im} U$ is called the **final subspace** of U. The adjoint U^*, which is also the inverse of U, is then a partial isometry with initial subspace $\operatorname{im} U$ and final subspace $(\ker U)^\perp$,

$$U^* \in \mathrm{U}\big(\operatorname{im} U, (\ker U)^\perp\big)$$ ■

Theorem 35 *Let $T \subseteq T^*$ be a closed symmetric operator on a Hilbert space $\mathcal{H}$. Then, the closed symmetric extensions of T are in one-to-one correspondence with the partial isometries of* $\ker(T^* - iI)$ *into* $\ker(T^* + iI)$,

$$\{S \supseteq T \mid S \subseteq S^* \textit{ and } S \textit{ is closed}\} \cong \mathrm{U}\big(\ker(T^* - iI),\ \ker(T^* + iI)\big)$$

Moreover, if $U \in \mathrm{U}\big(\ker(T^* - iI),\ \ker(T^* + iI)\big)$*, write* $(\ker U)^\perp$ *for the initial space of* U *in* $\ker(T^* - iI)$*, and write* T_U *for the associated closed symmetric extension of* T*. Then,* T_U *has domain*

$$\begin{aligned}\mathcal{D}(T_U) &= \mathcal{D}(T) \oplus_T (\ker U)^\perp \oplus_T \operatorname{im} U \\ &= \{x + x_+ + U(x_+) \mid x \in \mathcal{D}(T),\ x_+ \in (\ker U)^\perp\}\end{aligned}$$

and

$$T_U(x + x_+ + U(x_+)) = Tx + ix - iUx_+ \tag{3.92}$$

If $\dim(\ker U)^\perp < \infty$*, then the deficiency indices of* T_U *are related to those of* T *by*

$$n_\pm(T_U) = n_\pm(T) - \dim(\ker U)^\perp \tag{3.93}$$

Proof: If $S_1 \subseteq S_1^*$ is a closed symmetric extension of T, then by (3) of Theorem 34 we know that $\mathcal{D}(S_1^*) = \mathcal{D}(T) \oplus_T Y_1$ for some T-closed, T-symmetric subspace Y_1 of $\ker(T^* - iI) \oplus_T \ker(T^* + iI)$. Thus, any $x \in \mathcal{D}(S_1^*)$ can be written uniquely as $x = x_0 + x_+ + x_-$ with $x_0 \in \mathcal{D}(T)$, $x_+ + x_- \in Y_1$, and $x_+ \in \ker(T^* - iI)$, $x_- \in \ker(T^* + iI)$, and since Y_1 is T-symmetric, we have

$$\begin{aligned} 0 &= [x_+ + x_-, x_+ + x_-]_T \\ &= (T^*(x_+ + x_-, x_+ + x_-) - (x_+ + x_-, T^*(x_+ + x_-)) \\ &= i(x_+ - x_-, x_+ + x_-) + i(x_+ + x_-, x_+ - x_-) \\ &= 2i(x_+, x_+) - 2i(x_-, x_-) \\ &= 2i\|x_+\|^2 - 2i\|x_-\|^2 \end{aligned}$$

so $\|x_+\| = \|x_-\|$. Define the map

$$\begin{aligned} &U : Y_1 \cap \ker(T^* - iI) \to Y_1 \cap \ker(T^* + iI) \\ &U(x_+) := x_- \end{aligned}$$

and note that by the above considerations U is an isometry. If we extend U to all of $\ker(T^* - iI)$, then U is a partial isometry with initial space $Y_1 \cap \ker(T^* - iI) = (\ker U)^\perp$,

$$U \in \mathcal{PU}(\ker(T^* - iI),\ \ker(T^* + iI))$$

and note that

$$\mathcal{D}(S_1) = \{x_0 + x_+ + U(x_-) \mid x_0 \in \mathcal{D}(T),\ x_+ \in (\ker U)^\perp\}$$

and $S_1(x_0+x_++U(x_+)) = S_1(x_0)+ix_+-iU(x_+)$. Conversely, if $U \in \mathcal{PU}(\ker(T^*-iI), \ker(T^*+iI))$ with initial subspace $(\ker U)^\perp$, define S_1 and $\mathcal{D}(S_1)$ by (3.92) and (3.93), respectively, and note that $\mathcal{D}(S_1)$ is a T-closed, T-symmetric subspace of $\mathcal{D}(T^*)$, so by the previous theorem S_1 is a closed, symmetric extension of T.

Finally, to see the statement about the deficiency indices, note that

$$\begin{aligned} n_+(T) = \dim(\ker -iI) = \dim\ker U + \dim\operatorname{im} U = \dim\ker U + \dim(\ker U)^\perp \\ = \dim\ker(S_1 - iI) + \dim(\ker U)^\perp = n_+(S_1) + \dim(\ker U)^\perp \end{aligned}$$

and similarly with $n_-(T)$. ■

Corollary 19 *Let $T \subseteq T^*$ be a closed symmetric operator on a Hilbert space $\mathcal{H}$. If $n_\pm = \dim\ker(T \mp iI)$ are the deficiency indices of T, then*

(1) *$T = T^*$ iff $n_\pm = 0$.*

(2) *T has self-adjoint extensions iff $n_+ = n_-$, and these extensions are in one-to-one correspondence between self-adjoint extensions of T and unitary maps from $\ker(T - iI)$ to $\ker(T + iI)$,*

$$\{S \supseteq T \mid S = S^*\} \cong \mathrm{U}\big(\ker(T^* - iI),\ \ker(T^* + iI)\big)$$

(3) *If either $n_+ = 0 \neq n_-$ or $n_- = 0 \neq n_+$, then T is* **maximal symmetric**, *i.e. has no nontrivial symmetric extensions.* ■

Proof: (1) If $T = T^*$, then T^* is the only symmetric extension of T, and corresponding to it is the trivial partial isometry between $\ker(T^* - iI)$ and $\ker(T^* + iI)$. The corresponding subspace Y_1 of $\ker(T^* - iI) \oplus_T \ker(T^* + iI)$ is therefore trivial, so we clonclude that $\ker(T^* \pm iI) = \{0\}$, and therefore $n_\pm(T) = 0$.

(2) If S is a self-adjoint extension of T, then by (1) we must have $n_\pm(S) = 0$, so $n_\pm(T) = \dim\ker(U_S)^\perp$ for the corresponding partial isometry U_S, and conversely. Since this is the case

for all S, we see that $\dim(\ker U_S)^\perp$ is the same for all self-adjoint extensions S of T, and since $\dim\ker(S^* \pm iI) = 0$ because $S = S^*$, we must have $(\ker U_S)^\perp = \ker(T^* - iI)$, and all such U_S are in fact unitary, $U_S \in \mathrm{U}(\ker(T^* - iI),\, \ker(T^* + iI))$.

(3) If $n_+(T) = 0$ and $n_-(T) \neq 0$, then by (1) and (2) we see that the associated partial isometry U to a symmetric extensions S of T is the zero map, and cannot therefore be a unitary map, so cannot correspond to a self-adjoint operator. Thus, S is not self-adjoint. In fact, S itself must equal T, for otherwise we would reach a contradiction, with U having nontrivial kernel and a trivial kernel at the same time. Similarly, if $n_+(T) \neq 0$ and $n_-(T) = 0$, then the corresponding partial isometry must be the zero map and satisfy $\ker U = \ker(T^* - iI)$, which would make $S = T$. Thus, T is maximal symmetric. ■

Corollary 20 *A symmetric operator T is essentially self-adjoint iff T has exactly one self-adjoint extension, its closure $\overline{T}$. Moreover, a self-adjoint operator T is maximally symmetric, in that it has no proper symmetric extension.*

Proof: Suppose S is a self-adjoint extension of T, that is $T \subseteq S = S*$. Then S is closed and $S \supseteq \overline{T} = T^{**}$ because $\overline{T}$ is the smallest closed extension of T. On the other hand, since T is closable we have $(T^{**})^* = (\overline{T})^* = T*$, so taking adjoints of both sides of $T \subseteq S$ and of $T^{**} \subseteq T^*$, we have $S = S^* \subseteq T^* \subseteq (T^{**})^* = T^{**}$, so $S = T^{**}$. The converse follows from the previous corollary. ■

3.2.2 The Position and Momentum Operators

In this section we give two important applications of the above theorems, by showing that the position and momentum operators,

$$
\begin{aligned}
(Q_j f)(\mathbf{x}) &= x_j f(\mathbf{x}) \\
(P_j f)(\mathbf{x}) &= -i\frac{\partial f}{\partial x_j}(\mathbf{x}) \qquad \text{(weak derivative)}
\end{aligned}
$$

are essentially self-adjoint on the Hilbert space $L^2(\mathbb{R}^n)$ if defined initially on the dense subspace $\mathcal{D}(\mathbb{R}^n)$ (the space of test functions $C_c^\infty(\mathbb{R}^n)$ equipped with the inductive limit topology) or $\mathcal{S}(\mathbb{R}^n)$.

These results are intrinsically important, of course, but the significance for us is that P_j and Q_j are the building blocks of our symplectic Dirac operators D and $\tilde{D}$ considered below, at least locally. The difficulty with the local description of D and $\tilde{D}$ is that, while P_j and Q_j are individually essentially self-adjoint, compositions and linear combinations of these operators are not necessarily self-adjoint. The best illustration of this fact is the Harmonic oscillator $H_0 f = \frac{1}{2}(\Delta - \mathbf{x} \cdot \mathbf{x})f$, which is a linear combination of compositions of the position and momentum operators, and turns out to be essentially self-adjoint, but the proof of this requires completely different means. This is similar to our situation, as we will see below. It is at least clearl that not every linear combination of various compositions of the position and momentum operators is essentially self-adjoint; this is simply the observation that, when dealing with unbounded operators like P_j and Q_j, self-adjoint operators do not form an algebra.

In the process of proving the essential self-adjointness of P_j and Q_j, moreover, we will give the domains on which they are actually self-adjoint. Our treatment follows that of Moretti [82, pp. 219-224].

Remark 55 Note that these operators differ from those defined above in (4.29)-(4.30) by a factor of $-i$. This factor will have the effect of making those earlier operators skew-adjoint instead of self-adjoint. All of this will be explained clearly in what follows. ■

Theorem 36 (Self-Adjointness of the Position Operator) *For $j = 1, \ldots, n$ let Q_j be the position operator, the (unbounded) multiplication operator on $L^2(\mathbb{R}^n)$ given by*

$$(Q_j f)(\mathbf{x}) := x_j f(\mathbf{x})$$

with domain

$$\begin{aligned} D(Q_j) &:= \{f \in L^2(\mathbb{R}^n) \mid Q_j f \in L^2(\mathbb{R}^n)\} \\ &= \{f \in L^2(\mathbb{R}^n) \mid \int_{\mathbb{R}^n} |x_j f(\mathbf{x})|^2 \, d\mu < \infty\} \end{aligned}$$

Then the Q_j are all self-adjoint. Moreover, $\mathcal{D}(\mathbb{R}^n)$ and $\mathcal{S}(\mathbb{R}^n)$ are cores for the Q_j (so that if we *had defined them initially on these spaces the Q_j' would be essentially self-adjoint, with self-adjoint extensions $\overline{Q_j'} = Q_j$ on $D(Q_j)$ as defined above).*

Proof: Since $\mathcal{D}(\mathbb{R}^n)$ and $\mathcal{S}(\mathbb{R}^n)$ are dense in $L^2(\mathbb{R}^n)$ and contained in $D(Q_j)$, we know that Q_j is densely defined. Moreover, if $f, g \in D(Q_j)$, we have by an integration by parts that

$$\langle Q_j f, g\rangle = \int_{\mathbb{R}^n} x_j f\overline{g}\, d\mu = \int_{\mathbb{R}^n} f\overline{x_j g}\, d\mu = \langle f, Q_j g\rangle$$

so that Q_j is symmetric, $Q_j \subset Q_j^*$. But we can also show that $Q_j^* \subset Q$, since by the definition of the adjoint, $f \in D(Q_j^*)$ iff $\exists h \in L^2(\mathbb{R}^n)$, coinciding with $Q_j^* f$ by definition, such that for all $g \in D(Q_j)$

$$\int_{\mathbb{R}^n} x_j g\overline{f}\, d\mu = \langle Q_j g, f\rangle = \langle g, h\rangle = \int_{\mathbb{R}^n} g\overline{h}\, d\mu$$

i.e.

$$\int_{\mathbb{R}^n} g(\overline{x_j f - h})\, d\mu = 0, \qquad \forall g \in D(Q_j)$$

Since $D(Q_j^*)$ is dense in $L^2(\mathbb{R}^n)$, however, we have that $f \in D(Q_j^*)$ iff $x_j f(\mathbf{x}) = h(\mathbf{x})$ a.e. on $\mathbb{R}^n$, i.e. $h = Q_j f$ in $L^2(\mathbb{R}^n)$. Thus,

$$D(Q_j^*) = \{f \in L^2(\mathbb{R}^n) \mid Q_j f \in L^2(\mathbb{R}^n)\} = D(Q_j)$$

which shows that $Q_j = Q_j^*$.

Restricting Q_j to $\mathcal{D}(\mathbb{R}^n)$ or $\mathcal{S}(\mathbb{R}^n)$ (or defining them on these spaces initially) we still have that $Q_j \subset Q_j^*$, but we no longer have equality. We do, however, have $\ker(Q_j^* \pm iI) = \{0\}$, for if $x_j f(\mathbf{x}) = (Q_j^* f)(\mathbf{x}) = \pm i f(\mathbf{x})$ on $\mathbb{R}^n$ for some $f \in D(Q_j^*)$, then clearly we must have $f = 0$ a.e. on $\mathbb{R}^n$, or $f = 0 \in L^2(\mathbb{R}^n)$, and so $f = 0 \in D(Q_j^*)$. Thus, we know that Q_j is essentially self-adjoint in this case, and $\mathcal{D}(\mathbb{R}^n)$ and $\mathcal{S}(\mathbb{R}^n)$ are cores for Q_j. ∎

Lemma 13 *Let $h \in L^1_{loc}(\mathbb{R}^n)$. Then its weak partial derivative $\partial h/\partial x_j$ satisfies*

$$\frac{\partial h}{\partial x_j} = 0 \quad \Longleftrightarrow \quad h \text{ is a.e. on } \mathbb{R}^n \text{ constant in } x_j$$

Proof: Without loss of generality we may assume $j = 1$, and we write $\mathbf{x} = (x, \mathbf{y})$ for $x = x_1 \in \mathbb{R}$ and $\mathbf{y} = (x_2, \ldots, x_n) \in \mathbb{R}^{n-1}$. By definition of the weak derivative, we have for all $g \in \mathcal{D}(\mathbb{R}^n)$ that $\langle h, \overline{\partial_x g}\rangle =: \langle \partial_x h, \overline{g}\rangle = \langle 0, \overline{g}\rangle = 0$, i.e.

$$\int_{\mathbb{R}^n} h(x, \mathbf{y}) \partial_x g(x, \mathbf{y}) \, dx \otimes d\mathbf{y} = 0 \tag{3.94}$$

Choose $f \in \mathcal{D}(\mathbb{R}^n)$ and $a > 0$ so large that $\operatorname{supp}(f) \subseteq [-a, a]^n$ and a bump function $\xi \in \mathcal{D}(\mathbb{R})$ with $\operatorname{supp}(\xi) \subseteq [-a, a]$ and $\int_{\mathbb{R}} \xi \, d\mu = 1$, and note that the map

$$\begin{aligned} & g \in \mathcal{D}(\mathbb{R}^n) \\ & g(x, \mathbf{y}) := \int_{-\infty}^{x} f(u, \mathbf{y}) \, du - \left(\int_{-\infty}^{x} \xi(\nu) \, d\nu \right) \int_{\mathbb{R}} f(u, \mathbf{y}) \, du \end{aligned}$$

satisfies

$$\frac{\partial g}{\partial x} = f(x, \mathbf{y}) - \xi(x) \int_{\mathbb{R}} f(u, \mathbf{y}) \, du$$

Clearly $\operatorname{supp}(g) \subseteq [-a, a]^n$, too. Therefore, using this g in (3.94) we get, after an application of Fubini-Tonelli and a relabeling of variables,

$$\begin{aligned} 0 &= \int_{\mathbb{R}^n} h(x, \mathbf{y}) \partial_x g(x, \mathbf{y}) \, dx \otimes d\mathbf{y} \\ &= \int_{\mathbb{R}^n} h(x, \mathbf{y}) \left(f(x, \mathbf{y}) - \xi(x) \int_{\mathbb{R}} f(u, \mathbf{y}) \, du \right) dx \otimes d\mathbf{y} \\ &= \int_{\mathbb{R}^n} h(x, \mathbf{y}) f(x, \mathbf{y}) \, dx \otimes d\mathbf{y} - \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}} h(x, \mathbf{y}) \xi(x) \, dx \right) f(u, \mathbf{y}) \, du \otimes d\mathbf{y} \\ &= \int_{\mathbb{R}^n} \left\{ h(x, \mathbf{y}) - \underbrace{\left(\int_{\mathbb{R}} h(u, \mathbf{y}) \xi(u) \, du \right)}_{\in L^1_{loc}(\mathbb{R}^n)} \right\} f(x, \mathbf{y}) \, dx \otimes d\mathbf{y} \end{aligned}$$

Since $f \in \mathcal{D}(\mathbb{R}^n)$ was arbitrary, we conclude that $h(x, \mathbf{y}) - \int_{\mathbb{R}} h(u, \mathbf{y}) \xi(u) \, du = 0$ a.e. on $\mathbb{R}^n$, and therefore

$$h(x, \mathbf{y}) = k(\mathbf{y}) := \int_{\mathbb{R}} h(u, \mathbf{y}) \xi(u) \, du$$

a.e. on $\mathbb{R}^n$. ■

Theorem 37 (Self-Adjointness of the Momentum Operator) *For $j = 1, \ldots, n$ let P_j be the momentum operator, the partial differential operator on $L^2(\mathbb{R}^n)$ given by*

$$(P_j f)(\mathbf{x}) := -i\frac{\partial f}{\partial x_j}(\mathbf{x}) \qquad (\textit{weak erivative})$$

with domain

$$D(P_j) := \{f \in L^2(\mathbb{R}^n) \mid \partial f/\partial x_j \in L^2(\mathbb{R}^n)\}$$

Then the P_j are all self-adjoint. Moreover, $\mathcal{D}(\mathbb{R}^n)$ and $\mathcal{S}(\mathbb{R}^n)$ are cores for the P_j (so that if we had defined them initially on these spaces the P_j would be essentially self-adjoint, with self-adjoint extensions $\overline{P_j}$ on $D(Q_j)$ as defined above).

Proof: We will first prove that on $\mathcal{D}(\mathbb{R}^n)$ we have $\ker(P_j^* \pm iI) = \{0\}$, which will show that each P_j essentialy self-adjoint. Now,

$$\ker(P_j^* \pm iI) = \{f \in L^2(\mathbb{R}^n) \mid P_j^* f \pm if = 0\} = \{f \in L^2(\mathbb{R}^n) \mid \partial_j f \pm f = 0 \ \text{ weakly}\}$$

since the weak derivative $\partial_j f$ is by definition a function $h \in L^1_{loc}(\mathbb{R}^n)$ such that

$$\int_{\mathbb{R}^n} hg \, d\mu = -\int_{\mathbb{R}^n} f \partial_j g \, d\mu$$

or all $g \in \mathcal{D}(\mathbb{R}^n)$. Multiplying any $f \in \ker(P_j^* \pm iI)$ by an exponential $e^{\pm x_j}$ we get

$$\partial_j(e^{\pm x_j} f) = \pm e^{\pm x_j} f + e^{\pm x_j} \partial_j f = \pm e^{\pm x_j} f \mp e^{\pm x_j} f = 0$$

so by the above lemma we must have $f \in \ker(P_j^* \pm iI)$ iff $e^{\pm x_j} f$ is a.e. a constant function in x_j, and therefore must be a function of the form $f = e^{\mp x_j} h$ where h does not depend on x_j. Therefore, by Fubini-Tonelli we have

$$\|f\|^2_{L^2(\mathbb{R}^n)} = \|h\|^2_{L^2(\mathbb{R}^{n-1})} \int_{\mathbb{R}} e^{\mp 2x_j} \, dx_j$$

and since $e^{\mp 2x_j} \notin L^2(\mathbb{R})$ but $f \in L^2(\mathbb{R}^n)$, we must have $\|h\|_{L^2(\mathbb{R}^{n-1})} = 0$ and therefore $\|f\|_{L^2(\mathbb{R}^n)} = 0$, and so $f = 0 \in L^2(\mathbb{R}^n)$. This shows that $\ker(P_j^* \pm iI) = \{0\}$.

A similar argument shows that $\ker(P_j^* \pm iI) = \{0\}$ if instead we first define P_j on $\mathcal{S}(\mathbb{R}^n)$. The domain $D(P_j)$ given in the statement of the theorem is precisely the domain of P_j^*, however, as we have demonstrated in the course of this proof, and so on $D(P_j) \supset \mathcal{S}(\mathbb{R}^n) \supset \mathcal{D}(\mathbb{R}^n)$ we have that $P_j = P_j^*$, completing the proof. ■

3.3 Elliptic Differential Operators

In this section we describe an important class of linear differential operators, the elliptic operators. Let E and F be complex vector bundles over a smooth manifold M. If $\Gamma_c(E)$ is the space of compactly supported sections of E and $\Gamma_c(F)$ the space of compactly supported sections of F, then the Hermitian strucures on E and F give a natural L^2 inner product on these spaces, by $(\sigma, \tau) := \int_M h_E(\sigma, \tau)\, dV$ on $\Gamma_c(E)$ for example. An elliptic differential operator $P : L^2(E) \to L^2(F)$ is an unbounded operator acting on the L^2-closure of $\Gamma_c(E)$ with fiberwise invertible symbol $\sigma(P)(x, \xi) \in \mathrm{GL}(E_x, F_x)$ for nonzero $\xi \in T_x^*M$. Let $\mathsf{Ell}^k(E, F)$ denote the space of all kth order elliptic operators from E to F.

Let us list the extremely nice properties satisfied by elliptic operators:

(1) Elliptic operators always satisfy a particularly nice condition, which is local in nature, called elliptic regularity. This says that if $u \in L^2(E)$ satisfies an equation $Pu = v$ weakly and v is known to be smooth, then u must be smooth. In particular, $\ker P \subseteq \Gamma(E)$.

(2) If M is compact, any $P \in \mathsf{Ell}^k(E, F)$ is Fredholm, so the index of P

$$\operatorname{index} P := \dim \ker P - \dim \operatorname{coker} P$$

is well-defined. Once we put a topology on the space $\mathsf{Ell}^k(E, F)$, the index extends to a continuous map

$$\operatorname{index} : \mathsf{Ell}^k(E, F) \to \mathbb{Z}$$

(3) If M is compact and $P \in \mathsf{Ell}^k(E, F)$, the spaces $L^2(E)$ and $L^2(F)$ acquire orthogonal decompositions via

$$L^2(E) = \ker P \bigoplus \operatorname{im} P^* \qquad \text{and} \qquad L^2(F) = \ker P^* \bigoplus \operatorname{im} P$$

which is a kind of analog of the Hodge decomposition.

(4) If M is compact and $P \in \mathsf{Ell}^k(E, E)$ is symmetric (on the core $\Gamma_c(E) = \Gamma(E)$, which we usually take to be its domain), then P is essentially self-adjoint with self-adjoint exten-

sion having domain $H^k(E)$. Elliptic self-adjoint operators over compact manifolds have particularly nice spectral properties. Let $\sigma(P)$ denote the spectrum of P. Then,

(a) The spectrum $\sigma(P)$ of P is a closed, discrete, countable and unbounded subsets of $\mathbb{R}$.

(b) $\sigma(P)$ equals the point spectrum $\sigma_p(P)$ of P, that is $\sigma(P)$ contains only eigenvalues of P, each eigenvalue having finite geometric multiplicity, that is $\dim\ker(P-\lambda I)<\infty$ for each $\lambda\in\sigma(P)$.

(c) The finite-dimensional, hence closed, eigenspaces $\ker(P-\lambda I)$ form an orthogonal decomposition of $L^2(E)$,

$$L^2(E)=\bigoplus_{\lambda\in\sigma(P)}\ker(P-\lambda I)$$

so that we have a particularly simple resolution of the identity

$$\mathrm{id}_E=\sum_{\lambda\in\sigma(P)}P_\lambda$$

where $P_\lambda : L^2(E)\to\ker(P-\lambda I)$ is the orthogonal projection onto the closed λ-eigenspace. Therefore any L^2-section σ of E may be uniquely expressed as

$$\sigma=\sum_{\lambda\in\sigma(P)}\sigma_\lambda,\quad\text{where }\sigma_\lambda:=P_\lambda(\sigma)$$

Moreover, P also neatly decomposes as

$$P=\sum_{\lambda\in\sigma(P)}\lambda P_\lambda$$

Let us begin with elliptic regularity.

Theorem 38 *Let $P\in\mathsf{Ell}^k(E,F)$, thought of as a generalized operator $P:\mathcal{D}(E)\to\mathcal{D}'(F)$. If $Pu=v$ weakly for $u\in L^p_{\mathrm{loc}}(E)$ and $v\in L^p_{\mathrm{loc}}(F)$, $p\in(1,\infty)$, then in fact $u\in W^{k,p}_{\mathrm{loc}}(E)$. More generally, if $v\in W^{s,p}_{\mathrm{loc}}(F)$, then $u\in W^{s+k,p}(E)$, and we have the following local elliptic estimate: for all $0<r<R$,*

$$\|u\|_{W^{s+k,p}(E|_{B_r})}\leq C\big(\|v\|_{W^{s,p}(F|_{B_R})}+\|u\|_{L^p(E|_{B_r})}\big)$$

As a result, if $v\in\mathcal{E}(F)$, i.e. v is smooth, then $u\in\mathcal{E}(E)$, i.e. u is smooth. In particular, $\ker P$ consists of smooth sections.

Proof: Nicolaescu [85], Theorem 10.3.6, Corollaries 10.3.9-10.3.10. ■

Theorem 39 *Let $P \in \mathsf{Ell}^k(E,F)$, thought of as a generalized operator $P : \mathcal{D}(E) \to \mathcal{D}'(F)$. Extending it to $L^2(E) \to L^2(F)$ as a densely defined operator on a Hilbert space, P is Fredholm, so has a well-defined index*

$$\operatorname{index} P := \dim\ker P - \dim\operatorname{coker} P$$

and the index map

$$\operatorname{index} : \mathsf{Ell}^k(E,F) \to \mathbb{Z}, \quad P \mapsto \operatorname{index}(P)$$

is continuous, once we put the following metric topology on $\mathsf{Ell}^k(E,F)$: if $P,Q \in \mathsf{Ell}^k(E,F)$, let

$$\delta(E,F) = \sup_{\|u\|_{H^k_{\text{loc}}(E)}=1} \|Pu - Qu\|_{L^2(F)}$$

and let $d(P,Q) := \max\{\delta(P,Q),\ \delta(P^*,Q^*)\}$, then we have that $(\mathsf{Ell}^k(E,F), d)$ is a metric space. Moreover, the spaces $L^2(E)$ and $L^2(F)$ acquire orthogonal decompositions via

$$L^2(E) = \ker P \bigoplus \operatorname{im} P^* \qquad \textit{and} \qquad L^2(F) = \ker P^* \bigoplus \operatorname{im} P$$

Proof: Nicolaescu [85], Theorem 10.4.7, Corollary 10.4.10, Theorem 10.4.13. ■

Theorem 40 *If M is compact and $P \in \mathsf{Ell}^k(E,E)$ is symmetric (on the core $\Gamma_c(E) = \Gamma(E)$, which we usually take to be its domain), then P is essentially self-adjoint with self-adjoint extension having domain $H^k(E)$. Elliptic self-adjoint operators over compact manifolds have particularly nice spectral properties. Let $\sigma(P)$ denote the spectrum of P. Then,*

(1) *The spectrum $\sigma(P)$ of P is a closed, discrete, countable and unbounded subsets of $\mathbb{R}$.*

(2) *The spectrum $\sigma(P)$ equals the point spectrum $\sigma_p(P)$ of P, that is $\sigma(P)$ contains only eigenvalues of P, each eigenvalue having finite geometric multiplicity, that is $\dim\ker(P - \lambda I) < \infty$ for each $\lambda \in \sigma(P)$.*

(3) *The finite-dimensional, hence closed, eigenspaces $\ker(P - \lambda I)$ form an orthogonal decomposition of $L^2(E)$,*

$$L^2(E) = \bigoplus_{\lambda\in\sigma(P)} \ker(P - \lambda I)$$

so that we have a particularly simple resolution of the identity

$$\mathrm{id}_E = \sum_{\lambda \in \sigma(P)} P_\lambda$$

where $P_\lambda : L^2(E) \to \ker(P - \lambda I)$ *is the orthogonal projection onto the closed* λ*-eigenspace. Therefore any* L^2*-section* σ *of* E *may be uniquely expressed as*

$$\sigma = \sum_{\lambda \in \sigma(P)} \sigma_\lambda, \quad \textit{where } \sigma_\lambda := P_\lambda(\sigma)$$

Moreover, P *also neatly decomposes as*

$$P = \sum_{\lambda \in \sigma(P)} \lambda P_\lambda$$

Proof: Nicolaescu [85], Theorem 10.4.19. ■

Chapter 4

The Metaplectic Representation

4.1 The Weyl Algebra

By analogy with the definition of the Clifford algebra $\mathrm{Cl}(V,q)$ on a quadratic space (V,q), we define the **Weyl algebra**, or **symplectic Clifford algebra**, $\mathrm{Cl}(V,\omega)$ on a symplectic space (V,ω) as the solution to a universal problem, namely as the pair $(\mathrm{Cl}(V,\omega),j)$ consisting of

(1) an associative unital $\mathbb{R}$-algebra $\mathrm{Cl}(V,\omega)$

(2) an $\mathbb{R}$-linear map $j \in \mathcal{L}\big(V,\mathrm{Cl}(V,\omega)\big)$ satisfying

   (a) $j(u)j(v) - j(v)j(u) = -\omega(u,v)\cdot 1_{\mathrm{Cl}(V,\omega)}$

   (b) (*universal property*) If A is any other associative unital $\mathbb{R}$-algebra, and f is any real-linear map $f \in \mathcal{L}(V,A)$ satisfying $f(u)f(v) - f(v)f(u) = -\omega(u,v)\cdot 1_A$, then there is a unique $\mathbb{R}$-algebra homomorphism $F \in \mathrm{Hom}\big(\mathrm{Cl}(V,\omega),A\big)$ making the following diagram commute:

$$\begin{tikzcd} & \mathrm{Cl}(V,\omega) \arrow[dd, "F"] \\ V \arrow[ur, "j"] \arrow[dr, "f"'] & \\ & A \end{tikzcd}$$

Remark 56 As with any object satisfying a universal property, if it exists, the Weyl algebra is *unique up to isomorphism*, since if (C_1,j_1) and (C_2,j_2) are two solutions to the universal problem outlined above, then for the first we can take $A = C_2$ and for the second we can take $A = C_1$, and thus we get unique algebra homomorphisms $F_1 : C_1 \to C_2$ and $F_2 : C_2 \to C_1$, satisfying, respectively, $F_1 \circ j_1 = j_2$ and $F_2 \circ j_2 = j_1$. Combining these we get that $F_2 \circ F_1 : C_1 \to C_1$ and $F_1 \circ F_2 : C_2 \to C_2$ satisfy $F_1 \circ F_2 \circ j_1 = j_1$. But since the universal property applied to (C_1, j_1) and $A = C_1$ is clearly solved uniquely by id_{C_1}, we must have $F_2 \circ F_1 = \mathrm{id}_{C_1}$, and for similar reasons we have $F_1 \circ F_2 = \mathrm{id}_{C_2}$, which shows that $F_2 = F_1^{-1}$ and therefore F_1 is an isomorphism. ∎

Theorem 41 (Existence) *Let $T(V) := \bigoplus_{n=0}^{\infty} T^n(V)$ be the tensor algebra of a symplectic vector space (V, ω), and define the quotient algebra of $T(V)$ by the ideal generated by vectors in $T^1(V)$ of the form $u \otimes v - v \otimes u + \omega(u,v) \cdot 1$*

$$\mathrm{Cl}(V,\omega) := T(V)/(u \otimes v - v \otimes u + \omega(u,v) \cdot 1)$$

where $1 \in T^0(V) \subseteq T(V)$ is the identity element of $T(V)$. Let $i : V \hookrightarrow T(V)$ be the inclusion map, $i(V) = T^1(V)$, and let $\pi : T(V) \to \mathrm{Cl}(V,\omega)$ be the quotient map, and define $j : V \to \mathrm{Cl}(V,\omega)$ to be their composition, $j = \pi \circ i$. Then $\big(\mathrm{Cl}(V,\omega),\ j\big)$ is a Weyl algebra.

Proof: First, note that j satisfies property (2)(a): The identity element of $\mathrm{Cl}(V,\omega)$ is $1_{\mathrm{Cl}(V,\omega)} = j(1_{\mathbb{R}})$, and if we write I for the ideal $(u \otimes v - v \otimes u + \omega(u,v) \cdot 1)$ we have for all $u, v \in V$ that

$$\begin{aligned}
j(u)j(v) - j(v)j(u) &= \pi(i(u))\pi(i(v)) - \pi(i(v))\pi(i(u)) \\
&= \big(i(u) + I\big) \otimes \big(i(v) + I\big) - \big(i(v) + I\big) \otimes \big(i(u) + I\big) \\
&= \big(i(u) \otimes i(v) - i(v) \otimes i(u)\big) + I \\
&= \pi(i(u) \otimes i(v) - i(v) \otimes i(u)) \\
&= \pi(-\omega(u,v)) \\
&= -\omega(u,v)\pi(i(1_{\mathbb{R}}) \\
&= -\omega(u,v)j(1_{\mathbb{R}}) \\
&= -\omega(u,v) \cdot 1_{\mathrm{Cl}(V,\omega)}
\end{aligned}$$

Next, $\big(\mathrm{Cl}(V,\omega),\ j\big)$ satisfies the universal property of (2)(b): Suppoe A is an associative unital $\mathbb{R}$-algebra and $f \in \mathcal{L}(V, A)$ is a linear map satisfying $f(u)f(v) - f(v)f(u) = -\omega(u,v) \cdot 1_A$ for all $u, v \in V$. By the universal property of the tensor algebra $T(V)$, if $t : V \hookrightarrow T(V)$ is the tensor map (the inclusion), then for this A and this f we know that there exists a unique $\mathbb{R}$-algebra homomorphism $F \in \mathrm{Hom}(T(V), A)$ such that the following diagram commutes:

$$\begin{tikzcd} & T(V) \arrow[d, "F"] \\ V \arrow[ur, "t"] \arrow[r, "f"'] & A \end{tikzcd}$$

Consequently, since $f(u)f(v) - f(v)f(u) = -\omega(u,v) \cdot 1_A$ and $F \circ i = f$

$$F\big(i(u) \otimes i(v) - i(v) \otimes i(u) + \omega(u,v) \cdot 1\big) = f(u)f(v) - f(v)f(u) + \omega(u,v) \cdot 1_A = 0$$

which shows that $I \subseteq \ker F$. As a consequence, there is a unique $\mathbb{R}$-algebra homomorphism $\tilde{F} \in \mathrm{Hom}(T(V), A)$ such that the following diagram commutes:

$$\begin{tikzcd} T(V) \arrow[rr, "F"] \arrow[dr, "\pi"'] & & A \\ & T(V)/I \arrow[ur, "\tilde{F}"'] & \end{tikzcd}$$

Indeed, we define $\tilde{F}$ by $\tilde{F}(\tau + I) := F(\tau)$, and the fact that $I \subseteq \ker F$ ensures that $\tilde{F}$ is well defined. Thus, we have

$$\tilde{F} \circ j = \tilde{F} \circ (\pi \circ i) = (\tilde{F} \circ \pi) \circ i = F \circ i = f$$

i.e., since $\mathrm{Cl}(V, \omega) = T(V)/I$, we have shown that the diagram below commutes:

$$\begin{tikzcd} & \mathrm{Cl}(V,\omega) \arrow[d, "\tilde{F}"] \\ V \arrow[ur, "j"] \arrow[r, "f"'] & A \end{tikzcd}$$

■

Notation 7 *In the special case of* $\mathbb{R}^{2n}$ *with the standard symplectic form* ω_0 *we will write* Cl_n *for* $\mathrm{Cl}(\mathbb{R}^{2n}, \omega_0)$. ■

Theorem 42 *Let* (V, ω_V) *and* (W, ω_W) *be symplectic vector spaces, and let* $\mathrm{Cl}(V, \omega_V)$ *and* $\mathrm{Cl}(W, \omega_W)$ *be their respective Weyl algebras. If* $\mathrm{Cl}(V \oplus W, \omega_V \oplus \omega_W)$ *is the Weyl algebra of the direct sum* $(V \oplus W, \omega_V \oplus \omega_W)$*, then we have the following isomorphism of algebras:*

$$\mathrm{Cl}(V \oplus W, \omega_V \oplus \omega_W) \cong \mathrm{Cl}(V, \omega_V) \otimes \mathrm{Cl}(W, \omega_W) \tag{4.1}$$

Proof: Define the map $f : V \oplus W \to \mathrm{Cl}(V, \omega_V) \otimes \mathrm{Cl}(W, \omega_W)$ by

$$f(u + v) := u \otimes 1 + 1 \otimes v$$

Then f satisfies the distinguishing property: for all $(u + v), (u' + v') \in V \oplus W$,

$$\begin{aligned}
f(u+v)f(u'+v') - f(u'+v')f(u+v) &= (u \otimes 1 + 1 \otimes v)(u' \otimes 1 + 1 \otimes v') \\
&\qquad -(u' \otimes 1 + 1 \otimes v')(u \otimes 1 + 1 \otimes v) \\
&= uu' \otimes 1 + u \otimes v' + u' \otimes v + 1 \otimes vv' - u'u \otimes 1 \\
&\qquad -u' \otimes v - u \otimes v' - 1 \otimes v'v \\
&= (uu' - u'u) \otimes 1 + 1 \otimes (vv' - v'v) \\
&= (-\omega_V(u, u')1) \otimes +1 \otimes (-\omega_W(v, v')) \\
&= -(\omega_V(u, u') + \omega_W(v, v'))1 \otimes 1 \\
&= -(\omega_V \oplus \omega_W)((u + v), (u' + v'))1 \otimes 1
\end{aligned}$$

The universal property of the Weyl algebra $\mathrm{Cl}(V \oplus W, \omega_V \oplus \omega_W)$ then implies that there is a unique $\mathbb{R}$-algebra homomrphims $F : \mathrm{Cl}(V \oplus W, \omega_V \oplus \omega_W) \to \mathrm{Cl}(V, \omega_V) \otimes \mathrm{Cl}(W, \omega_W)$ making the following diagram commute:

$$\begin{tikzcd}
 & \mathrm{Cl}(V \oplus W, \omega_V \oplus \omega_W) \arrow[dd, "F"] \\
V \oplus W \arrow[ur, "j"] \arrow[dr, "f"'] & \\
 & \mathrm{Cl}(V, \omega_V) \otimes \mathrm{Cl}(W, \omega_W)
\end{tikzcd}$$

But this map is invertible, for we can construct its inverse $G : \mathrm{Cl}(V, \omega_V) \otimes \mathrm{Cl}(W, \omega_W) \to \mathrm{Cl}(V \oplus W, \omega_V \oplus \omega_W)$. Note first that the embeddings of V and W into $V \oplus W$ induce unique homomorphisms $g : \mathrm{Cl}(V, \omega_V) \to \mathrm{Cl}(V \oplus W, \omega_V \oplus \omega_W)$ and $h : \mathrm{Cl}(W, \omega_W) \to \mathrm{Cl}(V \oplus W, \omega_V \oplus \omega_W)$, satisfying $g(u+0) = u$ and $h(0+v) = v$ for all $u \in V$ and $v \in W$. Since the images of these homomorphisms commute, we obtain a homomorphism homomorphism $G = g \oplus h : \mathrm{Cl}(V, \omega_V) \otimes \mathrm{Cl}(W, \omega_W) \to \mathrm{Cl}(V \oplus W, \omega_V \oplus \omega_W)$. Finally, note that $F \circ G = \mathrm{id}$, for $F(G(u \otimes v)) = F((g \oplus h)(u \otimes v)) = F(u + v) = u \otimes v$. ■

Theorem 43 *Let (V,ω) be a symplectic vector space and let $(q_1,\ldots,q_n,p_1,\ldots,p_n)$ be a symplectic basis for V. Then in the Weyl algebra $\mathrm{Cl}(V,\omega)$ the set of all monomials of the form*

$$q_1^{a_1}q_2^{a_2}\cdots q_n^{a_n}p_1^{b_1}p_2^{b_2}\cdots p_n^{b_n}, \qquad \textit{where} \;\; a_i,b_j\in\mathbb{N}_0$$

forms a basis for $\mathrm{Cl}(V,\omega)$. Consequently, $\mathrm{Cl}(V,\omega)$ is infinite-dimensional.

Proof: By the previous theorem it is enough to prove the statement for the case $n=1$, i.e. $\dim(V)=2$. If (p,q) is the symplectic basis for V, then note first that the set of monomials p^aq^b, for $a,b\in\mathbb{N}_0$, form a spanning set, for, first of all, $\{x_1\otimes\cdots\otimes x_n \mid x_i=p \text{ or } q,\ n\in\mathbb{N}_0\}$ forms a basis for $T(V)=\bigoplus_{n=0}^{\infty}\bigotimes^n V$, and therefore all elements of $\mathrm{Cl}(V,\omega)=T(V)/I$ are words in p and q. But any word in p and q in $\mathrm{Cl}(V,\omega)$ can be reordered to have all ps on the left and all qs on the right, by use of the relation $pq-qp=-\omega(p,q)\cdot 1$, at the cost of introducing extra monomials with fewer letters.

Secondly, this set of monomials is linearly independent. To see this, we represent $\mathrm{Cl}(V,\omega)$ on the real vector space $\mathbb{R}[x,y]$ of polynomials in two variables, namely by letting q act by multiplication by x and letting p act by differentiation with respect to x plus multiplication by y,

$$p(x^ay^b)=x^{a+1}y^b$$

$$q(x^ay^b)=ax^{a-1}y^b+x^ay^{b+1}$$

That is, if $\varphi:\mathrm{Cl}(V,\omega)\to\mathrm{End}_{\mathbb{R}}\,\mathbb{R}[x,y]$ is the representation, then

$$\varphi(q)=x$$

$$\varphi(p)=\frac{d}{dx}+y$$

where x and y here denote multiplication, that is $\varphi(q)$ and $\varphi(p)$ are differential operators of order 0 and 1, respectively, on $\mathbb{R}[x,y]$. Now, the set of monomials x^ay^b for $a,b\in\mathbb{N}_0$ is a basis for $\mathbb{R}[x,y]$, and we note that φ is indeed a representation, for thinking of φ as a map from V to $\mathbb{R}[x,y]$, on these basis vectors φ satisfies

$$\begin{aligned}
(\varphi(q)\circ\varphi(p)-\varphi(p)\circ\varphi(q))x^ay^b &= x\Big(\frac{d}{dx}+y\Big)x^ay^b-\Big(\frac{d}{dx}+y\Big)xx^ay^b \\
&= x(ax^{-1}y^b+x^ay^{b+1})-\Big(\frac{d}{dx}+y\Big)x^{a+1}y^b \\
&= ax^ay^b+x^{a+1}y^{b+1}-(a+1)x^ay^b-x^{a+1}y^{b+1} \\
&= -x^ay^b \\
&= -\omega(p,q)x^ay^b
\end{aligned}$$

i.e. $\varphi(q)\circ\varphi(p)-\varphi(p)\circ\varphi(q)=-\omega(p,q)1$. Therefore, by the universal property there is a unique homomorphism $\Phi:\mathrm{Cl}(V,\omega)\to\mathbb{R}[x,y]$ satisfying $\Phi\circ j=\varphi$, and we can think of *this* map as our representation. In particular, this representation is faithful, as φ sends distinct basis elements to distinct operators and is $\mathbb{R}$-linear. Thus, Φ is injective. As a consequence, since the monomials x^ay^b form a basis for $\mathbb{R}[x,y]$, and therefore linearly independent, if we suppose

$$\sum_{a,b}c_{ab}x^ay^b=0$$

then of necessity $c_{ab}=0$. If, therefore,

$$\sum_{a,b}c_{ab}q^ap^b=0$$

then applying Φ gives $\sum_{a,b}c_{ab}x^a\big(\frac{d}{dx}+y\big)^b=0$ in $\mathrm{End}_{\mathbb{R}}\,\mathbb{R}[x,y]$, so applying it to the identity $1\in\mathbb{R}[x,y]$ (whose x-derivative is 0) gives $0=\sum_{a,b}c_{ab}x^a\big(\frac{d}{dx}+y\big)^b1=\sum_{a,b}c_{ab}x^ay^b$, which means $c_{ab}=0$. ■

Remark 57 This is in contrast to the symmetric case, where the Clifford algebra $\mathrm{Cl}(V,q)$ has finite dimension 2^n. In that case we embedded the spin group $\mathrm{Spin}(V,q)$, the double cover of the special orthogonal group $\mathrm{SO}(V,q)$, into $\mathrm{Cl}(V,q)$ and obtained thereby a faithful and irreducible matrix representation in $M_k(\mathbb{F})$ for some k, where $\mathbb{F}=\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$ depending on the choice of n. This was possible because $\mathrm{Cl}(V,q)$ was finite-dimensional and $\mathrm{Spin}(V,q)\hookrightarrow\mathrm{Cl}(V,q)$ therefore was also finite-dimensional. In the present case we do not have this finite-dimensionality at our disposal, and as it turns out the metaplectic group $\mathrm{Mp}(n,\mathbb{R})$, the analogue of the spin group $\mathrm{Spin}(n)$, does not have any faithful matrix representation. We will represent $\mathrm{Mp}(n,\mathbb{R})$ faithfully and irreducibly on $L^2(\mathbb{R}^n)$, via a unitary representation called the Segal-Shale-Weil representation or metaplectic representation. It happens that the the Lie algebra $\mathfrak{mp}(n,\mathbb{R})$ of the metaplectic group embeds into the Weyl algebra $\mathrm{Cl}(n) = \mathrm{Cl}(\mathbb{R}^{2n},\omega_0)$, which is linearly mapped into $\mathrm{End}(\mathcal{S}(\mathbb{R}^n))$ via an identification of the symplectic basis elements of $\mathbb{R}^{2n}$ with the quantum position and momentum operators. Hence, the groups $\mathrm{Mp}(n,\mathbb{R})$ and $\mathrm{Sp}(n,\mathbb{R})$, as well as their Lie algebras, and the Weyl algebra into which the Lie algebras are embedded, end up in $\mathrm{End}(\mathcal{S}(\mathbb{R}^n))$, if not in $U(L^2(\mathbb{R}^n))$. ■

4.2 The Harmonic Oscillator and the Hermite Functions

In this section we review the Harmonic oscillator and its eigenfunctions, the Hermite functions. The Harmonic oscillator is an unbounded self-adjoint operator, densely defined on $L^2(\mathbb{R}^n)$, whose spectrum *a fortiori* consists of real eigenvalues. The associated eigenspaces are finite dimensional and spanned by the corresponding Hermite functions. The set of all Hermite functions forms a complete orthonormal system for $L^2(\mathbb{R}^n)$, so $L^2(\mathbb{R}^n)$ has an orthogonal decomposition into the finite-dimensional eigenspaces of the Harmonic oscillator. The Hermite functions are also eigenfunctions of the Fourier transform, and since the Fourier transform is a unitary operator on $L^2(\mathbb{R}^n)$, it restricts to a unitary operator on each eigenspace. This will prove to be a very useful fact for us, and we will make good use of these properties of $\mathcal{F}$.

Definition 54 Define the **Harmonic oscillator** (also called the **Hamiltonian** or **Schrödinger operator** in the literature) to be the unbounded operator with initial domain the Schwartz space of rapidly decreasing functions $\mathcal{S}(\mathbb{R}^n)$ (Definition 25),

$$H_0 : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n) \tag{4.2}$$

$$(H_0 f)(\mathbf{x}) := \frac{1}{2}\big(\Delta - \langle \mathbf{x}, \mathbf{x}\rangle\big) f(\mathbf{x}) \tag{4.3}$$

where $\Delta = \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j^2}$ is the standard Laplacian. That is,

$$(H_0 f)(\mathbf{x}) = \frac{1}{2} \sum_{j=1}^{n} \left(\frac{\partial^2 f}{\partial x_j^2}(\mathbf{x}) - x_j^2 f(\mathbf{x}) \right)$$

That is, H_0 is just the Laplacian Δ *plus a potential* $V(\mathbf{x}) = -\langle \mathbf{x}, \mathbf{x} \rangle$, which is easily seen to lie in $L^2_{\text{loc}}(\mathbb{R}^n)$. A simple integration by parts shows that H_0 is symmetric, so that $(H_0 f, g)_{L^2(\mathbb{R}^n)} = (f, H_0 g)_{L^2(\mathbb{R}^n)}$ for all $f, g \in \mathcal{S}(\mathbb{R}^n)$, but it was not until 1951 that Kato [63] proved that it is essentially self-adjoint, which was a major breakthrough in the mathematical formalization of quantum mechanics (see Cycon et al. [26]). ∎

Remark 58 Physically, the Laplacian describes the position of a free particle, in the sense that a solution φ of the simple non-relativistic Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \varphi = -\Delta \varphi$$
$$\varphi(\mathbf{x}, 0) = \varphi_0(\mathbf{x})$$

has an associated probability, $\|\varphi(\cdot, t)\|^2_{L^2(\Omega)} = \int_\Omega |\varphi(\mathbf{x}, t)|^2 d\mathbf{x}$, which is interpreted as the probability of the particle lying in a region $\Omega \subseteq \mathbb{R}^n$ (here, of course, $\varphi(\cdot, t)$ is assumed to be an element of $L^2(\mathbb{R}^n)$ for each $t \in \mathbb{R}$).

Now, Δ is a symmetric operator (via integration by parts), and it is in fact essentially self-adjoint, which can be shown as follows: let $f \in \ker(\Delta^* \pm iI)$. Then, since the Fourier transform $\mathcal{F}$ is a unitary operator on $L^2(\mathbb{R}^n)$ which intertwines the position and momentum operators, x_j and $-i\partial_j$ (Proposition 60), we can apply $\mathcal{F}$ to the weak equality $\Delta^* f = \pm i f$ gives $\langle \mathbf{x}, \mathbf{x} \rangle \hat{f} = i\hat{f}$, which is only possible if $\hat{f} = 0$. Applying $\mathcal{F}^{-1}$ to this equality shows that $f = 0$. Thus, by Corollary 17 above we see that Δ is essentially self-adjoint. In other words, Δ is unitarily equivalent to a multiplication operator, which is known to be essentially self-adjoint. Consequently, we may exponentiate Δ, and all solutions of the Schrödinger equation are given in terms of an initial condition φ_0 by

$$\varphi(\mathbf{x}, t) = e^{-it\Delta} \varphi_0(\mathbf{x})$$

where $e^{-it\Delta}$ is defined via the functional calculus, as $e^{-it\Delta} = \sum_{n=0}^{\infty} e^{-i\lambda_n t} P_n$, with P_n the projection onto the λ_n-eigenspace E_{λ_n}. The fact that this is possible is a result of the self-adjointness and ellipticity of Δ (Theorem 40).

The harmonic oscillator H_0 has the additional term $V \in L^2_{\text{loc}}(\mathbb{R}^n)$ added to Δ. Physically, V is interpreted as an electric potential, and the resulting non-relativistic Schrödinger equation

$$i\hbar \frac{\partial}{\partial t}\varphi = H_0 \varphi$$

$$\varphi(\mathbf{x}, 0) = \varphi_0(\mathbf{x})$$

has solutions which describe the (probable) position of a single particle moving in an electric field. If we wish to solve this equation by the same methods, exponentiating H_0, we need to ensure that H_0 is self-adjoint. Now, H_0 is symmetric, being the sum of the symmetric operators Δ and V, but even though H_0 differs from Δ only by a potential function $V \in L^2_{\text{loc}}(\mathbb{R}^n)$, it is much harder to show that it is essentially self-adjoint—-self-adjoint operators do not necessarily add to self-adjoint operators. This is a feature of the more general fact that self-adjoint operators do not form an algebra. Kato's method of proof of the essential self-adjointness of H_0 relies on a novel method, an inequality called the Δ-boundedness of V, which is the existence of real numbers $a, b \geq 0$ such that $\|Vf\| \leq a\|\Delta f\| + b\|f\|$. This inequality is then combined with von Neumann's basic idea, Corollary 17, to show essential self-adjointness by showing that $\ker(H_0^* \pm iI) = \{0\}$. ■

Definition 55 Let $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$ be a multi-index, and define the **Hermite functions on** $\mathbb{R}^n$ by

$$h_\alpha \in \mathcal{S}(\mathbb{R}^n) \tag{4.4}$$

$$h_\alpha(\mathbf{x}) := h_{\alpha_1}(x_1) h_{\alpha_2}(x_2) \cdots h_{\alpha_n}(x_n) \tag{4.5}$$

where $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and each $h_{\alpha_j} \in \mathcal{S}(\mathbb{R})$ is a **Hermite function on** $\mathbb{R}$,

$$h_{\alpha_j}(x_j) := e^{x_j^2/2} \frac{\partial^{\alpha_j}}{\partial x_j^{\alpha_j}} e^{-x_j^2} \tag{4.6}$$

As we will show below, the Hermite functions form a complete orthonormal system for $L^2(\mathbb{R}^n)$, and serve as eigenfunctions for both the Harmonic oscillator H_0 and the Fourier transform $\mathcal{F}$. ■

To see that the Hermite functions are eigenfunctions of H_0, we need to make some preliminary observations. Let

$$D_j := \frac{\partial}{\partial x_j}$$

$$X_j := x_j \qquad \text{(multiplication by } x_j\text{)}$$

and define

$$Z_j := D_j + X_j$$

$$Z_j^* := D_j - X_j$$

These Z_j and Z_j^* operators are not to be confused with our operators Z and Z^* on symplectic spinors defined below. For simplicity, consider the case $n = 1$ first, so that we may write $Z = D + X$ and $Z^* = D - X$. Let us also define $h_n \equiv 0$ for $n < 0$ so as to simplify matters below.

Proposition 67 *For all $n \in \mathbb{N}_0$ we have*

(1) $Z^* h_n(x) = h_{n+1}(x)$. *More generally,*

$$Z_j^* h_\alpha(\mathbf{x}) = h_{\alpha+\varepsilon(j)}(\mathbf{x}) \tag{4.7}$$

where $\varepsilon(j) = (0, \ldots, 1, \ldots, 0) \in \mathbb{N}^n$ has a 1 in the jth slot and 0 elsewhere.

(2) $Z h_n(x) = -2n h_{n-1}(x)$. *More generally,*

$$Z_j h_\alpha(\mathbf{x}) = -2\alpha_j h_{\alpha-\varepsilon(j)}(\mathbf{x}) \tag{4.8}$$

(3) $\frac{1}{2}[Z, Z^*] = -I$. *Consequently,*

$$\frac{1}{2} Z Z^* = \frac{1}{2} Z^* Z - I = \frac{1}{2}(D^2 - X^2) - \frac{1}{2} I = H_0 - \frac{1}{2} I$$

where $H_0 = \frac{1}{2}(D^2 - x^2)$ *is the* 1*-dimensional harmonic oscillator. More generally, the* n*-dimensional harmonic oscillator is given by*

$$\begin{aligned} H_0 &= \frac{1}{2}\sum_{j=1}^{n}(D_j^2 - X_j^2) && (4.9) \\ &= \frac{1}{2}\sum_{j=1}^{n}(Z_j Z_j^* + I) && (4.10) \\ &= \frac{1}{2}\sum_{j=1}^{n}(Z_j^* Z_j - I) && (4.11) \end{aligned}$$

Proof: (1) For any $f \in C^\infty(\mathbb{R})$ we have

$$D\big(e^{-x^2/2}f(x)\big) = -xe^{-x^2/2}f(x) + e^{-x^2/2}Df(x)$$

so applying this to $f(x) = h_n(x) = e^{x^2/2}D^n e^{-x^2}$ we have

$$\begin{aligned} D^{n+1}e^{-x^2} = D\big(D^n e^{-x^2}\big) = D\big(e^{-x^2/2}\big[e^{x^2/2}D^n e^{-x^2}\big]\big) \\ = D\big(e^{-x^2/2}h_n(x)\big) = -xe^{-x^2/2}h_n(x) + e^{-x^2/2}Dh_n(x) \end{aligned}$$

which implies that

$$\begin{aligned} h_{n+1}(x) &= e^{x^2/2}D^{n+1}e^{-x^2} \\ &= e^{x^2/2}\big[-xe^{-x^2/2}h_n(x) + e^{-x^2/2}Dh_n(x)\big] \\ &= -xh_n(x) + Dh_n(x) \\ &= (D - X)h_n(x) \\ &= Z^* h_n(x) \end{aligned}$$

(2) We will prove this by induction. For $n = 1$, we have

$$\begin{aligned}
Zh_1(x) &= (D + X)h_1(x) \\
&= D\big(e^{x^2/2}De^{-x^2}\big) + xe^{x^2/2}De^{-x^2} \\
&= -2D\big(xe^{-x^2/2}\big) - 2x^2e^{-x^2/2} \\
&= -2(e^{-x^2/2} - xDe^{-x^2/2}) - 2x^2e^{-x^2/2} \\
&= -2e^{-x^2/2} + 2x^2e^{-x^2/2} - 2x^2e^{-x^2/2} \\
&= -2e^{-x^2/2} \\
&= -2 \cdot 1h_0(x)
\end{aligned}$$

Now assume that $(D + X)h_k(x) = -2kh_{k-1}(x)$ for all $n \geq k \geq 1$ and consider the $(n + 1)$st case. Then

$$\begin{aligned}
-2nh_{n-1}(x) = (D + X)h_n(x) = Dh_n(x) + xh_n(x) = D(e^{x^2/2}D^ne^{-x^2}) + xh_n(x) \\
= xe^{x^2/2}D^ne^{-x^2} + e^{x^2}D^{n+1}e^{-x^2} + xh_n(x) = 2xh_n(x) + h_{n+1}(x)
\end{aligned}$$

which is a recurrence relation that may be solved for h_{n+1},

$$h_{n+1}(x) = -2nh_{n-1} - 2xh_n(x) \tag{4.12}$$

Also, from $-2nh_{n-1}(x) = (D + X)h_n(x) = Dh_n(x) + xh_n(x)$ we have

$$Dh_n(x) = -2nh_{n-1}(x) - xh_n(x) \tag{4.13}$$

From these we get

$$\begin{aligned}
(D+X)h_{n+1} &= (D+X)(-2nh_{n-1}(x)-2xh_n(x)) \\
&= -2nDh_{n-1}(x)-2nxh_{n-1}(x)-2D(xh_n(x))-2x^2h_n(x) \\
&= -2nDh_{n-1}(x)-2nxh_{n-1}(x)-2h_n(x)-2xDh_n(x)-2x^2h_n(x) \\
&= -2nDh_{n-1}(x)-2nxh_{n-1}(x)-2h_n(x)-2x(-2nh_{n-1}(x)-xh_n(x))-2x^2h_n(x) \\
&= -2nDh_{n-1}(x)-2nxh_{n-1}(x)-2h_n(x)+4nxh_{n-1}(x)+2x^2h_n(x)-2x^2h_n(x) \\
&= -2nDh_{n-1}(x)-2nxh_{n-1}(x)-2h_n(x)+2nxh_{n-1}(x) \\
&= -2n(D-x)h_{n-1}(x)-2h_n(x) \\
&= -2nh_n(x)-2h_n(x) \\
&= -2(n+1)h_n(x)
\end{aligned}$$

where we used (1) in the penultimate equality. This completes the induction and proves (2).

(3) For any $f \in C^\infty(\mathbb{R}^n)$ we have

$$\begin{aligned}
[Z,Z^*]f &= (D+X)(D-X)f-(D-X)(D+X)f \\
&= D^2f+xDf-D(xf)-x^2f-D^2f+xDf-D(xf)+x^2f \\
&= 2xDf-2D(xf) \\
&= 2xDf-2f-2xDf \\
&= -2f
\end{aligned}$$

The other statements follow by applying these results to each coordinate x_j. ■

Corollary 21 *The Hermite functions $h_\alpha \in \mathcal{S}(\mathbb{R}^n)$ are eigenfunctions of the harmonic oscillator H_0, and*

$$H_0h_\alpha = -\Big(|\alpha|+\frac{n}{2}\Big)h_\alpha \tag{4.14}$$

for all $\alpha \in \mathbb{N}_0^n$. The $-\big(|\alpha|+\frac{n}{2}\big)$-eigenspaces wiill be denoted $\mathfrak{W}_\ell$, where $\ell = |\alpha| \in \mathbb{N}_0$. These are all finite dimensional of dimension

$$\dim \mathfrak{W}_\ell = \binom{n+\ell-1}{\ell} \tag{4.15}$$

which is just the number of n-tuples $\alpha \in \mathbb{N}_0^n$ of size $|\alpha| = \ell$.

Proof: Since $Z_j h_\alpha = -2\alpha_j h_{\alpha-\varepsilon(j)}$ and $Z_j^* h_\alpha = h_{\alpha+\varepsilon(j)}$, we have

$$(Z_j Z_j^* + I)h_\alpha = Z_j h_{\alpha+\varepsilon(j)} + h_\alpha = \big(-2(\alpha_j + 1) + 1\big)h_\alpha = -(2\alpha_j + 1)h_\alpha$$

so

$$H_0 h_\alpha = \frac{1}{2}\sum_{j=1}^{n}(Z_j Z_j^* + I)h_\alpha = \sum_{j=1}^{n} -\frac{1}{2}(2\alpha_j + 1)h_\alpha = -\Big(|\alpha| + \frac{n}{2}\Big)h_\alpha$$

The dimension of the eigenspace $\mathfrak{W}_\ell$ for the eigenvalue $-\big(\ell + \frac{n}{2}\big)$ of H_0 is thus the size of the set $S = \{\alpha \in \mathbb{N}_0^n \mid |\alpha| = \ell\}$, which is $\binom{n+\ell-1}{\ell}$ by the method of stars and bars in combinatorics. ■

Notation 8 *Let $Z_j = D_j + X_j$ and $Z_j^* = D_j - X_j$ as above, and let us define the operators*

$$Z^\alpha := Z_1^{\alpha_1} Z_2^{\alpha_x} \cdots Z_n^{\alpha_n}$$

$$(Z^*)^\alpha := (Z_1^*)^{\alpha_1}(Z_2^*)^{\alpha_x} \cdots (Z_n^*)^{\alpha_n}$$

■

As in the proof of (1) in Proposition 67, $D\big(e^{-x^2/2}f(x)\big) = -xe^{-x^2/2}f(x) + e^{-x^2/2}Df(x)$, so multiplying both sides by $e^{x^2/2}$ we have

$$Z^* f(x) = (D - X)f(x) = -xf(x) + Df(x) = e^{x^2/2}D\big(e^{-x^2/2}f(x)\big)$$

Therefore, in n dimensions,

$$Z_j^* f(\mathbf{x}) = e^{x_j^2/2}D_j\big(e^{-x_j^2/2}f(\mathbf{x})\big)$$

We have also that $[Z_i, Z_j^*] = -2\delta_{ij}I$, so from all this we get

$$[Z_i, (Z^*)^\alpha] = -2\alpha_i (Z^*)^{\alpha-\varepsilon_i}$$

by induction on $|\alpha|$. Let us use this observation to show the following:

Theorem 44 *The Hermite functions $h_\alpha \in \mathcal{S}(\mathbb{R}^n)$ form an orthonormal basis (or a complete orthonormal system) for $L^2(\mathbb{R}^n)$.*

Proof: Consider the one-dimensional case first. Since $Z^*h_n = h_{n+1}$ for all n, we have $(Z^*)^n h_0 = h_n$. In general, then, $(Z^*)^\alpha h_0 = h_\alpha$. In the L^2 inner product, therefore, since the h_α are real-valued, we have $(h_\alpha, h_\beta)_{L^2(\mathbb{R}^n)} = ((Z^*)^\alpha h_0, (Z^*)^\beta h_0)_{L^2(\mathbb{R}^n)} = -(h_0, Z^\alpha, (Z^*)^\beta h_0)_{L^2(\mathbb{R}^n)}$, since Z_j and $-Z_j^*$ are formal adjoints since D_j is skew-symmetric and X_j is symmetric. If $\alpha_j \neq \beta_j$ for any j, then $(h_\alpha, h_\beta)_{L^2(\mathbb{R}^n)} = 0$, while if $\alpha = \beta$, then

$$(h_\alpha, h_\beta)_{L^2(\mathbb{R}^n)} = -(h_0, Z^\alpha (Z^*)^\beta h_0)_{L^2(\mathbb{R}^n)} = -(-2)^{|\alpha|}\alpha!$$

which we can accordingly normalize to unit length. Thus, at the very least the h_α are orthogonal in $L^2(\mathbb{R}^n)$. They are also linearly independent, for if $h_{\alpha^1}, \dots, h_{\alpha^k}$ are distinct Hermite functions and $\sum_{j=1}^k a_j h_{\alpha^j} = 0$ in $\mathcal{S}(\mathbb{R}^n)$, then writing $h_{\alpha^j} = (Z^*)^{\alpha^j} h_0$ we see that $(\sum_{j=1}^k a_j (Z^*)^{\alpha^j})h_0 = 0$. But this can only happen if all $a_j = 0$, by (1) of Proposition 67. Finally, let $V = \mathrm{span}_{\mathbb{C}}(h_\alpha \,|\, \alpha \in \mathbb{N}_0^n,\, n \in \mathbb{N})$, and note that $V^\perp = \{0\}$, or equivalently $V = L^2(\mathbb{R}^n)$: If $f \in V^\perp$, then $(f, h_\alpha)_{L^2(\mathbb{R}^n)} = 0$ for all α, which means, expanding the expressions for h_α by taking the derivative D^α of $e^{-\langle \mathbf{x}, \mathbf{x} \rangle}$, that $e^{\langle \mathbf{x}, \mathbf{x} \rangle/2} h_\alpha(\mathbf{x})$ is a degree $|\alpha|$ polynomial, call it $H_\alpha(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$ (a **Hermite polynomial**). An induction argument over $|\alpha|$ shows that these polynomials span all polynomials in $\mathbb{R}[\mathbf{x}]$. Since polynomials are dense in $L^2(\mathbb{R}^n)$, the result follows. ■

4.3 The Metaplectic Representation

Definition 56 By Proposition 26 we have that $\mathrm{Sp}(n, \mathbb{R})$ is connected and has fundamental group $\pi_1\big(\mathrm{Sp}(n, \mathbb{R})\big) \cong \mathbb{Z}$. Of course $\mathrm{Sp}(n, \mathbb{R})$ is actually path connected and semi-locally simply connected (because the unitary group is, since it has a universal cover). The Galois correspondence between the subgroups of $\pi_1\big(\mathrm{Sp}(n, \mathbb{R})\big)$ and the covers of $\mathrm{Sp}(n, \mathbb{R})$ then implies that, in particular, there corresponds to the subgroup $2\mathbb{Z}$ of $\mathbb{Z}$, which has index $[\mathbb{Z} : 2\mathbb{Z}] = 2$, a unique cover of $\mathrm{Sp}(n, \mathbb{R})$ of multiplicity 2, known as the **metaplectic group**,

$$\rho : \mathrm{Mp}(n, \mathbb{R}) \to \mathrm{Sp}(n, \mathbb{R}) \tag{4.16}$$

Here $\mathrm{Mp}(n,\mathbb{R})$ is the metaplectic group and ρ is the covering map. It's derivative

$$d\rho : \mathfrak{mp}(n,\mathbb{R}) \to \mathfrak{sp}(n,\mathbb{R}) \tag{4.17}$$

which of course satisfies

$$d\rho(\mathbf{u})(\mathbf{v}) = [\mathbf{u},\mathbf{v}] \tag{4.18}$$

is an isomorphism of Lie algebras, $\mathfrak{mp}(n,\mathbb{R}) \cong \mathfrak{sp}(n,\mathbb{R})$. These Lie algebras are characterized by Proposition 20. ■

Definition 57 Let Cl_n be the Weyl algebra, or symplectic Clifford algebra, and define

$$\mathfrak{a}(n) := \mathrm{span}(\{\mathbf{u}\cdot\mathbf{v} + \mathbf{v}\cdot\mathbf{u} \mid \mathbf{u},\mathbf{v} \in \mathbb{R}^{2n}\}) \subseteq \mathrm{Cl}_n \tag{4.19}$$

Put a Lie bracket on Cl_n is by

$$[u,v] := u\cdot v - v\cdot u$$

for all $u,v \in \mathrm{Cl}_n$. We will show that $\mathfrak{a}(n)$ is a Lie subalgebra of Cl_n which is isomorphic to $\mathfrak{sp}(n,\mathbb{R})$, so that we may view $\mathfrak{sp}(n,\mathbb{R})$ as embedded in the Weyl algebra. As we will see in the next section, Cl_n acts on $\mathcal{S}(\mathbb{R}^n)$ by polynomials in the position and momentum operators, an extension of the symplectic Clifford multiplication on $\mathcal{S}(\mathbb{R}^n)$—the Cl_n-module structure on $\mathcal{S}(\mathbb{R}^n)$—hence $\mathrm{span}(n,\mathbb{R})$, too, will act on $\mathcal{S}(\mathbb{R}^n)$. ■

Proposition 68 *We have the isomorphisms*

$$\mathfrak{a}(n) \cong (\mathbb{R}^{2n})^{\odot 2} \cong \mathfrak{sp}(n,\mathbb{R}) \tag{4.20}$$

given by

$$\mathbf{u}\cdot\mathbf{v} + \mathbf{v}\cdot\mathbf{u} \;\mapsto\; \mathbf{u}\odot\mathbf{v} \;\mapsto\; \omega_0(\cdot,\mathbf{u})\mathbf{v} + \omega_0(\cdot,\mathbf{v})\mathbf{u} \tag{4.21}$$

and using these we can see that

$$d\rho(\mathbf{u}\cdot\mathbf{v} + \mathbf{v}\cdot\mathbf{u})(\mathbf{x}) = (\mathbf{u}\odot\mathbf{v})(\mathbf{x}) = \omega_0(\mathbf{x},\mathbf{u})\mathbf{v} + \omega_0(\mathbf{x},\mathbf{v})\mathbf{u} \tag{4.22}$$

for any $\mathbf{x} \in \mathbb{R}^{2n}$*. Consequently, the inverse of* $d\rho$ *may be expressed as*

$$d\rho^{-1}(A) = \frac{1}{2}\sum_{j=1}^{n}\big(p_j \cdot (Aq_j) - q_j \cdot (Ap_j)\big) \tag{4.23}$$

in terms of the standard symplectic basis $(q,p) = (q_1, \ldots, q_n, p_1, \ldots, p_n)$ *of* $\mathbb{R}^{2n}$*.*

Proof: The second isomorphism was already given in Propositions 22 and 23, so it remains to prove the first. Consider a basic vector $\mathbf{v}_1 \cdot \mathbf{v}_2 + \mathbf{v}_2 \cdot \mathbf{v}_1$ in $\mathfrak{a}(n)$ and a vector $\mathbf{v} \in \mathbb{R}^{2n}$. We first show that the commutator $[\mathbf{v}_1 \cdot \mathbf{v}_2 + \mathbf{v}_2 \cdot \mathbf{v}_1,\ \mathbf{v}]$ may be identified with the action of $2(\mathbf{v}_1 \odot \mathbf{v}_2)$ on $\mathbf{v}$. This will show that $\mathfrak{a}(n)$ is isomorphic to $\mathfrak{sp}(n,\mathbb{R})$, since we have the obvious isomorphism between $\mathfrak{a}(n)$ and $\mathbb{R}^{2n} \odot \mathbb{R}^{2n}$, in view of the fact that the map $\psi : \mathfrak{a}(n) \to \mathbb{R}^{2n} \odot \mathbb{R}^{2n}$ has trivial kernel and is surjective.

To see the identity, note first that

$$\begin{aligned}
[\mathbf{v}_1 \cdot \mathbf{v}_2, \mathbf{v}] &= \mathbf{v}_1 \cdot \mathbf{v}_2 \cdot \mathbf{v} - \mathbf{v} \cdot \mathbf{v}_1 \cdot \mathbf{v}_2 \\
&= \mathbf{v}_1 \cdot \mathbf{v} \cdot \mathbf{v}_2 + \omega_0(\mathbf{v}, \mathbf{v}_2)\mathbf{v}_1 - \mathbf{v} \cdot \mathbf{v}_1 \cdot \mathbf{v}_2 \\
&= \omega_0(\mathbf{v}, \mathbf{v}_1)\mathbf{v}_2 + \omega_0(\mathbf{v}, \mathbf{v}_2)\mathbf{v}_1 \\
&= (\mathbf{v}_1 \odot \mathbf{v}_2)(\mathbf{v})
\end{aligned}$$

where $(\mathbf{v}_1 \odot \mathbf{v}_2)(\mathbf{v})$ means $\varphi(\mathbf{v}_1 \odot \mathbf{v}_2)(\mathbf{v})$ for φ the isomorphism $(\mathbb{R}^{2n})^{\odot 2} \to \mathfrak{sp}(n,\mathbb{R})$ of Proposition 22. Consequently,

$$[\mathbf{v}_1 \cdot \mathbf{v}_2 + \mathbf{v}_2 \cdot \mathbf{v}_1,\ \mathbf{v}] = 2(\mathbf{v}_1 \odot \mathbf{v}_2)(\mathbf{v})$$

which is an element of $\mathbb{R}^{2n}$. Extending by linearity we get that $[u, \mathbf{v}] \in \mathbb{R}^{2n}$ for all $v \in \mathfrak{a}(n)$, and as a consequence we may apply the Clifford product to such elements, which, by the derivation property of the commutator, gives

$$[\mathbf{v}_1 \cdot \mathbf{v}_2,\ v] = \mathbf{v}_1 \cdot [\mathbf{v}_2, v] + [\mathbf{v}_1, v] \cdot \mathbf{v}_2$$

and therefore

$$\begin{aligned}
[\mathbf{v}_1 \cdot \mathbf{v}_2 + \mathbf{v}_2 \cdot \mathbf{v}_1,\ v] &= \mathbf{v}_1 \cdot [\mathbf{v}_2, v] + [\mathbf{v}_1, v] \cdot \mathbf{v}_2 + \mathbf{v}_2 \cdot [\mathbf{v}_1, v] + [\mathbf{v}_2, v] \cdot \mathbf{v}_1 \\
&= \Big(\mathbf{v}_1 \cdot [\mathbf{v}_2, v] + [\mathbf{v}_2, v] \cdot \mathbf{v}_1\Big) + \Big(\mathbf{v}_2 \cdot [\mathbf{v}_1, v] + [\mathbf{v}_1, v] \cdot \mathbf{v}_2\Big)
\end{aligned}$$

which is clearly a sum of elements in $\mathfrak{a}(n)$, and therefore in $\mathfrak{a}(n)$. Thus, $\mathfrak{a}(n)$ is closed under the Lie bracket of Cl_n. ■

Remark 59 Since the unitary group $\mathrm{U}(n)$ is a subgroup of the symplectic group $\mathrm{Sp}(n,\mathbb{R})$ (Proposition 25), we can pull $\mathrm{U}(n)$ back via the double-covering map ρ to the double cover $\hat{\mathrm{U}}(n)$ of $\mathrm{U}(n)$:

$$\begin{tikzcd}
\hat{\mathrm{U}}(n) := \rho^{-1}\big(\mathrm{U}(n)\big) \arrow[r, hook, "i"] \arrow[d, "\rho"] & \mathrm{Mp}(n,\mathbb{R}) \arrow[d, "\rho"'] \\
\mathrm{U}(n) \arrow[r, hook, "i"'] & \mathrm{Sp}(n,\mathbb{R})
\end{tikzcd}$$

and we can then look at the Lie algebras $\mathfrak{u}(n)$ and $\hat{\mathfrak{u}}(n)$ of $\mathrm{U}(n)$ and $\hat{\mathrm{U}}(n)$, respectively. ■

Since $\mathrm{U}(n) \subset \mathrm{Sp}(n,\mathbb{R})$, we have $\mathfrak{u}(n) \subseteq \mathfrak{sp}(n,\mathbb{R})$ as well. Let us now characterize $\mathfrak{u}(n)$ and $\hat{\mathfrak{u}}(n)$.

Proposition 69 *We have the following characterizations of $\mathfrak{u}(n)$, the Lie algebra of $\mathrm{U}(n)$, under the isomorphisms $\mathfrak{a}(n) \cong (\mathbb{R}^{2n})^{\odot 2} \cong \mathfrak{sp}(n,\mathbb{R})$,*

$$\begin{aligned}
\mathfrak{u}(n) &= \mathfrak{sp}(n,\mathbb{R}) \cap \mathfrak{o}(n) \\
&= \mathrm{span}(\{X - J_0 X J_0 \mid X \in \mathfrak{sp}(n,\mathbb{R})\}) \\
&\cong \mathrm{span}(\{\mathbf{u} \odot \mathbf{v} + J_0\mathbf{u} \odot J_0\mathbf{v} \mid \mathbf{u},\mathbf{v} \in \mathbb{R}^{2n}\}) \\
&\cong \mathrm{span}(\{\mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{u} + J_0\mathbf{u} \cdot J_0\mathbf{v} + J_0\mathbf{v} \cdot J_0\mathbf{u} \mid \mathbf{u},\mathbf{v}\mathbb{R}^{2n}\}) \\
&\cong M_n(\mathbb{C})
\end{aligned}$$

while the Lie algebra $\hat{\mathfrak{u}}(n)$ of $\hat{\mathrm{U}}(n)$ is characterized by

$$\begin{aligned}
\hat{\mathfrak{u}}(n) &= d\rho^{-1}(\mathfrak{u}(n)) \\
&= \mathrm{span}(\{\mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{u} + J_0\mathbf{u} \cdot J_0\mathbf{v} + J_0\mathbf{v} \cdot J_0\mathbf{u} \mid \mathbf{u},\mathbf{v} \in \mathbb{R}^{2n}\}) \\
&= \mathrm{span}(\{q_i \cdot q_j + p_i \cdot p_j \mid 1 \le i \le j \le n\} \cup \{q_j \cdot p_i - p_i \cdot q_j \mid 1 \le i < j \le n\})
\end{aligned}$$

Proof: The first characterization follows from the observation that $\mathrm{U}(n)$ is identified as a subgroup of $\mathrm{Sp}(n,\mathbb{R})$ (Proposition 25), so it satisfies both the characterizations its own Lie algebra and that of $\mathrm{Sp}(n,\mathbb{R})$. Thus, for any $X \in \mathfrak{u}(n)$ we have both

$$\langle X\mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{u}, X\mathbf{v}\rangle = 0$$

and

$$\omega_0(X\mathbf{u}, \mathbf{v}) + \omega_0(\mathbf{u}, X\mathbf{v}) = 0$$

Now, the first can be rewritten as $\mathbf{v}^T X\mathbf{u} + \mathbf{u}^T X\mathbf{v} = \mathbf{0}$, or $\mathbf{v}^T X\mathbf{u} + \mathbf{v}^T X^T\mathbf{u} = \mathbf{0}$, for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{2n}$, i.e.

$$X + X^T = 0$$

while the second implies that

$$J_0 X + X^T J_0 = 0$$

by the same reasoning (see Proposition 20). Combining these we get that

$$\begin{aligned} \mathfrak{u}(n) &= \{X \in \operatorname{End}\mathbb{R}^{2n} \mid J_0 X = X J_0\} \\ &= \{X \in \operatorname{End}\mathbb{R}^{2n} \mid J_0 X J_0^{-1} = X\} \end{aligned}$$

That is, $\mathfrak{u}(n)$ is the set of (not necessarily invertible) endomorphisms of $\mathbb{R}^{2n}$ which commute with J_0. By Theorem 12 we conclude that

$$\mathfrak{u}(n) = \operatorname{End}_{\mathbb{C}}(\mathbb{R}^{2n}_J) \cong \operatorname{End}_{\mathbb{C}}(\mathbb{C}^n) \cong M_n(\mathbb{C})$$

Of course, this is natural, since the unitary group is the maximal compact subgroup not only of $\mathrm{Sp}(n,\mathbb{R})$, but of $\mathrm{GL}(n,\mathbb{C})$.

Notice, too, that if $A = X - J_0 X J_0$ for any $X \in \operatorname{End}(\mathbb{R}^{2n})$, then $A \in \mathfrak{u}(n)$, for

$$J_0 A = J_0(X - J_0 X J_0) = J_0 X - J_0^2 X = J_0 X + X J_0 = (X - J_0 X J_0) J_0 = A J_0$$

so we also have

$$\mathfrak{u}(n) = \operatorname{span}(\{X - J_0 X J_0 \mid X \in \mathfrak{sp}(n,\mathbb{R})\})$$

Now, under the identification $\mathfrak{sp}(n,\mathbb{R}) \cong (\mathbb{R}^{2n})^{\odot 2}$, $\mathbf{u}\odot\mathbf{v} \mapsto \omega_0(\cdot,\mathbf{u})\mathbf{v} + \omega_0(\cdot,\mathbf{v})\mathbf{u}$, whose inverse is $X \mapsto \frac{1}{2}\sum_{j=1}^n (Xq_j \odot p_j - q_j \odot Xp_j)$, the fact that $J_0 \in \mathfrak{sp}(n,\mathbb{R})$ means

$$0 = 0 \cdot J_0\mathbf{v}_2 = \Big(\omega_0(J_0\mathbf{v},\mathbf{v}_1) + \omega_0(\mathbf{v},J_0\mathbf{v}_1)\Big)J_0\mathbf{v}_2$$

or

$$\omega_0(\mathbf{v},J_0\mathbf{v}_1)J_0\mathbf{v}_2 = -\omega_0(J_0\mathbf{v},\mathbf{v}_1)J_0\mathbf{v}_2 = \omega_0(\mathbf{v},J_0\mathbf{v}_1)J_0\mathbf{v}_2 = -J_0\Big(\omega_0(J_0\mathbf{v},\mathbf{v}_1)\mathbf{v}_2\Big)$$

Therefore, the expression $X - J_0XJ_0$ in $\mathfrak{sp}(n,\mathbb{R})$ becomes $\mathbf{v}_1\odot\mathbf{v}_2 + J_0\mathbf{v}_1\odot J_0\mathbf{v}_2$, since

$$\begin{aligned}
\mathbf{v}_1\odot\mathbf{v}_2 + J_0\mathbf{v}_1\odot J_0\mathbf{v}_2 \quad &\leftrightarrow \quad \omega_0(\cdot,\mathbf{v}_1)\mathbf{v}_2 + \omega_0(\cdot,\mathbf{v}_2)\mathbf{v}_1 + \omega_0(\cdot,J_0\mathbf{v}_1)J_0\mathbf{v}_2 + \omega_0(\cdot,J_0\mathbf{v}_2)J_0\mathbf{v}_1 \\
&= \quad \omega_0(\cdot,\mathbf{v}_1)\mathbf{v}_2 + \omega_0(\cdot,\mathbf{v}_2)\mathbf{v}_1 - J_0\Big(\omega_0(J_0\cdot,\mathbf{v}_1)\mathbf{v}_2\Big) - J_0\Big(\omega_0(J_0\cdot,\mathbf{v}_2)\mathbf{v}_1\Big) \\
&\leftrightarrow \quad X - J_0XJ_0
\end{aligned}$$

Finally, recall the isomorphism $\mathfrak{sp}(n,\mathbb{R}) \cong \mathfrak{a}(n)$, which gives for all $\mathbf{u},\mathbf{v}\in\mathbb{R}2n$

$$\mathbf{v}_1\odot\mathbf{v}_2 + J_0\mathbf{v}_1\odot J_0\mathbf{v}_2 \mapsto (\mathbf{v}_1\cdot\mathbf{v}_2 + \mathbf{v}_2\cdot\mathbf{v}_1) + (J_0\mathbf{v}_1\cdot J_0\mathbf{v}_2 + J_0\mathbf{v}_2\cdot J_0\mathbf{v}_1)$$

and this gives the last characterization of $\mathfrak{u}(n)$. Another way to see this characterization, which will also apply to $\hat{\mathfrak{u}(n)} = d\rho^{-1}(\mathfrak{u}(n))$, is precisely by applying $d\rho^{-1}$ to $\mathbf{v}_1\odot\mathbf{v}_2 + J_0\mathbf{v}_1\odot J_0\mathbf{v}_2$ and using Proposition 68. The last characterization of $\hat{\mathfrak{u}(n)}$ follows from the canonical relations among the symplectic basis vectors (2.39), by noting that

$$q_i\cdot q_j + q_j\cdot q_i + J_0q_i\cdot J_0q_j + J_0q_j\cdot J_0q_i = 2(q_i\cdot q_j + p_ip_j)$$

and

$$q_i\cdot p_j + p_j\cdot q_i + J_0q_i\cdot J_0p_j + J_0p_j\cdot J_0q_i = 2(q_i\cdot p_j - q_j\cdot p_i)$$

This completes the proof. ■

Our interest in these groups and their Lie algebras naturally has to do with their representations, for ultimately we wish to use the representations to construct associated vector bundles to the principal bundles with structure groups $\mathrm{Mp}(n,\mathbb{R})$ and $\hat{\mathrm{U}}(n)$. The important representation of $\mathrm{Mp}(n,\mathbb{R})$ (and so its subgroup $\hat{\mathrm{U}}(n)$) is the **metaplectic**, or **Segal-Shale-Weil**, **representation**,

$$\mathfrak{m} : \mathrm{Mp}(n,\mathbb{R}) \to \mathrm{U}(L^2(\mathbb{R}^n)) \tag{4.24}$$

which is the unique unitary representation of $\mathrm{Mp}(n,\mathbb{R})$ commuting with the Schrödinger representation $\mathfrak{r}_S : \mathrm{H}(n) \to \mathrm{U}(L^2(\mathbb{R}^n))$ of the Heisenberg group, in the sense that

$$\mathfrak{m}(q) \circ \mathfrak{r}_S(\mathbf{v}, s) = \mathfrak{r}_S(\rho(q)\mathbf{v}, s) \circ \mathfrak{m}(q)$$

for all $q \in \mathrm{Mp}(n,\mathbb{R})$ and $(\mathbf{v}, s) \in \mathrm{H}(n)$. The unitary representation is faithful, and the space of **smooth vectors** of $\mathfrak{m}$, i.e. those $f \in L^2(\mathbb{R}^n)$ such that the map $\mathbf{r}_S(\cdot)f : \mathrm{Mp}(n,\mathbb{R}) \to L^2(\mathbb{R}^n)$ given by $\big(\mathbf{r}_S(\cdot)f\big)(q) := \mathfrak{r}_S(q)f$ is smooth, consists precisely of the Schwartz space $\mathcal{S}(\mathbb{R}^n)$.

Theorem 45 (Metaplectic Representation) *There exists a unique unitary representation*

$$\mathfrak{m} : \mathrm{Mp}(n,\mathbb{R}) \to U(L^2(\mathbb{R}^n)) \tag{4.25}$$

which satisfies

$$\mathfrak{m}(q) \circ \mathfrak{r}_S(\mathbf{v}, s) = \mathfrak{r}_S\big(\rho(q)\mathbf{v}, s\big) \circ \mathfrak{m}(q) \tag{4.26}$$

for all $q \in \mathrm{Mp}(n,\mathbb{R})$ and $(\mathbf{v}, s) \in H(n)$, where $H(n)$ is the Heisenberg group, $\mathfrak{r}_S(\mathbf{v}, s)$ is its Schrödinger representation, and $\rho : \mathrm{Mp}(n,\mathbb{R}) \to \mathrm{Sp}(n,\mathbb{R})$ is the double covering map. ■

Theorem 46 *The metaplectic representation $\mathfrak{m}$ is faithful. It decomposes into the sum of two inequivalent irreducible unitary representations, which are the restrictions of $\mathfrak{m}$ to the spaces of even and odd functions, respectively. Furthermore, $\mathcal{S}(\mathbb{R}^n)$ is precisely the space of smooth vectors of $\mathfrak{m}$. In particular, $\mathcal{S}(\mathbb{R}^n)$ is $\mathfrak{m}$-invariant.* ■

The significance of the above fuss over the unitary subgroup $\mathrm{U}(n)$ of $\mathrm{Sp}(n,\mathbb{R})$ and it's lift $\hat{\mathrm{U}}(n)$ to a subgroup of $\mathrm{Mp}(n,\mathbb{R})$ is that, when we work out our operators on the $\mathfrak{m}$-bundle associated to the principal $\mathrm{Mp}(n,\mathbb{R})$-bundle, we can advantageously reduce the structure group to $\hat{\mathrm{U}}(n)$ and consider the associated $\mathfrak{m}|_{\hat{\mathrm{U}}(n)}$-bundle. This is why we consider the restriction of the metaplectic representation to $\hat{\mathrm{U}}(n)$:

$$\mathfrak{u} := \mathfrak{m}|_{\hat{\mathrm{U}}(n)} : \hat{\mathrm{U}}(n) \to \mathrm{U}(L^2(\mathbb{R}^n)) \tag{4.27}$$

Of course, we also note that, in order to work with smooth vectors only, we may consider the representations $\mathfrak{u}$ and $\mathfrak{m}$ on $\mathcal{S}(\mathbb{R}^n)$. That is we may advantageously 'reduce' the vector space on which we represent our groups, in addition to reducing the group.

4.4 Symplectic Clifford Multiplication

This bring us to the last technical point, the quantization of the symplectic coordinates which finds use in the definition of symplectic Clifford multiplication. This quantization is achieved first by means of the linear map

$$\sigma : \mathbb{R}^{2n} \to \mathrm{End}(\mathcal{S}(\mathbb{R}^n)) \tag{4.28}$$

defined on the symplectic basis vectors by

$$Q_j := \sigma(q_j) := ix_j \qquad \text{(multiplication operator)} \tag{4.29}$$

$$P_j := \sigma(p_j) := \frac{\partial}{\partial x_j} \qquad \text{(differential operator)} \tag{4.30}$$

$$\sigma(1) := i \qquad \text{(multiplication operator)} \tag{4.31}$$

That is, if $q = (q_1, \ldots, q_n)$ and $p = (p_1, \ldots, p_n)$, then

$$(q, p) \xrightarrow{\sigma} (Q, P) = (i\mathbf{x}, \mathrm{grad}) \tag{4.32}$$

We refer to Section 3.2.2 above for the analytical properties of P_j and Q_j (which differ by an i from our operators here, making the above operators *skew*-symmetric, and even skew-adjoint by the results of that section).

Recall now the definition of the Weyl algebra on a symplectic vector space (which we may take to be $(\mathbb{R}^{2n}, \omega_0)$ without loss of generality by Proposition 32): it is a pair $(\mathrm{Cl}(\mathbb{R}^{2n}, \omega_0), j)$ consisting of an associative unital $\mathbb{R}$-algebra $\mathrm{Cl}(\mathbb{R}^{2n}, \omega_0)$ and an $\mathbb{R}$-linear map $j : \mathbb{R}^{2n} \to \mathrm{Cl}(\mathbb{R}^{2n}, \omega_0)$ satisfying

$$j(\mathbf{u})j(\mathbf{v}) - j(\mathbf{v})j(\mathbf{u}) = -\omega_0(\mathbf{u}, \mathbf{v}) \cdot 1$$

for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{2n}$, and the following universal property: if A is any other associative unital $\mathbb{R}$-algebra and $f : \mathbb{R}^{2n} \to A$ is linear map satisfying $f(\mathbf{u})f(\mathbf{v}) - f(\mathbf{v})f(\mathbf{u}) = -\omega_0(\mathbf{u}, \mathbf{v}) \cdot 1$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{2n}$, then there is a unique algebra homomorphism $F : \mathrm{Cl}(\mathbb{R}^{2n}, \omega_0) \to A$ such that $f = F \circ j$

$$\begin{tikzcd} & \mathrm{Cl}(\mathbb{R}^{2n}, \omega_0) \arrow[d, "F"] \\ \mathbb{R}^{2n} \arrow[ur, "j"] \arrow[r, "f"'] & A \end{tikzcd}$$

Our map σ extends by linearity to all of $\mathbb{R}^{2n}$, of course. We now want to show that it satisfies these extra two conditions, and so extends to a linear map (but not an algebra homomorphism)

$$\tilde{\sigma} : \mathrm{Cl}(\mathbb{R}^{2n}, \omega_0) \to \mathrm{End}(\mathcal{S}(\mathbb{R}^n)) \tag{4.33}$$

To see this, recall first the **canonical commutation relations**

$$[P_i, P_j] = [Q_i, Q_j] = 0, \qquad [P_i, Q_j] = i\delta_{ij} \tag{4.34}$$

(these are immediate, but see Folland [34, pp. 12-17] for further discussion). From these relations we get (Habermann and Habermann [54, Lemmas 1.4.1-1.4.2]) the following proposition.

Proposition 70 *The quantization map* $\sigma : \mathbb{R}^{2n} \to \mathrm{End}(\mathcal{S}(\mathbb{R}^n))$ *satisfies*

(1) *For all* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{2n}$ *we have*

$$\sigma(\mathbf{u}) \circ \sigma(\mathbf{v}) - \sigma(\mathbf{v}) \circ \sigma(\mathbf{u}) = -i\omega_0(\mathbf{u}, \mathbf{v}) \tag{4.35}$$

Consequently, since we have already shown the existence of $\mathrm{Cl}(\mathbb{R}^{2n}, \omega_0)$ *(Theorem 41), we have that* σ *extends to an* $\mathbb{R}$*-linear map* $\tilde{\sigma} : \mathrm{Cl}(\mathbb{R}^{2n}, \omega_0) \to \mathrm{End}(\mathcal{S}(\mathbb{R}^n))$ *making the following diagram commute:*

$$\begin{tikzcd} & \mathrm{Cl}(\mathbb{R}^{2n}, \omega_0) \arrow[d, "\tilde{\sigma}"] \\ \mathbb{R}^{2n} \arrow[ur, "j"] \arrow[r, "\sigma"'] & \mathrm{End}(\mathcal{S}(\mathbb{R}^n)) \end{tikzcd}$$

It is this extended $\tilde{\sigma}$ which appears in (4)-(6) below (we shall use σ for $\tilde{\sigma}$ understanding it to be the extended map). For future purposes also require the extended σ to respect symplectic Clifford mutliplication,

$$\sigma(\mathbf{v}_1 \cdot \mathbf{v}_2 \cdot \dots \cdot \mathbf{v}_k) := \sigma(\mathbf{v}_1) \circ \sigma(\mathbf{v}_2) \circ \dots \circ \sigma(\mathbf{v}_k) \tag{4.36}$$

(2) *With respect to the L^2-inner product on $\mathcal{S}(\mathbb{R}^n)$ the map σ sends vectors in $\mathbb{R}^{2n}$ to skew-symmetric operators,*

$$\big(\sigma(\mathbf{u})f, g\big)_{L^2} = -\big(f, \sigma(\mathbf{u})g\big)_{L^2} \tag{4.37}$$

for all $f, g \in \mathcal{S}(\mathbb{R}^n)$ and all $\mathbf{u} \in \mathbb{R}^n$.

(3) *The Fourier transform interacts with σ as*

$$\mathcal{F} \circ \sigma(J_0\mathbf{u}) = \sigma(\mathbf{u}) \circ \mathcal{F} \tag{4.38}$$

for all $\mathbf{u} \in \mathbb{R}^n$.

(4) *If $\rho : \mathrm{Mp}(n,\mathbb{R}) \to \mathrm{Sp}(n,\mathbb{R})$ is the double cover and $d\rho : \mathfrak{mp}(n,\mathbb{R}) \to \mathfrak{sp}(n,\mathbb{R})$ is its derivative, $d\rho(\mathbf{u})(\mathbf{v}) = [\mathbf{u}, \mathbf{v}]$, then*

$$\sigma(u) \circ \sigma(\mathbf{v}) - \sigma(\mathbf{v}) \circ \sigma(u) = i\sigma\big(d\rho(u)\mathbf{v}\big) \tag{4.39}$$

for all $u \in \mathfrak{mp}(n,\mathbb{R})$ and all $\mathbf{v} \in \mathbb{R}^{2n}$.

(5) *With respect to the L^2-inner product on $\mathcal{S}(\mathbb{R}^n)$, the map σ sends the metaplectic lie algebra into symmetric operators:*

$$\big(\sigma(u)f, g\big)_{L^2} = \big(f, \sigma(u)g\big)_{L^2} \tag{4.40}$$

for all $f, g \in \mathcal{S}(\mathbb{R}^n)$ and all $u \in \mathfrak{mp}(n,\mathbb{R})$. Here, we use the isomorphism $\mathfrak{mp}(n,\mathbb{R}) \cong \mathfrak{sp}(n,\mathbb{R}) \cong \mathfrak{a}(n) \subseteq \mathrm{Cl}(\mathbb{R}^{2n}, \omega_0)$.

(6) *The kernel of the extended homomorphism $\sigma : \mathrm{Cl}(\mathbb{R}^n, \omega_0) \to \mathrm{End}(\mathcal{S}(\mathbb{R}^n))$ is trivial,*

$$\ker \sigma = \{0\} \tag{4.41}$$

Proof: (1) Write any $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{2n}$ in the standard symplectic coordinates as

$$\mathbf{u} = (\mathbf{a}, \mathbf{b}) = \sum_{j=1}^{n} a_j q_j + b_j p_j \qquad \text{and} \qquad \mathbf{u} = (\mathbf{a}', \mathbf{b}') = \sum_{j=1}^{n} a'_j q_j + b'_j p_j$$

By the canonical commutation relations, we then have

$$\begin{aligned}
&\sigma(\mathbf{u}) \circ \sigma(\mathbf{v}) - \sigma(\mathbf{v}) \circ \sigma(\mathbf{u}) \\
&\qquad = \sum_{i,j=1}^{n} \big[a_i Q_i + b_i P_i\big] \circ \big[a'_j Q_j + b'_j P_j\big] - \big[a'_j Q_j + b'_j P_j\big] \circ \big[a_i Q_i + b_i P_i\big] \\
&\qquad = \sum_{i=1}^{n} (a_j b'_j - a'_j b_j)[Q_j, P_j] \\
&\qquad = -i \sum_{i=1}^{n} (a_j b'_j - a'_j b_j) \\
&\qquad = -i\big[\big\langle \mathbf{a}, \mathbf{b}' \big\rangle - \big\langle \mathbf{a}', \mathbf{b} \big\rangle\big] \\
&\qquad = -i\omega_0(\mathbf{u}, \mathbf{v})
\end{aligned}$$

the last equality by (4) of Proposition 18.

(2) This statement follows by integration by parts, noting that x_j and $i\partial/\partial x_j$ are symmetric operators (in fact essentially self-adjoint, by Theorems 36 and 37), so that ix_j and $\partial/\partial x_j$ are skew-symmetric.

(3) This statement follows from the fact that the Fourier transform intertwines the position and momentum operators, $Q_j \circ \mathcal{F} = \mathcal{F} \circ P_j$ and $P_j \circ \mathcal{F} = -\mathcal{F} \circ Q_j$ (Proposition 60), along with the fact that J_0 operates on a symplectic basis as $J_0 q_j = p_j$ and $J_0 p_j = -q_j$.

(4) From (1) we have for all $\mathbf{v}_1, \mathbf{v_2}, \mathbf{v} \in \mathbb{R}^{2n}$

$$\begin{aligned}
\sigma(\mathbf{v}_1) \circ \sigma(\mathbf{v}_2) \circ \sigma(\mathbf{v}) &= \sigma(\mathbf{v}_1) \circ \sigma(\mathbf{v}) \circ \sigma(\mathbf{v}_2) - i\omega_0(\mathbf{v}_2, v)\sigma(\mathbf{v}_1) \\
&= \sigma(\mathbf{v}) \circ \sigma(\mathbf{v}_1) \circ \sigma(\mathbf{v}_2) - i\omega_0(\mathbf{v}_1, \mathbf{v})\sigma(\mathbf{v}_2) - i\omega_0(\mathbf{v}_2, \mathbf{v})\sigma(\mathbf{v}_1)
\end{aligned}$$

Using the isomorphism $\mathfrak{mp}(n, \mathbb{R}) \cong \mathfrak{a} = \operatorname{span}(\mathbf{v_1} \cdot \mathbf{v}_2 + \mathbf{v_2} \cdot \mathbf{v}_1 \mid \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^{2n})$, we therefore have for all $u = \mathbf{v_1} \cdot \mathbf{v}_2 + \mathbf{v_2} \cdot \mathbf{v_1} \in \mathfrak{mp}(n,\mathbb{R})$ that

$$\begin{aligned}\sigma(u)\circ\sigma(\mathbf{v}) - \sigma(\mathbf{v})\circ\sigma(u) &= \sigma(\mathbf{v}_1\cdot\mathbf{v}_2+\mathbf{v}_2\cdot\mathbf{v}_1)\circ\sigma(\mathbf{v}) - \sigma(\mathbf{v})\circ\sigma(\mathbf{v}_1\cdot\mathbf{v}_2+\mathbf{v}_2\cdot\mathbf{v}_1)\\ &= 2i\sigma(\omega_0(\mathbf{v},\mathbf{v}_1)\mathbf{v}_2+\omega_0(\mathbf{v},\mathbf{v}_2)\mathbf{v}_1)\\ &= i\sigma(d\rho(\mathbf{v}_1\cdot\mathbf{v}_2+\mathbf{v}_2\cdot\mathbf{v}_1)\mathbf{v})\end{aligned}$$

(5) This follows by two applications of (2), in view of the isomorphism $\mathfrak{mp}(n,\mathbb{R}) \cong \mathfrak{a}$.

(6) Suppose $u \in \ker\sigma$. Then $\sigma(u)f = 0$ for all $f \in \mathcal{S}(\mathbb{R}^n)$. Let us express v in the basis described in Theorem 43, which consists of monomials of the form $q_1^{a_1}q_2^{a_2}\cdots q_n^{a_n}p_1^{b_1}p_2^{b_2}\cdots p_n^{b_n}$, which we can write more succinctly using multi-index notation as $q^\alpha p^\beta$. Since σ respects symplectic Clifford multiplication by definition (see (1)), we have that $\sigma(q^\alpha p^\beta) = \sigma(q)^\alpha\sigma(p)^\beta$. Applying $\sigma(u)$ thus to functions f that are polynomial near 0, which are dense in $L^2(\mathbb{R}^n)$ and therefore in $\mathcal{S}(\mathbb{R}^n)$, we see that u must be 0. ■

Definition 58 In what follows, we write $v \cdot f$ for $\sigma(v)f$ whenever $v \in \mathfrak{mp}(n,\mathbb{R})$ and $f \in \mathcal{S}(\mathbb{R}^n)$. Thus, symplectic Clifford multiplication of a Schwartz function is understood to be the induced action of $\mathrm{Cl}_n := \mathrm{Cl}(\mathbb{R}^{2n},\omega_0)$ on $\mathcal{S}(\mathbb{R}^n)$ applied to f. We consequently get a **Cl_n-module structure on $\mathcal{S}(\mathbb{R}^n)$**, whose associated action we call **symplectic Clifford multiplication**:

$$\mu_0 : \mathrm{Cl}_n \otimes \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n) \tag{4.42}$$

$$\mu_0(v\otimes f) \equiv v\cdot f := \sigma(v)f \tag{4.43}$$

This is the symplectic analog to 'Clifford module' and 'Clifford multiplication' on quadratic spaces and thence on Riemannian manifolds. Notice that if $\mathbf{u} \in \mathbb{R}^{2n}$ is written in the standard symplectic basis as $(\mathbf{a},\mathbf{b})$, i.e. $\mathbf{u} = \sum_{j=1}^n a_jq_j + b_jp_j$, then symplectic Clifford multiplication takes a particularly simple form,

$$\begin{aligned}(\mathbf{u}\cdot f)(\mathbf{x}) &= \sum_{j=1}^n \big(a_j\sigma(q_j)+b_j\sigma(p_j)\big)f(\mathbf{x})\\ &= \sum_{j=1}^n\Big(ia_jx_jf(\mathbf{x}) + b_j\frac{\partial f}{\partial x_j}(\mathbf{x})\Big)\\ &= i\langle\mathbf{a},\mathbf{x}\rangle + \langle\mathbf{b},\mathrm{grad}\, f(\mathbf{x})\rangle\end{aligned}$$

This makes clear that symplectic Clifford multiplication is an application of a linear combination of the standard multiplication and differentiation operators. This will come in handy later, when we look at the local behavior of the symplectic Dirac operators. ■

Remark 60 One nice immediate result of this symplectic Clifford module construction for $\mathcal{S}(\mathbb{R}^n)$ is that *we may express the harmonic oscillator in terms of symplectic Clifford products*:

$$H_0 := \frac{1}{2}\left(\sum_{j=1}^{n} \frac{\partial^2}{\partial x_j^2} - x_j^2\right) = \sigma\left(\frac{1}{2}\sum_{j=1}^{n} q_j \cdot q_j + p_j \cdot p_j\right) \tag{4.44}$$

i.e.

$$\sigma^{-1}(H_0) = \frac{1}{2}\sum_{j=1}^{n} q_j \cdot q_j + p_j \cdot p_j \tag{4.45}$$

which is an element of $\hat{\mathfrak{u}}(n)$. The operator H_0 is interesting for us, of course, because of its action on $L^2(\mathbb{R}^n)$ and $\mathcal{S}(\mathbb{R}^n)$. It's spectral analysis gives us a decomposition of $L^2(\mathbb{R}^n)$ into countably many finite-dimensional eigenspaces $\mathfrak{W}_\ell$ (Corollary 21 and Theorem 44 above), and since its eigenfunctions are the Hermite functions h_α, which are elements of $\mathcal{S}(\mathbb{R}^n)$, we also get a decomposition of $\mathcal{S}(\mathbb{R}^n)$ into countably many finite-dimensional subspaces $\mathfrak{S}_\ell = \mathcal{S}(\mathbb{R}^n) \cap \mathfrak{W}_\ell$. ■

Remark 61 Two smaller results about σ are:

$$\sigma = id\rho \tag{4.46}$$

on $\mathbb{R}^{2n}$, which is just the observation that (1) and (4) give the same expression on $\mathbb{R}^{2n}$, and

$$\sigma^{-1}(J_0) = \sigma^{-1}(H_0) \tag{4.47}$$

as elements of $\hat{\mathfrak{u}}(n)$. The left-hand side follows from Proposition 23. ■

Chapter 5

The Symplectic Dirac Operators

5.1 The Symplectic Spinor Bundle

The Lie groups $\mathrm{Sp}(n,\mathbb{R})$, $\mathrm{U}(n)$, $\mathrm{Mp}(n,\mathbb{R})$ and $\hat{\mathrm{U}}(n)$ each gives rise to a principal bundle:

(1) The **symplectic frame bundle**, an analog of the orthonormal frame bundle $P_{\mathrm{O}(n)}$ on a Riemannian manifold,

$$P_{\mathrm{Sp}(n,\mathbb{R})} \tag{5.1}$$

(2) The **unitary frame bundle**, a reduction of the $P_{\mathrm{Sp}(n,\mathbb{R})}$, perhaps analogous to the $\mathrm{SO}(n)$-reduction of the $P_{\mathrm{O}(n)}$ bundle (the *oriented* orthonormal frame bundle on a Riemannian manifold),

$$P_{\mathrm{U}(n)} \tag{5.2}$$

(3) The **principal** $\mathrm{Mp}(n,\mathbb{R})$**-bundle**,

$$P_{\mathrm{Mp}(n,\mathbb{R})} \tag{5.3}$$

(4) The **principal** $\hat{\mathrm{U}}(n)$**-bundle**,

$$P_{\hat{\mathrm{U}}(n)} \tag{5.4}$$

Locally on (M,ω), over a small enough local trivializaton, it is always possible to double cover the symplectic frame bundle by the metaplectic bundle, that is locally we can have both the $P_{\mathrm{Sp}(n,\mathbb{R})}|_U$ and $P_{\mathrm{Mp}(n,\mathbb{R})}|_U$ bundles and a fiber-preserving map between the two, but globally there may be topological obstructions. These obstructions are encoded in the second Stiefel-Whitney class $w_2 \in H^2(M;\mathbb{Z}_2)$, which is zero if and only if there is a globally defined bundle map

$$F : P_{\mathrm{Mp}(n,\mathbb{R})} \to P_{\mathrm{Sp}(n,\mathbb{R})} \tag{5.5}$$

which we call a **metaplectic structure** on (M, ω).

Remark 62 Consider an ω-compatible almost compatible J, that is $\omega(X, JY)$ defines a Riemannian metric g on M. By Darboux's theorem we know that, locally, with respect to a local symplectic frame $(e, f) = (e_1, \ldots, e_n, f_1, \ldots, f_n)$, J looks like the standard almost complex structure $J_0 = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$, so

$$Je_j = f_j, \qquad \text{and} \qquad Jf_j = -e_j$$

Hence, the relations

$$\omega_0(e_j, e_k) = \omega_0(f_j, f_k) = 0, \qquad \omega_0(e_j, f_k) = \delta_{jk}$$

imply the relations $\omega_0(e_j, Jf_k) = \omega_0(f_j, Je_k) = 0$ and $\omega_0(e_j, Je_k) = \delta_{jk}$, i.e.

$$g(e_j, f_k) = 0, \qquad g(e_j, e_k) = \delta_{jk}$$

Thus, if we reduce the structure group from $\mathrm{Sp}(n, \mathbb{R})$ to $\mathrm{U}(n)$, then any local unitary frame $(e_1, \ldots, e_{2n})$ is completely determined by the relations

$$g(e_j, f_k) = 0, \qquad g(e_j, e_k) = \delta_{jk}, \qquad Je_j = e_{n+j}$$

which obviously involve the choice of J. We conclude, therefore, that *there is a one-to-one correspondence between unitary reductions of the symplectic frame bundle and the set of almost complex structures J on (M, ω).* We accordingly denote by

$$P^J_{U(n)} \tag{5.6}$$

the unitary reduction of $P_{\mathrm{Sp}(n,\mathbb{R})}$ associated to J. ∎

The representations $\mathfrak{m} : \mathrm{Mp}(n, \mathbb{R}) \to \mathrm{U}(L^2(\mathbb{R}^n))$ and $\mathfrak{u} = \mathfrak{m}|_{\hat{\mathrm{U}}(n)} : \hat{\mathrm{U}}(n) \to \mathrm{U}(L^2(\mathbb{R}^n))$ now allow us to construct the associated vector bundles with typical fiber $L^2(\mathbb{R}^n)$ or $\mathcal{S}(\mathbb{R}^n)$:

(1) $\mathbf{Q} := P_{\mathrm{Mp}(n,\mathbb{R})} \times_{\mathfrak{m}} L^2(\mathbb{R}^n)$

(2) $\mathbf{S} := P_{\mathrm{Mp}(n,\mathbb{R})} \times_{\mathfrak{m}} \mathcal{S}(\mathbb{R}^n)$

(3) $\mathbf{Q}^J := P_{\hat{\mathrm{U}}(n)} \times_{\mathfrak{u}} L^2(\mathbb{R}^n)$

(4) $\mathbf{S}^J := P_{\hat{\mathrm{U}}(n)} \times_{\mathfrak{u}} \mathcal{S}(\mathbb{R}^n)$

These bundles all have a **Hermitian structure**, via the L^2 inner product on $L^2(\mathbb{R}^n)$:

$$h([p,f],\ [p,g]) := \langle f, g\rangle_{L^2(\mathbb{R}^n)} := \int_{\mathbb{R}^n} f\overline{g}\, d\mu_L \tag{5.7}$$

Thus, $(\mathbf{Q}, h)$, for example, is a *hermitian vector bundle*.

Let us brienfly describe the sections of these bundles. It is a general feature of vector bundles $P_G \times_\rho V$ associated to principal G-bundles P_G that sections $s \in \Gamma(P_G \times_\rho V)$ are in one-to-one correspondence with G-equivariant maps from P_G to V, that is maps $\hat{s} : P_G \to V$ satisfying $\hat{s}(pg) = \rho(g^{-1})\hat{s}(p)$ for all $p \in P_G$ and $g \in G$, with $\rho : G \to \mathrm{GL}(V)$ the given representation. Locally, over an open set $U \subseteq M$, sections s, or their equivariant counterparts $\hat{s}$, can be described as follows: Let $\varphi \in \Gamma(U, P_{\mathrm{Sp}(n,\mathbb{R})})$ be a local symplectic frame, and let $\overline{\varphi} \in \Gamma(U, P_{\mathrm{Mp}(n,\mathbb{R})})$ be its lift to the metaplectic bundle,

$$\begin{tikzcd} & P_{\mathrm{Mp}(n,\mathbb{R})} \arrow[d, "F"] \\ M \supseteq U \arrow[ur, "\overline{\varphi}"] \arrow[r, "\varphi"'] & P_{\mathrm{Sp}(n,\mathbb{R})} \end{tikzcd}$$

Elements of $P_G \times_\rho V$ are equivalence classes $[p, f]$ under the equivalence relation $(pg, f) \sim (p, \rho(g)f)$, i.e. orbits of the G-action on $P_G \times V$ given by $(p, f) \cdot g := (pg, \rho(g^{-1})f)$. Locally, in our case, we therefore have

$$\hat{s} \circ \overline{\varphi} : U \to \mathcal{S}(\mathbb{R}^n)$$

and so

$$s = [\overline{\varphi},\ \hat{s} \circ \overline{\varphi}] \in \Gamma(U, \mathbf{S}) = \Gamma(U,\ P_{\mathrm{Mp}(n,\mathbb{R})} \times_{\mathfrak{m}} \mathcal{S}(\mathbb{R}^n)) \tag{5.8}$$

and similarly with the other associated vector bundles.

We need this description of the sections in order to understand how the differential operators acting on such sections behave. Before we do this, however, let us note two further features of the spaces of sections of $\mathbf{Q}$, $\mathbf{S}$, $\mathbf{Q}^J$ and $\mathbf{S}^J$. First of all, the *compactly supported* sections $s \in \Gamma_c(\mathbf{Q})$ (or $\Gamma_c(\mathbf{S})$) have a natural L^2 inner product defined using the hermitian structure h on $\mathbf{Q}$ (5.7) given above:

$$(s,t) := \int_M h(s,t)\,dV = \int_M \langle f_s(x), f_t(x)\rangle_{L^2(\mathbb{R}^n)}\,dV(x) \tag{5.9}$$

where dV is the volume density on M (using its metric structure) and $f_s(x) := (\hat{s}\circ\overline{\varphi})(x)$, $f_t(x) := (\hat{t}\circ\overline{\varphi})(x) \in L^2(\mathbb{R}^n)$ or $\mathcal{S}(\mathbb{R}^n)$ for $x \in U \subseteq M$ and $\varphi \in \Gamma(U, P_{\mathrm{Sp}(n,\mathbb{R})})$ a local symplectic frame (or more generally a local frame $\varphi = (x^1, \ldots, x^{2n}) : U \to F(TM) = P_{\mathrm{GL}(2n,\mathbb{R})})$. Secondly, the Hermite functions

$$h_\alpha(\mathbf{x}) = h_{\alpha_1}(x_1)h_{\alpha_2}(x_2)\cdots h_{\alpha_n}(x_n)$$

with $h_{\alpha_j}(x_j) = e^{x_j^2/2}\big(\frac{\partial^{\alpha_j}}{\partial x_j^{\alpha_j}} e^{-x_j^2}\big)$, are the eigenfunctions of the harmonic oscillator,

$$H_0 h_\alpha(\mathbf{x}) = -\left(|\alpha| + \frac{n}{2}\right)h_\alpha(\mathbf{x})$$

and they form a complete orthonormal set for $L^2(\mathbb{R}^n)$,

$$L^2(\mathbb{R}^n) = \bigoplus_{\alpha\in\mathbb{N}_0^n} \operatorname{span}(h_\alpha) = \bigoplus_{\ell=0}^{\infty}\Big(\bigoplus_{|\alpha|=\ell}\operatorname{span}(h_\alpha)\Big) = \bigoplus_{\ell=0}^{\infty}\mathfrak{W}_\ell \tag{5.10}$$

and $\dim\mathfrak{W}_\ell = \binom{n+\ell-1}{\ell}$. Of course ℓ here corresponds to the eigenvalue $\lambda = -(\ell + n/2)$ of H_0, thus $\mathfrak{W}_\ell = E_{-(\ell+n/2)}$ is the corresponding eigenspace. It is clear that the harmonic oscillator can be defined on the Schwartz space bundles in a natural way, fiberwise:

$$\mathcal{H} : \mathbf{S} \to \mathbf{S} \tag{5.11}$$

$$\mathcal{H}([\mathbf{p}, f]) := [\mathbf{p}, H_0 f] \tag{5.12}$$

On sections, therefore, $\mathcal{H}(s) = \mathcal{H}([\overline{\varphi}, f_s]) := [\overline{\varphi}, H_0 f_s]$. The upshot, then, is that each of our four associated vector bundles decompose into a direct sum of finite-dimensional subbundles. For example, the symplectic spinor bundle decomposes as

$$\mathbf{Q} = \bigoplus_{\ell=0}^{\infty}\mathbf{Q}_\ell := \bigoplus_{\ell=0}^{\infty} P_{\mathrm{Mp}(n,\mathbb{R})} \times_{\mathfrak{m}} \mathfrak{W}_\ell \tag{5.13}$$

5.2 The Spinor Derivative

First, let us define the **(symplectic) spinor derivative**, that is the connecton on the symplectic spinor bundle

$$\nabla : \Gamma(\mathbf{Q}) \to \Gamma(T^*M \otimes \mathbf{Q}) \tag{5.14}$$

which we may, for convenience, also define on the Schwartz space spinors,

$$\nabla : \Gamma(\mathbf{S}) \to \Gamma(T^*M \otimes \mathbf{S}) \tag{5.15}$$

We begin by generalizing the one-to-one correspondence between smooth $\mathrm{Mp}(n,\mathbb{R})$-equivariant functions from $P_{\mathrm{Mp}(n,\mathbb{R})}$ to $\mathcal{S}(\mathbb{R}^n)$ (or $L^2(\mathbb{R}^n)$) and the smooth sections of the symplectic spinor bundle $\Gamma(\mathbf{Q})$ (or $\Gamma(\mathbf{S})$),

$$\Gamma(M, \mathbf{Q}) \cong C^\infty(P_{\mathrm{Mp}(n,\mathbb{R})}, L^2(\mathbb{R}^n))^{\mathrm{Mp}(n,\mathbb{R})} \tag{5.16}$$

We establish a similar correspondence for differential forms on M valued in $\mathbf{Q}$ (or $\Gamma(\mathbf{S})$) and **basic** differential forms on $P_{\mathrm{Mp}(n,\mathbb{R})}$ valued in $L^2(\mathbb{R}^n)$ (or $\mathcal{S}(\mathbb{R}^n)$):

$$\Omega^k(M, \mathbf{Q}) \cong \Omega^k(P_{\mathrm{Mp}(n,\mathbb{R})}; L^2(\mathbb{R}^n))_{\mathrm{bas}} \tag{5.17}$$

This, again, is a standard isomorphism for associated bundles $P_G \times_\rho V$. Given such a bundle, where $\rho : G \to \mathrm{GL}(V)$ or $\mathrm{U}(V)$ is the given representation, we define the correspondence

$$\Omega^k(M, P_G \times_\rho V) \ni \alpha_M \;\; \longleftrightarrow \;\; \alpha \in \Omega^k(P_G; V)_{\mathrm{bas}} \tag{5.18}$$

by

$$\pi^*\alpha_M(X_1, \ldots, X_k)_x = [p,\ \alpha(X_1, \ldots, X_k)_p] \tag{5.19}$$

where $\pi : P_G \to M$ is the projection map, $x \in M$ and $p \in \pi^{-1}(x)$. A **basic k-form** $\alpha \in \Omega^k(P_G; V)$, then, is one that, in addition to being G-equivariant, i.e. $R_g^*\alpha = \rho(g^{-1})\alpha$ for all $g \in G$, also vanishes on vertical vectors, that is $\alpha(X_1, \ldots, X_k) = 0$ if any $X_j \in \Gamma(VP_G)$ (equivalently $T\pi(X) = 0$, since $T_p\pi : H_pP_G \to T_{\pi(p)}M$ is an isomorphism). Thus, the above correspondence is indeed a bijection.

Now consider a symplectic connection $\nabla : \Gamma(TM) \to \Gamma(T^*M \otimes TM)$, i.e. one satisfying $\nabla\omega = 0$, or equivalently $d(\omega(X,Y)) = \omega(\nabla X, Y) + \omega(X, \nabla Y)$ for all $X, Y \in \Gamma(TM)$. It has an associated connection 1-form, also denoted here ω, valued in $\mathfrak{sp}(n, \mathbb{R}^n)$, $\omega \in \Omega^1(P_{\mathrm{Sp}(n,\mathbb{R})}; \mathfrak{sp}(n, \mathbb{R}))$, which we pull back via the bundle map $F : P_{\mathrm{Mp}(n,\mathbb{R})} \to P_{\mathrm{Sp}(n,\mathbb{R})}$ to a form $F^*\omega \in \Omega^1(P_{\mathrm{Mp}(n,\mathbb{R})}; \mathfrak{sp}(n, \mathbb{R}))$, then lift it to a 1-form on $P_{\mathrm{Mp}(n,\mathbb{R})}$ valued in $\mathfrak{mp}(n, \mathbb{R})$,

$$\tilde{\omega} := d\rho^{-1} \circ F^*\omega \in \Omega^1(P_{\mathrm{Mp}(n,\mathbb{R})}; \mathfrak{mp}(n, \mathbb{R}))$$

With this connection 1-form in hand we take a local symplectic frame $\varphi = (e, f) \in \Gamma(U, P_{\mathrm{Sp}(n,\mathbb{R})})$ and use it to pull $\tilde{\omega}$ back to $U \subseteq M$, giving us a local $\mathfrak{mp}(n, \mathbb{R})$-valued 1-form on U:

$$\alpha := \varphi^*\tilde{\omega} \in \Omega^1(U; \mathfrak{mp}(n, \mathbb{R})) = \Gamma(U, T^*M \otimes \mathfrak{mp}(n, \mathbb{R})) \tag{5.20}$$

Finally, using α we can define a covariant derivative on $\mathbf{Q} = P_{\mathrm{Mp}(n,\mathbb{R})} \times_{\mathfrak{m}} L^2(\mathbb{R}^n)$ by

$$\nabla s = d\hat{s} + d\mathfrak{m}(\tilde{\omega}) \wedge \hat{s} \tag{5.21}$$

$$= [\overline{\varphi},\ d(\hat{s} \circ \overline{\varphi}) + d\mathfrak{m}(\alpha)(\hat{s} \circ \overline{\varphi})] \tag{5.22}$$

That is, at a point $x \in U \subseteq M$, with $p \in \pi^{-1}(x) \subseteq P_{\mathrm{Mp}(n,\mathbb{R})}$,

$$\begin{aligned}(\nabla_X s)_p &= d\hat{s}(X)_p + d\mathfrak{m}(\tilde{\omega}_p(X)) \wedge \hat{s}(p) \\ &= X_p\hat{s} + d\mathfrak{m}(\tilde{\omega}_p(X))(\hat{s}(p))\end{aligned}$$

where $X \in \Gamma(TP_{\mathrm{Mp}(n,\mathbb{R})})$, or if we wish to think of $X \in \Gamma(TM)$,

$$(\nabla_X s)_x = [\overline{\varphi}(x),\ d(\hat{s} \circ \overline{\varphi})_x(X) + d\mathfrak{m}(\alpha_x(X))((\hat{s} \circ \overline{\varphi})(x))]$$

Schematically, we define the spinor derivative ∇ on $\mathbf{Q}$ by the following diagram:

$$\begin{tikzcd}
\Gamma\big(M, P_{\mathrm{Mp}(n,\mathbb{R})} \times_{\mathfrak{m}} L^2(\mathbb{R}^n)\big) \arrow[r, "\nabla"] \arrow[d, "\sim"'] & \Gamma\big(M, T^*M \otimes P_{\mathrm{Mp}(n,\mathbb{R})} \times_{\mathfrak{m}} L^2(\mathbb{R}^n)\big) \\
\Omega^0(P_{\mathrm{Mp}(n,\mathbb{R})}; L^2(\mathbb{R}^n))_{\mathrm{bas}} \arrow[r, "d + d\mathfrak{m}(\tilde{\omega})"'] & \Omega^1(P_{\mathrm{Mp}(n,\mathbb{R})}; L^2(\mathbb{R}^n))_{\mathrm{bas}} \arrow[u, "\sim"']
\end{tikzcd}$$

i.e. transferring the problem from the vector bundle to the bundle of basic 0-forms and using the exterior derivative there.

Remark 63 We should note that, generally speaking, we may define a spinor derivative ∇ for any choice of connection 1-form $\omega \in \Omega^1(P_{\mathrm{Mp}(n,\mathbb{R})}; \mathfrak{mp}(n,\mathbb{R}))$, not just the one gotten from the chosen symplectic connection. Indeed, even the symplectic connection had to be chosen, so even $\tilde{\omega}$ involved a choice. Thus, we see that ∇ should really be defined with respect to the choice of ω, and should be denoted ∇^ω. ■

More generally, since the isomorphism on forms holds for all k, not just 0 and 1, we may define the **exterior covariant derivative** by the same means, using the exterior derivative on basic forms:

$$d^\nabla : \Omega^k(M; \mathbf{Q}) \to \Omega^{k+1}(M; \mathbf{Q}) \tag{5.23}$$

Let us describe the local expression of $(\nabla_X s)_x$ in a local symplectic basis $\varphi = (e, f)$. First of all, note that for any $X \in \Gamma(TM)$ we have

$$\alpha(X) = \varphi^*(d\rho^{-1} \circ F^*\omega(X)) = d\rho^{-1}(\varphi^* F^*\omega(X)) = d\rho^{-1}(\varphi^*\omega(X))$$

Secondly, we note that by Proposition 1.4.5 in Habermann and Habermann, [54], the action of $d\mathfrak{m}(v)$ in $\mathrm{End}(\mathcal{S}(\mathbb{R}^n))$, where $v \in \mathfrak{mp}(n, \mathbb{R}^n)$, is given by Clifford multiplication:

$$d\mathfrak{m}(v)f = -iv \cdot f \tag{5.24}$$

i.e. $-i\sigma(v)(f)$. Combining these results, and using the fact noted above that $d\rho^{-1}$ is an isomorphism between $\mathfrak{sp}(n,\mathbb{R})$ and $\mathfrak{mp}(n,\mathbb{R}) \cong \mathfrak{a}(n) \subseteq \mathrm{Cl}_n$ with inverse given by $d\rho^{-1}(A) = \frac{1}{2}\sum_{j=1}^n \big(p_j \cdot (Aq_j) - q_j \cdot (Ap_j)\big)$, cf (4.23), we get, by letting $A = \varphi^*\omega(X)$ and $f_s := \hat{s} \circ \overline{\varphi}$, that

$$d\mathfrak{m}(\alpha(X))(f_s) = -i\alpha(X) \cdot f_s = -\frac{i}{2}\left(\sum_{j=1}^n (p_j \cdot \varphi^*\omega(X)(q_j) - q_j \cdot \varphi^*\omega(X)(p_j)\right) \cdot f_s$$

Therefore, writing a local section as $s = [\overline{\varphi}, f_s]$, we have

$$\begin{aligned}
(\nabla_{\tilde{X}} s)_p = (\nabla_X s)_x &= [\overline{\varphi}(x),\ X_x f_s + d\mathfrak{m}(\alpha_x(X))(f_s)] \qquad (5.25)\\
&= \left[\overline{\varphi}(x),\ X_x f_s - \frac{i}{2}\left(\sum_{j=1}^n (p_j \cdot \varphi^*\omega(X)(q_j) - q_j \cdot \varphi^*\omega(X)(p_j)\right) \cdot f_s\right]\\
&= ds_p(\tilde{X}) - \frac{i}{2}\sum_{j=1}^n (f_j \cdot \nabla_{\tilde{X}} e_j - e_j \cdot \nabla_{\tilde{X}} f_j) \cdot s
\end{aligned}$$

where $\tilde{X} \in \Gamma(HP_{\mathrm{Mp}(n,\mathbb{R})})$ is the horizontal lift of $X \in \Gamma(TM)$, i.e. $T\pi(\tilde{X}) = X$.

5.3 The Symplectic Dirac Operators and Related Operators

Since the symplectic form ω on M is by definition nondegenerate, we can use it to define **flattening** and **sharpening operators**. The flattening operator

$$\omega^\flat : \Gamma(TM) \to \Gamma(T^*M) = \Omega^1(M) \tag{5.26}$$

is given by

$$\omega^\flat(X) := \omega(X, \cdot) \tag{5.27}$$

and using it we get the sharpening operator:

$$\omega^\sharp := (\omega^\flat)^{-1} : \Omega^1(M) \to \Gamma(TM) \tag{5.28}$$

Also, given an ω-compatible almost complex structure J on TM, we get a Riemannian metric,

$$g(\cdot, \cdot) := \omega(\cdot, J\cdot)$$

and so a second set of flattening and sharpening operators:

$$g^\flat : \Gamma(TM) \to \Gamma(T^*M) = \Omega^1(M) \tag{5.29}$$

$$g^\flat(X) := g(X, \cdot) \tag{5.30}$$

and using this

$$g^\sharp := (g^\flat)^{-1} : \Omega^1(M) \to \Gamma(TM) \tag{5.31}$$

We remark that, for a given $\alpha \in \Omega^1(M)$, we have $g^\sharp(\alpha) = X$ iff $\alpha = g^\flat(X)$ iff for all $Y \in \Gamma(TM)$ we have $\alpha(Y) = g(X, Y) = \omega(X, JY) = \omega(JX, J^2Y) = \omega(JX, -Y) = \omega(-JX, Y) = \omega^\flat(-JX)(Y)$, which shows that

$$-JX = \omega^\sharp(\alpha) \iff X = g^\sharp(\alpha)$$

or equivalently, since $J^2 = -I$,

$$X = \omega^\sharp(\alpha) \iff JX = g^\sharp(\alpha) \tag{5.32}$$

With these isomorphisms in hand, as well as the the spinor derivative and the Clifford multiplication map $\mu : \Gamma(TM \otimes \mathbf{Q}) \to \Gamma(\mathbf{Q})$, $\mu(X \otimes s) := X \cdot s := \sigma(X)(s)$, we can define our two **symplectic Dirac operators**:

$$D := \mu \circ \omega^\sharp \circ \nabla : \Gamma(\mathbf{Q}) \to \Gamma(\mathbf{Q}) \tag{5.33}$$

$$\tilde{D} := \mu \circ g^\sharp \circ \nabla : \Gamma(\mathbf{Q}) \to \Gamma(\mathbf{Q}) \tag{5.34}$$

That is, they are the two compositions, respectively,

$$\Gamma(\mathbf{Q}) \xrightarrow{\nabla} \Gamma(T^*M \otimes \mathbf{Q}) \xrightarrow[g^\sharp]{\omega^\sharp} \Gamma(TM \otimes \mathbf{Q}) \xrightarrow{\mu} \gamma(\mathbf{Q})$$

Using the local expression (5.25) of the spinor derivative ∇ above, we get local expressions for the symplectic Dirac operators, relative to a symplectic frame (e, f):

$$\begin{aligned} D &= \sum_{j=1}^n e_j \cdot \nabla_{f_j} - f_j \cdot \nabla_{e_j} && (5.35)\\ &= -\Big(\sum_{j=1}^n (Je_j) \cdot \nabla_{e_j} + (Jf_j) \cdot \nabla_{f_j}\Big) && (5.36)\end{aligned}$$

and

$$\begin{aligned} \tilde{D} &= \sum_{j=1}^n (Je_j) \cdot \nabla_{f_j} - (Jf_j) \cdot \nabla_{e_j} && (5.37)\\ &= \sum_{j=1}^n e_j \cdot \nabla_{e_j} + f_j \cdot \nabla_{f_j} && (5.38)\end{aligned}$$

From these two operators we get the following two important Dirac-type operators:

$$Z := D + i\tilde{D} : \Gamma(\mathbf{Q}) \to \Gamma(\mathbf{Q}) \tag{5.39}$$

and

$$Z^* := D - i\tilde{D} : \Gamma(\mathbf{Q}) \to \Gamma(\mathbf{Q}) \tag{5.40}$$

Let us also define a few other operators in terms of these:

$$\begin{aligned} P &:= i[\tilde{D}, D] = i(\tilde{D}D - D\tilde{D}) && (5.41)\\ Q &:= i(\tilde{D}D + D\tilde{D}) && (5.42)\\ R &:= [Z, Z^*]_{\mathrm{gr}} := ZZ^* + Z^*Z && (5.43)\end{aligned}$$

Lastly, we have the **Fourier transform**, defined fiberwise:

$$\mathcal{F} : \mathbf{Q} \to \mathbf{Q} \tag{5.44}$$

$$\mathcal{F}([\mathbf{p}, f]) := [\mathbf{p}, \mathcal{F}f] \tag{5.45}$$

as well as the harmonic oscillator

$$\mathcal{H} : \mathbf{Q} \to \mathbf{Q} \tag{5.46}$$

$$\mathcal{H}([\mathbf{p}, f]) := [\mathbf{p}, H_0 f] \tag{5.47}$$

From the local expression for D and $\tilde{D}$ it follows (cf. Habermann and Habermann [54, Proposition 5.3.1]) that when $\nabla J = 0$ we have

$$[\mathcal{H}, D] = i\tilde{D} \tag{5.48}$$

$$[\mathcal{H}, \tilde{D}] = -iD \tag{5.49}$$

and hence

$$[\mathcal{H}, Z] = [\mathcal{H}, D] + i[\mathcal{H}, \tilde{D}] = i\tilde{D} + i(-i)D = Z \tag{5.50}$$

$$[\mathcal{H}, Z^*] = [\mathcal{H}, D] - i[\mathcal{H}, \tilde{D}] = i\tilde{D} + (-i)^2 D = -(D - i\tilde{D}) = -Z^* \tag{5.51}$$

$$[\mathcal{H}, P] = i[\mathcal{H}, [\tilde{D}, D]] = -i[\tilde{D}, [D, \mathcal{H}]] - i[D, [\mathcal{H}, \tilde{D}] = -i[\tilde{D}, -i\tilde{D}] - i[D, -iD] = 0 \tag{5.52}$$

the last by the Jacobi identity. Thus, the operator P, which is of Laplace type, preserves the eigenbundles $\mathbf{Q}_\ell$, while the operators Z and Z^* decrease and increase, respectively, the degree of the eigenbundles, $Z : \Gamma(\mathbf{Q}_\ell) \to \Gamma(\mathbf{Q}_{\ell-1})$ and $Z^* : \Gamma(\mathbf{Q}_\ell) \to \Gamma(\mathbf{Q}_{\ell+1})$:

$$\mathcal{H}Zs_\ell = Z\mathcal{H}s_\ell + Zs_\ell = -\Big(\ell + \frac{n}{2}\Big)Zs_\ell + Zs_\ell = -\Big((\ell - 1) + \frac{n}{2}\Big)Zs_\ell \tag{5.53}$$

$$\mathcal{H}Z^*s_\ell = Z^*\mathcal{H}s_\ell - Z^*s_\ell = -\Big(\ell + \frac{n}{2}\Big)Z^*s_\ell - Z^*s_\ell = -\Big((\ell + 1) + \frac{n}{2}\Big)Z^*s_\ell \tag{5.54}$$

Moreover, since $D = \frac{1}{2}(Z + Z^*)$ and $\tilde{D} = \frac{1}{2i}(Z - Z^*)$, we see that both D and $\tilde{D}$ move eigensections $s_\ell \in \Gamma(\mathbf{Q}_\ell)$ to $\Gamma(\mathbf{Q}_{\ell+1})$ and $\Gamma(\mathbf{Q}_{\ell+1})$.

5.4 Properties and Relations among the Operators

We begin by fixing the function spaces on which our operators will be assumed to act. We recall from Section 5.1 above that we have constructed two (infinite-dimensional hermitian) vector bundles, the symplectic spinor bundle $\mathbf{Q} \to M$ and the symplectic Schwartz space spinor bundle $\mathbf{S} \to M$, on whose sections we suppose our operators will act. But we will of course be concerned with *differential* operator, so smoothness of sections is a concern, and we also want our operators to be densely defined on a Hilbert space $\mathcal{H}$, presumably with those smooth sections dense in $\mathcal{H}$.

Let us begin with smoothness: From Lemma 3.2.3 in Habermann and Habermann [54], we know that the *smooth* sections of $\mathbf{Q}$, denoted $\Gamma(\mathbf{Q})$ and defined by the isomorphism (5.16), $\Gamma(\mathbf{Q}) \cong C^\infty(P_{\mathrm{Mp}(n,\mathbb{R})}, L^2(\mathbb{R}^n))^{\mathrm{Mp}(n,\mathbb{R})}$ [1] , are in fact the (continuous) sections of $\mathbf{S}$

$$\Gamma(\mathbf{Q}) = C(M, \mathbf{S}) \cong C(P_{\mathrm{Mp}(n,\mathbb{R})}, \mathcal{S}(\mathbb{R}^n))^{\mathrm{Mp}(n,\mathbb{R})} \tag{5.55}$$

and the *smooth* sections of $\mathbf{S}$ are analogously defined to be those (continuous) sections s of $\mathbf{S}$ corresponding to smooth maps $\hat{s} \in C^\infty(P_{\mathrm{Mp}(n,\mathbb{R})}, \mathcal{S}(\mathbb{R}^n))^{\mathrm{Mp}(n,\mathbb{R})}$,

$$\Gamma(\mathbf{S}) \cong C^\infty(P_{\mathrm{Mp}(n,\mathbb{R})}, \mathcal{S}(\mathbb{R}^n))^{\mathrm{Mp}(n,\mathbb{R})} \tag{5.56}$$

Clearly $\Gamma(\mathbf{S}) \subseteq \Gamma(\mathbf{Q})$. Next, we define the *compactly supported sections*

$$\Gamma_c(\mathbf{Q}) \tag{5.57}$$

$$\Gamma_c(\mathbf{S}) \tag{5.58}$$

of $\mathbf{Q}$ and $\mathbf{S}$ to be those compactly supported sections whose corresponding $\mathrm{Mp}(n,\mathbb{R})$-equivariant maps from $P_{\mathrm{Mp}(n,\mathbb{R})}$ to $L^2(\mathbb{R}^n)$, respectively $\mathcal{S}(\mathbb{R}^n)$, are smooth.

Lastly, the Hilbert spaces of sections on which our operators will be densely defined (on $\Gamma(\mathbf{Q}$ or $\Gamma_c(\mathbf{Q})$ or their Schwartz analogs) will be the L^2 completions of $\Gamma_c(\mathbf{Q})$ and $\Gamma_c(\mathbf{S})$ with respect to

[1] This congruence, (5.16), may be given simply in terms of continuous functions, $C(M,\mathbf{Q}) \cong C(P_{\mathrm{Mp}(n,\mathbb{R})}, L^2(\mathbb{R}^n))^{\mathrm{Mp}(n,\mathbb{R})}$, and then we distinguish the *smooth* sections from the merely *continuous* by saying they are smooth if their corresponding $\mathrm{Mp}(n,\mathbb{R})$-equivariant functions are smooth. This is the sense in which Habermann and Habermann say a section s of $\mathbf{Q}$ is smooth (Definition 3.2.2 in [54]).

$(\cdot,\cdot)$ defined in (5.9) above, $(s,t) = \int_M \langle f_s(x), f_t(x)\rangle_{L^2(\mathbb{R}^n)}\, dV(x)$,

$$L^2(\mathbf{Q}) = \overline{\Gamma_c(\mathbf{Q})} \tag{5.59}$$

$$L^2(\mathbf{S}) = \overline{\Gamma_c(\mathbf{S})} \tag{5.60}$$

We will mostly work with $\Gamma(\mathbf{S})$ and $\Gamma_c(\mathbf{S})$, since $\Gamma_c(\mathbf{Q}) = C_c(M,\mathbf{S})$ contains the dense (by definition) subspace $\Gamma_c(\mathbf{S}) = C_c^\infty(M,\mathbf{S})$ of $L^2(\mathbf{Q})$. Throughout this section, unless otherwise specified, we will assume the operators to be acting on $\Gamma_c(\mathbf{S})$.

One last assumption, which we may later weaken, is that M be compact. This will give us $\Gamma_c(\mathbf{Q}) = \Gamma(\mathbf{Q})$ and $\Gamma_c(\mathbf{S}) = \Gamma(\mathbf{S})$, and make statements concerning formal adjoints of operators less problematic, because once we know a section is smooth, it will also be compactly supported and thus in the domains of the operators.

Theorem 47 *Let $t = Zs = Z^*\tau \in \operatorname{im} Z \cap \operatorname{im} Z^*$. Then there are unique symplectic spinor fields s' and τ' such that $Ds' = \tilde{D}\tau' = t' \in \operatorname{im} D \cap \operatorname{im} \tilde{D}$, namely*

$$s' = s - \tau \tag{5.61}$$

$$\tau' = -is - i\tau \tag{5.62}$$

and conversely, if $t' = Ds' = \tilde{D}\tau'$, then there are unique s and τ such that $Zs = Z^*\tau = t \in \operatorname{im} Z \cap \operatorname{im} Z^*$, namely

$$s = \frac{1}{2}s' + \frac{i}{2}\tau' \tag{5.63}$$

$$\tau = -\frac{1}{2}s' + \frac{i}{2}\tau' \tag{5.64}$$

That is, $\begin{pmatrix} s' \\ \tau' \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -i & -i \end{pmatrix} \begin{pmatrix} s \\ \tau \end{pmatrix}$ *and* $\begin{pmatrix} s \\ \tau \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & i \\ -1 & i \end{pmatrix}\begin{pmatrix} s' \\ \tau' \end{pmatrix}$, *where*

$$\frac{1}{2}\begin{pmatrix} 1 & i \\ -1 & i \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -i & -i \end{pmatrix}^{-1}$$

which is to say sources (s,τ) and (s',τ') of common images are related by an invertible linear map between $\Gamma(\mathbf{Q}) \oplus \Gamma(\mathbf{Q})$ and itself.

Proof: Let $t = Zs = Z^*\tau$, and define $s' = s - \tau$ and $\tau' = -is - i\tau$. Then,

$$Ds + i\tilde{D}s = Zs = t = Z^*\tau = D\tau - i\tilde{D}\tau$$

which implies

$$Ds' = D(s-\tau) = \tilde{D}(-is - i\tau) = \tilde{D}\tau'$$

Conversely, if $Ds' = \tilde{D}\tau'$, define $s = \frac{1}{2}s' + \frac{i}{2}\tau'$ and $\tau = -\frac{1}{2}s' + \frac{i}{2}\tau'$ and note that

$$\begin{aligned}
Zs &= \frac{1}{2}Zs' + \frac{i}{2}Z\tau' \\
&= \frac{1}{2}Ds' + \frac{i}{2}\tilde{D}s' + \frac{i}{2}D\tau' - \frac{1}{2}\tilde{D}\tau' \\
&= \frac{1}{2}\tilde{D}\tau' + \frac{i}{2}\tilde{D}s' + \frac{i}{2}D\tau' - \frac{1}{2}Ds' \\
&= -\frac{1}{2}\big(Ds' - i\tilde{D}s'\big) + \frac{i}{2}\big(D\tau' - i\tilde{D}\tau'\big) \\
&= -\frac{1}{2}Z^*s' + \frac{i}{2}Z^*\tau' \\
&= Z^*\tau
\end{aligned}$$

which proves the claim. ■

Corollary 22 *With the above notation we have*

$$t' = Ds' = \tilde{D}\tau' = \frac{1}{2}\big(Z\tau - Z^*s\big) \tag{5.65}$$

Proof: Since $s' = s - \tau$, we have $\tau = s - s'$ and $s = \tau + s'$, so $Z\tau = Zs - Zs'$ and $Z^*s = Z^*\tau + Z^*s' = Zs + Z^*s'$. Thus, $Z^*s - Z\tau = Z^*s' + Zs' = 2Ds' = 2\tilde{D}\tau' = 2t'$. ■

Theorem 48 *If $Ds' = \tilde{D}\tau'$, then $Ds' = \tilde{D}\tau' = 0$, so that*

$$\operatorname{im} D \cap \operatorname{im} \tilde{D} = \{0\} \tag{5.66}$$

and in addition

$$Ds = D\tau = \frac{i}{2}D\tau' \tag{5.67}$$

$$\tilde{D}s = -\tilde{D}\tau = \frac{1}{2}\tilde{D}s' \tag{5.68}$$

$$Z\tau = Z^*s \tag{5.69}$$

$$Zs' = -Z^*s' \tag{5.70}$$

$$Z\tau' = Z^*\tau' \tag{5.71}$$

Proof: Firstly, from equation (5.65) we have that $2Ds' = 2\tilde{D}\tau' = Z\tau - Z^*s$, which gives us the identities

$$\begin{aligned} Z\tau &= Z^*s + 2Ds' && (5.72)\\ &= Z^*s + 2\tilde{D}\tau' && (5.73)\end{aligned}$$

and

$$\begin{aligned} Z^*s &= Z\tau - 2Ds' && (5.74)\\ &= Z\tau - 2\tilde{D}\tau' && (5.75)\end{aligned}$$

Also, since $s' = s - \tau$ and $\tau' = -is - i\tau$, we have $\tilde{D}s' = \tilde{D}s - \tilde{D}\tau$ and $D\tau' = -iDs - iD\tau$, and consequently

$$\begin{aligned} \tilde{D}s' + D\tau' &= \tilde{D}s - \tilde{D}\tau - iDs - iD\tau && (5.76)\\ &= -i(D + i\tilde{D})s - i(D\tau - i\tilde{D}\tau)\\ &= -iZs - iZ^*\tau\\ &= -2iZs\\ &= -2iZ^*\tau\end{aligned}$$

and, using the identities (5.72)-(5.75),

$$\begin{aligned}\tilde{D}s' - D\tau' &= \tilde{D}s - \tilde{D}\tau + iDs + iD\tau \\ &= i(D - i\tilde{D})s + i(D\tau + i\tilde{D}\tau) \\ &= iZ^*s + iZ\tau \\ &= 2iZ^*s + 2iDs' \\ &= 2iZ\tau - 2iDs'\end{aligned} \tag{5.77}$$

We can add (5.76) and (5.77) in three different ways:

$$\begin{aligned}2\tilde{D}s' &= (-iZs - iZ^*\tau) + (iZ^*s + iZ\tau) = -i(Z - Z^*)s + i(Z - Z^*)\tau \\ &= 2\tilde{D}s - 2\tilde{D}\tau\end{aligned}$$

$$\begin{aligned}&= (-2iZs) + (2iZ^*s + 2iDs') = -2i(Z - Z^*)s + 2iDs' \\ &= -2i(2i\tilde{D}s) + 2iDs' \\ &= 4\tilde{D}s + 2iDs'\end{aligned}$$

$$\begin{aligned}&= (-2iZ^*\tau) + (2iZ\tau - 2iDs') = 2i(Z - Z^*)\tau - 2iDs' \\ &= 2i(2i\tilde{D}\tau) - 2iDs' \\ &= -4\tilde{D}\tau - 2iDs'\end{aligned}$$

which give us three identities:

$$\tilde{D}s' = \tilde{D}s - \tilde{D}\tau \tag{5.78}$$

$$= 2\tilde{D}s + iDs' \tag{5.79}$$

$$= -2\tilde{D}\tau - iDs' \tag{5.80}$$

(We remark here that that (5.78) is consistent with merely applying $\tilde{D}$ to $s' = s - \tau$, while (5.79) and (5.80) change if instead we use the identities $\tau = i\tau' - s$ and $s = i\tau' - \tau$, which follow from

(5.62), giving instead $\tilde{D}s' = 2\tilde{D}s - i\tilde{D}s' = -2\tilde{D}\tau + iDs'$. The reason, as we can see in what follows, is that $Ds' = D\tau' = 0$.) Subtracting (5.80) from (5.79) we get that

$$0 = 2\tilde{D}s + 2\tilde{D}\tau + 2iDs'$$

so that by (5.62)

$$Ds' = i(\tilde{D}s + \tilde{D}\tau) = -(-i)(\tilde{D}s + \tilde{D}\tau) = -\tilde{D}(-is - i\tau) = -\tilde{D}\tau'$$

But then $\tilde{D}\tau' = Ds' = -\tilde{D}\tau'$, where the first equality is by our starting assumption and the second is from the equation just above. This proves our first claim,

$$\tilde{D}\tau' = Ds' = 0 \tag{5.81}$$

Alternatively, we could have subtracted (5.76) and (5.77) in three different ways:

$$\begin{aligned}
2D\tau' &= (-iZs - iZ^*\tau) - (iZ^*s + iZ\tau) = -i(Z + Z^*)s - i(Z + Z^*)\tau = -2iDs - 2iD\tau \\
&= (-2iZs) - (2iZ^*s + 2iDs') = -2i(Z + Z^*)s - 2iDs' = -4iDs - 2iDs'' \\
&= (-2iZ^*\tau) - (2iZ\tau - 2iDs') = -2i(Z + Z^*)\tau + 2iDs' = -4iD\tau + 2iDs'
\end{aligned}$$

which give us three further identities:

$$\begin{aligned}
D\tau' &= -iDs - iD\tau & (5.82)\\
&= -2iDs - iDs' & (5.83)\\
&= -2iD\tau + iDs' & (5.84)
\end{aligned}$$

and subtracting (5.83) from (5.84) would similarly give $Ds' = \tilde{D}\tau' = 0$. These identities prove useful in other ways, though, for we can plug $Ds' = 0$ back into them to get

$$Ds = D\tau = \frac{i}{2}D\tau' \tag{5.85}$$

$$\tilde{D}s = -\tilde{D}\tau = \frac{1}{2}\tilde{D}s' \tag{5.86}$$

while plugging $Ds' = 0$ back into (5.79) and (5.80) gives

$$2\tilde{D}s = \tilde{D}s' = -2\tilde{D}\tau$$

and, of course, plugging $Ds' = 0$ into (5.65) gives

$$Z^* s = Z\tau \tag{5.87}$$

Since also $Zs = Z^*\tau$, however, we conclude

$$2Ds = Zs + Z^* s = Z^*\tau + Z\tau = 2D\tau$$

Finally, since $Zs = Z^*\tau$ and $Z\tau = Z^* s$, we have $Z(s - \tau) = -Z^*(s - \tau)$ and $Z(s + \tau) = Z^*(s + \tau)$, which, in light of equations (5.61)-(5.62), means $Zs' = -Z^* s'$ and $Z\tau' = Z^*\tau'$. ■

Proposition 71 *With D, $\tilde{D}$, Z, Z^*, P, and R as above, we have the following simple relations:*

$$D = \frac{1}{2}(Z + Z^*) \tag{5.88}$$

$$\tilde{D} = \frac{1}{2i}(Z - Z^*) \tag{5.89}$$

$$P = \frac{1}{2}[Z, Z^*] \tag{5.90}$$

$$Z^2 = D^2 + Q - \tilde{D}^2 \tag{5.91}$$

$$Z^{*2} = D^2 - Q - \tilde{D}^2 \tag{5.92}$$

$$Z^2 + Z^{*2} = 2(D^2 - \tilde{D}^2) \tag{5.93}$$

$$Z^2 - Z^{*2} = 2Q \tag{5.94}$$

$$2Q + 2P = (Z - Z^*)(Z + Z^*) = 4i\tilde{D}D \tag{5.95}$$

$$2Q - 2P = (Z + Z^*)(Z - Z^*) = 4iD\tilde{D} \tag{5.96}$$

$$ZZ^* = D^2 + P + \tilde{D}^2 \tag{5.97}$$

$$Z^*Z = D^2 - P + \tilde{D}^2 \tag{5.98}$$

$$R = 2(D^2 + \tilde{D}^2) \tag{5.99}$$

Proof:

(1) $Z + Z^* = D + i\tilde{D} + D - i\tilde{D} = 2D$.

(2) $Z - Z^* = D + i\tilde{D} - (D - i\tilde{D}) = 2i\tilde{D}$.

(3) $Z^2 = (D + i\tilde{D})^2 = D^2 + i(\tilde{D}D + D\tilde{D}) + i^2\tilde{D}^2 = D^2 + Q - \tilde{D}^2$.

(4) $[Z, Z^*] = [D + i\tilde{D}, D - i\tilde{D}] = [D, D] + i[\tilde{D}, D] - i[D, \tilde{D}] + i^2[\tilde{D}, \tilde{D}] = 2i[\tilde{D}, D] = 2P$.

(5) $Z^{*2} = (D - i\tilde{D})^2 = D^2 - i(\tilde{D}D + D\tilde{D}) + i^2\tilde{D}^2 = D^2 - Q - \tilde{D}^2$.

(6)-(7) follow from (3) and (5).

(8) $2Q + 2P = (Z^2 - Z^{*2}) + (ZZ^* - Z^*Z) = (Z - Z^*)(Z + Z^*) = (2i\tilde{D})(2D) = 4i\tilde{D}D$.

(9) $2Q - 2P = (Z^2 - Z^{*2}) - (ZZ^* - Z^*Z) = (Z + Z^*)(Z - Z^*) = (2D)(2i\tilde{D}) = 4iD\tilde{D}$.

(10) $ZZ^* = (D + i\tilde{D})(D - i\tilde{D}) = D^2 + i(\tilde{D}D - D\tilde{D}) - i^2\tilde{D}^2 = D^2 + P + \tilde{D}^2$.

(11) $Z^*Z = (D - i\tilde{D})(D + i\tilde{D}) = D^2 - i(\tilde{D}D - D\tilde{D}) - i^2\tilde{D}^2 = D^2 - P + \tilde{D}^2$.

(12) Follows from (10) and (11). ■

A fixed choice of $J \in \Gamma(\mathcal{J}(M,\omega))$ gives a reduction of the structure group of the symplectic spinor bundle $\mathbf{Q}$ to $\hat{\mathrm{U}}(n)$. To see this, recall Theorem 16, which says that each reduction of the structure group of TM to $\mathrm{U}(n)$ corresponds to a unique ω-compatible almost complex structure $J \in \Gamma(\mathcal{J}(M,\omega))$. Lifting this reduction to $\hat{\mathrm{U}}(n)$ by the restriction of the double covering map $\rho : \mathrm{Mp}(n,\mathbb{R}) \to \mathrm{Sp}(n,\mathbb{R})$, $\hat{\mathrm{U}}(n) = \rho^{-1}(\mathrm{U}(n))$ and representing it on $L^2(\mathbb{R}^n)$ by the metaplectic representation $\mathfrak{m}$ gives the $\hat{\mathrm{U}}(n)$-reduction on $\mathbf{Q}$. If we also require of our J that it be parallel, $\nabla J = 0$, then Proposition 57 tells us that J commutes with ∇, $J(\nabla_X Y) = \nabla_X(JY)$, and this in turn gives us the commutation relation $\mathcal{F}(\nabla_X s) = \nabla_X(\mathcal{F}s)$. To call attention to this choice of J, Habermann and Habermann write $\mathcal{F}^J$. We shall avoid this extra notation, because for our purposes we shall mostly require $\nabla J = 0$ anyway, and often we shall also need the $\mathrm{U}(n)$-reduced symplectic spinor bundles. In any case, under these assumptions, we get the following commutation relations between our operators and $\mathcal{F}$:

Proposition 72 *If we choose a symplectic connection ∇ and an almost complex structure J which is parallel with respect to this connection, i.e. $\nabla J = 0$, then the relations between the Fourier* *transform $\mathcal{F}$ and the other operators is as follows:*

$$\mathcal{F}D = -\tilde{D}\mathcal{F} \tag{5.100}$$

$$\mathcal{F}\tilde{D} = D\mathcal{F} \tag{5.101}$$

$$\mathcal{F}Q = -Q\mathcal{F} \tag{5.102}$$

$$\mathcal{F}P = P\mathcal{F} \tag{5.103}$$

$$\mathcal{F}Z = iZ\mathcal{F} \tag{5.104}$$

$$\mathcal{F}Z^* = -iZ^*\mathcal{F} \tag{5.105}$$

$$\mathcal{F}D^2 = \tilde{D}^2\mathcal{F} \tag{5.106}$$

$$\mathcal{F}\tilde{D}^2 = D^2\mathcal{F} \tag{5.107}$$

$$\mathcal{F}Z^2 = -Z^2\mathcal{F} \tag{5.108}$$

$$\mathcal{F}Z^{*2} = -Z^{*2}\mathcal{F} \tag{5.109}$$

$$\mathcal{F}R = R\mathcal{F} \tag{5.110}$$

Proof: (5.100), (5.101) and (5.103) are Proposition 7.3.1 in Habermann and Habermann, [54]. The rest follow from the first two:

(1) $\mathcal{F}Q = i\mathcal{F}\tilde{D}D + i\mathcal{F}D\tilde{D} = -iD\tilde{D}\mathcal{F} - i\tilde{D}D\mathcal{F} = -Q\mathcal{F}$.

(2) $\mathcal{F}Z = \mathcal{F}D + i\mathcal{F}\tilde{D} = -\tilde{D}\mathcal{F} + iD\mathcal{F} = i(D + i\tilde{D})\mathcal{F} = iZ\mathcal{F}$.

(3) $\mathcal{F}Z^* = \mathcal{F}D - i\mathcal{F}\tilde{D} = -\tilde{D}\mathcal{F} - iD\mathcal{F} = -i(D - i\tilde{D})\mathcal{F} = -iZ^*\mathcal{F}$.

(4) $\mathcal{F}D^2 = (-1)^2\tilde{D}^2\mathcal{F} = \tilde{D}^2\mathcal{F}$.

(5) $\mathcal{F}\tilde{D}^2 = D^2\mathcal{F}$.

(6) $\mathcal{F}Z^2 = i^2Z^2\mathcal{F} = -Z^2\mathcal{F}$.

(7) $\mathcal{F}Z^{*2} = (-i)^2Z^{*2}\mathcal{F} = -Z^{*2}\mathcal{F}$.

(8) $\mathcal{F}R = \mathcal{F}ZZ^* + \mathcal{F}Z^*Z = i(-i)ZZ^* + (-i)iZ^*Z\mathcal{F} = (ZZ^* + Z^*Z)\mathcal{F} = R\mathcal{F}$. ■

Proposition 73 *On a compact symplectic manifold (M, ω) we have the following kernel identities:*

(1) $\ker D = \ker D^2$ *and* $\ker \tilde{D} = \ker \tilde{D}^2$ *when these operators are restricted to their domains* $\Gamma_c(\mathbf{S})$*.*

(2) $\ker D \cap \ker \tilde{D} = \ker D^2 \cap \ker \tilde{D}^2 = \ker Z \cap \ker Z^* = \ker R = \ker(D^2 + \tilde{D}^2)$.

(3) $\ker ZZ^* = \ker Z^*$ *and* $\ker Z^*Z = \ker Z$

Proof: (1) We clearly have $\ker D \subseteq \ker D^2$, so suppose $s \in \ker D^2$, by which we mean $D^2 s = 0$ weakly, or $0 = (0, D^2\sigma)$ for all $\sigma \in \Gamma_c(\mathbf{S})$. Then, $0 = (0, s) = (D^2 s, s) = (Ds, Ds) = \|Ds\|^2$, so $Ds = 0$. Similarly with $\tilde{D}^2$.

(2) Since $\subseteq$ is clear, suppose that $s \in \ker Z \cap \ker Z^*$. Then $s \in \ker(Z + Z^*) = \ker D$ and $s \in \ker(Z - Z^*) = \ker \tilde{D}$.

(3) Since $\supseteq$ is clear, suppose $s \in \ker R$. Then $0 = (Rs, s) = (ZZ^*s, s) + (Z^*Zs, s) = (Z^*s, Z^*s) + (Zs, Zs) = \|Z^*s\|^2 + \|Zs\|^2$, which implies $Zs = Z^*s = 0$.

(4) Since $\subseteq$ is clear, let $s \in \ker D^2 \cap \ker \tilde{D}^2$. Then $s \in \ker(D^2 + \tilde{D}^2) = \ker R = \ker Z \cap \ker Z^* = \ker D \cap \ker \tilde{D}$, since $R = 2(D^2 + \tilde{D}^2)$.

(5) Clearly $\ker Z^* \subseteq \ker ZZ^*$, so suppose $s \in \ker ZZ^*$. Then, $0 = (ZZ^*s, s) = (Z^*s, Z^*s) = \|Z^*s\|^2$, which implies $Z^*s = 0$. Similarly, if $s \in \ker Z^*Z$, then $0 = (Z^*Zs, s) = (Zs, Zs) = \|Zs\|^2$, which implies $Zs = 0$. ∎

Theorem 49 *The following kernel identities hold on any compact symplectic manifold (M,ω):*

$$\begin{aligned}
\ker D \cap \ker \tilde{D} &= \ker D^2 \cap \ker \tilde{D}^2 && (5.111)\\
&= \ker(D^2 \pm \tilde{D}^2) && (5.112)\\
&= \ker Z && (5.113)\\
&= \ker Z^* && (5.114)\\
&= \ker ZZ^* && (5.115)\\
&= \ker Z^*Z && (5.116)\\
&= \ker Z^2 \cap \ker (Z^*)^2 && (5.117)\\
&= \ker(Z^2 + (Z^*)^2) && (5.118)\\
&= \ker R && (5.119)\\
&\subseteq \ker P && (5.120)\\
&= \ker Q && (5.121)\\
&= \ker(Z^2 - (Z^*)^2) && (5.122)\\
&= \ker \tilde{D}D \cap \ker D\tilde{D} && (5.123)
\end{aligned}$$

Consequently, since P is elliptic, its kernel consists of smooth sections,

$$\ker P \subseteq \Gamma(\mathbf{S}) \tag{5.124}$$

so if we assume M is compact, this means all kernel elemets of Z and Z^*, being contained in $\ker P$*, are smooth and compactly supported, and therefore in the domains of all the operators (see also Corollary 23 below for weak solutions). Finally, if $s \in \ker P$, then $Ds \perp \tilde{D}s$, i.e. $(Ds, \tilde{D}s) = 0$:*

$$D(\ker P) \perp \tilde{D}(\ker P) \tag{5.125}$$

Proof: (1) If $s \in \ker Z = \ker(D + i\tilde{D})$, then $Ds = -i\tilde{D}s = \tilde{D}(-is) \in \operatorname{im} D \cap \operatorname{im} \tilde{D} = \{0\}$, so $Ds = \tilde{D}s = 0$. Similarly, if $s \in \ker Z^* = \ker(D - i\tilde{D})$, then $Ds = i\tilde{D}s = \tilde{D}(is) = 0$, so $Ds = \tilde{D}s = 0$. Consequently, since $\ker D \cap \ker \tilde{D} = \ker Z \cap \ker Z^* \subseteq \ker Z$ and $\ker D \cap \ker \tilde{D} = \ker Z \cap \ker Z^* \subseteq$

$\ker Z^*$, we have $\ker Z \subseteq \ker Z^*$ and $\ker Z^* \subseteq \ker Z$, and so

$$\ker Z = \ker Z^* = \ker Z \cap \ker Z^* = \ker D \cap \ker \tilde{D}$$

The equalities

$$\ker D \cap \ker \tilde{D} = \ker D^2 \cap \ker \tilde{D}^2 = \ker R = \ker(D^2 + \tilde{D}^2)$$

and

$$\ker Z^* = \ker ZZ^*, \qquad \ker Z = \ker Z^* Z$$

were proved in Proposition 73, so it remains to show that $\ker(D^2 - \tilde{D}^2)$ equals any of these, from which will follow also that $\ker(Z^2 + (Z^*)^2) = \ker(D^2 - \tilde{D}^2)$, since $Z^2 + (Z^*)^2 = 2(D^2 - \tilde{D}^2)$, and also that $\ker Z^2 \cap \ker(Z^*)^2$ equals any of these, since $\ker Z^2 \cap \ker(Z^*)^2 \subseteq \ker(Z^2 + (Z^*)^2)$ and $\supseteq$ follows from the fact that $\ker D^2 \cap \ker \tilde{D}^2$ is contained in $\ker(Z^2 + (Z^*)^2)$. We will show that $\ker(D^2 \pm \tilde{D}^2) = \ker D^2 \cap \ker \tilde{D}^2$: If $s \in \ker(D^2 \pm \tilde{D}^2)$, then $D^2 s = \pm \tilde{D}^2 s$, i.e. $D(Ds) = \tilde{D}(\pm \tilde{D}s)$. By Theorem 48 this equals 0, so $D^2 s = \tilde{D}^2 s = 0$ and $s \in \ker D^2 \cap \ker \tilde{D}^2 = \ker D = \ker \tilde{D}$. The reverse inclusion is obvious. Since $Z^2 + (Z^*)^2 = 2(D^2 - \tilde{D}^2)$, we finally have that $\ker(Z^2 + (Z^*)^2) = \ker(D^2 - \tilde{D}^2)$.

(2) We clearly have $\ker \tilde{D}D \cap \ker D\tilde{D} \subseteq \ker P = i[\tilde{D}, D]$, so suppose $s \in \ker P$. Then $0 = Ps = i\tilde{D}Ds - iD\tilde{D}s$, so $\tilde{D}Ds = D\tilde{D}s = 0$ by (5.66), which shows that $\ker P = \ker \tilde{D}D \cap \ker D\tilde{D}$. Thus, if $\sigma \in \ker P$, then $D\tilde{D}\sigma = \tilde{D}D\sigma = 0$, and therefore $Qs = i\tilde{D}D\sigma + iD\tilde{D}\sigma = 0$, so $\ker P \subseteq \ker Q$. Similarly, if $\sigma \in \ker Q$, then $D\tilde{D}\sigma = -\tilde{D}D\sigma = \tilde{D}(-D\sigma) = 0$, so $P\sigma = i\tilde{D}D\sigma - D\tilde{D}\sigma = 0$, and $\ker Q \subseteq \ker P$. The last equality follows from (5.94).

(3) The last assertion is shown as follows: If $s \in \ker P$, then $\tilde{D}Ds = D\tilde{D}s = 0$, so using the symmetry of D and $\tilde{D}$, along with the ellipticity of P, which by elliptic regularity ensures all elements in its kernel are smooth, and so in the domains of D and $\tilde{D}$ on a compact manifold M, we get

$$0 = (D\tilde{D}s, s) = (\tilde{D}s, Ds) \qquad \blacksquare$$

Since the Fourier transform appears as a sort of intertwiner of the two symplectic Dirac operators (5.100)-(5.101), it is inevitable that its properties will play a role in our analysis of D and $\tilde{D}$. Two of the more important of its properties that will be of help to us are, (1) $\mathcal{F}$ is an isometry of $L^2(\mathbf{Q})$, which follows from its definition fiberwise on $L^2(\mathbb{R}^n)$. In the Euclidean setting $\mathcal{F}$ also restricts to an invertible bounded operator on $\mathcal{S}(\mathbb{R}^n)$, so it is so also on the Schwartz spinor bundle $\Gamma(\mathbf{S})$. (2) The spectral properties of $\mathcal{F}$: the Fourier transform has four eigenvalues, ± 1 and $\pm i$, and the associated eigenspaces form an orthogonal decomposition of $L^2(\mathbb{R}^n)$, which restrict to an orthogonal decomposition of $\mathcal{S}(\mathbb{R}^n)$. Moreover, the eigenfunctions of $\mathcal{F}$ on $L^2(\mathbb{R}^n)$ are the same as the eigenfunctions of the Harmonic oscillator H_0, the Hermite functions. We now wish to describe the extension of this decomposition to $L^2(\mathbf{Q})$ and $\Gamma(\mathbf{S})$. The importance of this decomposition will become apparent when we employ it to study the behavior of D and $\tilde{D}$ on the associated eigen-subbundles.

Lemma 14 *The four eigenspaces of the Fourier transform $\mathcal{F}$ on $L^2(\mathbf{Q})$ defined in (5.44) above give an orthogonal decomposition of $L^2(\mathbf{Q})$,*

$$L^2(\mathbf{Q}) = \bigoplus_{j=0}^{3} \ker(\mathcal{F} - (-i)^j I) \tag{5.126}$$

and the jth orthogonal projection onto $\ker(\mathcal{F} - (-i)^j I)$ is given by

$$P_j = \frac{1}{4}\sum_{k=0}^{3} (i^j \mathcal{F})^k \tag{5.127}$$

Proof: This follows from the fact that we have these statements for the Fourier transform on $L^2(\mathbb{R}^n)$, in combination with our definition of $\mathcal{F}$ on $L^2(\mathbf{Q})$ fiberwise (see (5.44) above), which uses the fact that the unitary representation $\mathfrak{u} = \mathfrak{m}|_{\hat{U}(n)}$ (the restriction of the metraplectic representation to the lift $\hat{U}(n)$ of $U(n)$ in $\mathrm{Mp}(n,\mathbb{R})$) decomposes into irreducible componenets $\mathfrak{u}_\ell$, the restrictions of $\mathfrak{u}$ to the $\mathfrak{u}$-invariant subspaces $\mathfrak{W}_\ell = \bigoplus_{|\alpha|=\ell} \mathrm{span}(h_\alpha)$—these happen to be subspaces of the eigenspaces of $\mathcal{F}$, since the Hermite functions are also eigenfunctions of $\mathcal{F}$ (see (5.10) above and Habermann and Habermann [54, Cor. 1.5.2]). To see that the projections are those stated above, consider any $f \in L^2(\mathbb{R}^n)$, and let $f_j = P_j f$. Then,

$$\mathcal{F}f_0 = \frac{1}{4}\mathcal{F}\sum_{k=0}^{3}\mathcal{F}^k f = \frac{1}{4}\sum_{k=0}^{3}\mathcal{F}^{k+1} f = \frac{1}{4}\left(\sum_{k=0}^{3}\mathcal{F}^k f + \mathcal{F}^4 f - \mathcal{F}^0 f\right) = f_0$$

$$\mathcal{F}f_1 = \frac{1}{4}\mathcal{F}\sum_{k=0}^{3}(i\mathcal{F})^k f = \frac{1}{4}\sum_{k=0}^{3} i^k\mathcal{F}^{k+1} f = -\frac{i}{4}\sum_{k=0}^{3}(i\mathcal{F})^k f = -if_1$$

$$\mathcal{F}f_2 = \frac{1}{4}\mathcal{F}\sum_{k=0}^{3}(-\mathcal{F})^k f = -\frac{1}{4}\sum_{k=0}^{3}(-\mathcal{F})f^k = -f_2$$

$$\mathcal{F}f_3 = \frac{1}{4}\mathcal{F}\sum_{k=0}^{3}(-i\mathcal{F})^k f = \frac{1}{4}\sum_{k=0}^{3}(-i)^k\mathcal{F}^{k+1} f = \frac{i}{4}\sum_{k=0}^{3}(-i\mathcal{F})^k f = if_3$$

where we used the fact that $\mathcal{F}^4 = I$. Consequently,

$$\sum_{j=0}^{3} f_0 = \frac{1}{4}\sum_{j=0}^{3}\sum_{k=0}^{3}(i^j\mathcal{F})^k f = \frac{1}{4}\sum_{k=0}^{3}(1 + i^k + i^{2k} + i^{3k})\mathcal{F}^k f = \frac{1}{4}4f = f$$

It follows from these considerations that any spinor field $s \in L^2(\mathbf{Q})$ decomposes uniquely into four orthogonal pieces, $s = s_0 + s_1 + s_2 + s_3$, on which $\mathcal{F}$ acts as $\mathcal{F}s_j = (-i)^j s_j$. ∎

Theorem 50

(1) $\ker D \cong \ker\tilde{D}$, *in fact*

$$\ker D = \mathcal{F}(\ker\tilde{D}) \qquad \textit{and} \qquad \ker\tilde{D} = \mathcal{F}(\ker D)$$

Moreover, if $s = s_0 + s_1 + s_2 + s_3 \in \ker D$, *where* $s_j \in \ker(\mathcal{F} - (-i)^j I)$, *then*

$$s_0 + s_2,\ s_1 + s_3 \in \ker D$$

$$s_0 - s_2,\ s_1 - s_3 \in \ker\tilde{D}$$

as well, and similarly if $t = t_0 + t_1 + t_2 + t_3 \in \ker\tilde{D}$, *then*

$$t_0 + t_2,\ t_1 + t_3 \in \ker\tilde{D}$$

$$t_0 - t_2,\ t_1 - t_3 \in \ker D$$

(2) *If* $s \ker D \cap \ker\tilde{D}$, *then writing* $s = s_0 + s_1 + s_2 + s_3 \in$ *where* $s_j \in \ker(\mathcal{F} - (-i)^j I)$, *and* $s = \sum_{\ell=0}^{\infty}\sigma_\ell$, *where* $\sigma_\ell \in \Gamma(\mathbf{S}_\ell)$, *we will have that each* s_j *and* σ_ℓ *lies in* $\ker D \cap \ker\tilde{D}$ *as*

well, and

$$\begin{aligned}
\ker D \cap \ker \tilde{D} &= \{s \in \ker D \mid \mathcal{F}s \in \ker D\} \\
&= \{s \in \ker \tilde{D} \mid \mathcal{F}s \in \ker \tilde{D}\} \\
&= \{s \in \Gamma(\mathbf{S}) \mid \mathcal{F}^j s \in \ker D \cap \tilde{D}, \forall j = 0,1,2,3\} \\
&= \{s \in \Gamma(\mathbf{S}) \mid s_j \in \ker D \cap \ker \tilde{D}, \forall j = 0,1,2,3\} \\
&= \{s \in \Gamma(\mathbf{S}) \mid s_\ell \in \ker D \cap \ker \tilde{D}, \forall \ell \in \mathbb{N}_0\}
\end{aligned}$$

Proof: (1) The first statement follows the commutation relations (5.100)-(5.101) and the fact that $\mathcal{F}$ is an isometry on $\Gamma(\mathbf{Q})$, $Ds = 0$ implies $0 = \mathcal{F}0 = \mathcal{F}Ds = -\tilde{D}\mathcal{F}s$, and similarly $\tilde{D}s = 0$ implies $0 = \mathcal{F}0 = \mathcal{F}\tilde{D}s = D\mathcal{F}s$. Now, by definition of s_j being an element in $\ker(\mathcal{F} - (-i)^j I)$, we have

$$\begin{aligned}
&s = \mathcal{F}^0 s = s_0 + s_1 + s_2 + s_3 \\
&\mathcal{F}s = s_0 - is_1 - s_2 + is_3 \\
&\mathcal{F}^2 s = s_0 - s_1 + s_2 - s_3 \\
&\mathcal{F}^3 s = s_0 + is_1 - s_2 - is_3
\end{aligned}$$

so if $s \in \ker D$, then $Ds = 0$, and applying the intertwining relations (5.100)-(5.101) and $\mathcal{F}^j$, $j = 1,2,3$, to this equality and using the fact that $\mathcal{F}$ is an isometry, we get

$$\begin{aligned}
&D(s_0 + s_1 + s_2 + s_3) = Ds = 0 \\
&\tilde{D}(s_0 - is_1 - s_2 + is_3) = \tilde{D}\mathcal{F}s = -\mathcal{F}Ds = 0 \\
&D(s_0 - s_1 + s_2 - s_3) = D\mathcal{F}^2 s = -\mathcal{F}^2 Ds = 0 \\
&\tilde{D}(s_0 + is_1 - s_2 - is_3) = \tilde{D}\mathcal{F}^3 s = \mathcal{F}^3 Ds = 0
\end{aligned}$$

so from the first and third identities we have $-D(s_0 + s_2) = D(s_1 + s_3) = D(s_0 + s_2)$, which must therefore equal 0,

$$D(s_0 + s_2) = D(s_1 + s_3) = 0$$

and similarly from the second and fourth identities we have $-\tilde{D}(s_0-s_2) = i\tilde{D}(s_1-s_3) = \tilde{D}(s_0-s_2)$, which must therefore equal 0,

$$\tilde{D}(s_0 - s_2) = \tilde{D}(s_1 - s_3) = 0$$

Reversing the roles of D and $\tilde{D}$ gives the other set of identities.

(2) Since $D\sigma = \tilde{D}\sigma = 0 = \mathcal{F}0 = \mathcal{F}\tilde{D}\sigma = -D\mathcal{F}\sigma$, and similarly $\tilde{D}\sigma = 0 = -\tilde{D}\mathcal{F}\sigma$, the result follows. Now use Lemma 14. First, note that if we write $s = s_0 + s_1 + s_2 + s_3$ where $s_j \in \ker(\mathcal{F} - (-i)^j)$, then $s \in \ker D \cap \ker \tilde{D}$ implies each $s_j \in \ker D \cap \ker \tilde{D}$ by applying $\mathcal{F}^j$ to both sides of $Ds = \tilde{D}s = 0$. For, from $D\mathcal{F}^j s = 0$ for $j = 0, 1, 2, 3$ we get

$$D(s_0 + s_2) = -D(s_1 + s_3)$$

$$D(s_0 - s_2) = iD(s_1 - s_3)$$

$$D(s_0 + s_2) = D(s_1 + s_2)$$

$$D(s_0 - s_2) = -iD(s_1 - s_3)$$

We conclude that $-D(s_1 + s_3) = D(s_0 + s_2) = D(s_1 + s_3) = 0$ and $-iD(s_1 - s_3) = D(s_0 - s_2) = iD(s_1 - s_3) = 0$, which gives $-Ds_0 = Ds_2 = Ds_0 = 0$ and $-Ds_1 = Ds_3 = -Ds_3 = 0$. Thus $Ds_j = 0$ for all $j = 0, 1, 2, 3$. We similarly we get $\tilde{D}s_j = 0$ for all j. Since $s_0 + s_2$ consists of the even sections $\sigma_\ell \in \Gamma(\mathbf{S}_\ell)$, $\ell = 2k$ (because $\mathcal{F}\sigma_\ell = (-i)^\ell \sigma_\ell$), and $s_1 + s_3$ consists of the odds, we see that $D\sigma_\ell = \tilde{D}\sigma_\ell = 0$ for all $\ell \in \mathbb{N}_0$, because $2D = Z + Z^*$ and $2i\tilde{D} = Z - Z^*$ switch the parity, with Z moving degrees down by one and Z^* moving degrees up by one, by (5.50)-(5.54). The converse of these statements is clear, since $s_j \in \ker D \cap \ker \tilde{D}$ for each j implies $s \in \ker D \cap \ker \tilde{D}$, and similarly with the σ_ℓ. ∎

5.5 Essential Self-Adjointness of the Symplectic Dirac Operators

Let us now consider again our sympectic Dirac operators D and $\tilde{D}$. Throughout this section we assume that the symplectic connection ∇ on M is torsion-free and that $\nabla J = 0$, so that Theorem 4.5.3, Habermann and Habermann [54] applies to give us that D and $\tilde{D}$ are symmetric on $\mathcal{H} = L^2(\mathbf{Q})$. We begin by showing that D and $\tilde{D}$ both have equal deficiency indices (so that Theorem 32 above applies to give the existence of self-adjoint extensions of D and $\tilde{D}$). Following this and a detailed description of the relations between the deficiency subspaces $\ker(D^* \pm iI)$ and $\ker(\tilde{D}^* \pm iI)$, we show that each deficiency subspace is in fact trivial, so there is a *unique* self-adjoint extension to each of D and $\tilde{D}$.

Theorem 51 *The deficiency indices of D and $\tilde{D}$ are the same, so D and $\tilde{D}$ have self-adjoint extensions. Moreover, if $n_\pm(D)$ are the deficience indices of D and $n_\pm(\tilde{D})$ those of $\tilde{D}$, we have*

$$n_+(D) = n_-(D) = n_+(\tilde{D}) = n_-(\tilde{D}) \tag{5.128}$$

Proof: We begin by showing that

$$\mathrm{im}(D + iI) \cong \mathrm{im}(\tilde{D} - iI) \cong \mathrm{im}(D - iI) \cong \mathrm{im}(\tilde{D} + iI)$$

for $D \pm iI$, $\tilde{D} \pm iI$ on their common domain $\Gamma_c(\mathbf{S})$. Indeed, from (5.100)-(5.101) we know that $\mathcal{F}\tilde{D} = D\mathcal{F}$ and $\mathcal{F}D = -\tilde{D}\mathcal{F}$, and moreover $\mathcal{F}$ is a unitary (and so an isometric) operator on $\Gamma_c(\mathbf{S})$ and $L^2(\mathbf{Q})$, since it is defined fiberwise and it is so on the fibers, $\mathcal{S}(\mathbb{R}^n)$ or $L^2(\mathbb{R}^n)$. Therefore, all powers of $\mathcal{F}$ are also unitary, and since by Proposition 7.2.3 in Habermann and Habermann [54] $\mathcal{F}$ preserves the subbundles $\mathbf{Q}_\ell$ and satisfies $\mathcal{F}s_\ell = i^{-\ell}s_\ell$ on these subbundles, we also have $\mathcal{F}^4 = I$. Therefore, for all $s \in \Gamma_c(\mathbf{S})$,

$$\mathcal{F}(D + iI)s = (\tilde{D} - iI)(-\mathcal{F}s)$$

$$\mathcal{F}(\tilde{D} - iI)s = (D - iI)\mathcal{F}s$$

$$\mathcal{F}(D - iI)s = (\tilde{D} + iI)(-\mathcal{F}s)$$

$$\mathcal{F}(\tilde{D} + iI)s = (D + iI)\mathcal{F}s$$

which shows the first claim (in particular $\mathcal{F}^2(D+iI)s = (D-iI)(-\mathcal{F}^2 s)$ and $F^2(\tilde{D}+iI)s = (\tilde{D}+iI)(-\mathcal{F}^2 s)$ give the explicit isomorphisms $-\mathcal{F}^2 : \operatorname{im}(D+iI) \to \operatorname{im}(D-iI)$ and $-\mathcal{F}^2 : \operatorname{im}(\tilde{D}+iI) \to \operatorname{im}(\tilde{D}-iI)$).

Next, to show that

$$\operatorname{im}(D+iI)^\perp \cong \operatorname{im}(\tilde{D}-iI)^\perp \cong \operatorname{im}(D-iI)^\perp \cong \operatorname{im}(\tilde{D}+iI)^\perp$$

we use the fact that $\mathcal{F}$, and therefore every power of $\mathcal{F}$, is an isometry, i.e. preserves the inner product on $L^2(\mathbf{Q})$. For suppose that $s \in \operatorname{im}(D+iI)^\perp$. Then $((D+iI)\tau, s) = 0$ for all $\tau \in \Gamma_c(\mathbf{S})$, and applying $\mathcal{F}$ to the inner product gives

$$0 = ((D+iI)\tau, s) = (\mathcal{F}(D+iI)\tau, \mathcal{F}s) = -((\tilde{D}-iI)(\mathcal{F}\tau), \mathcal{F}s)$$

so $\mathcal{F}s \in \operatorname{im}(\tilde{D}-iI)^\perp$. Since $\mathcal{F}$ is invertible, we see that this defines an isomorphism between $\operatorname{im}(D+iI)^\perp$ and $\operatorname{im}(\tilde{D}-iI)^\perp$. The other isomorphisms follow similarly. ■

We now begin to investigate in detail the isomorphisms given above. We will see (Lemma 18 and Proposition 74) that, given any $t \in \ker(D^* - iI)$, if we write $t = t_0 + t_1 + t_2 + t_3$, where each $t_j \in \ker(\mathcal{F} - (-i)^j)$ is in one of the four eigenspaces of the Fourier transform, then we will have

$$t_0 + t_2,\ t_1 + t_3 \in \mathcal{D}(D^*)$$
$$t_0 - t_2,\ t_1 - t_3 \in \mathcal{D}(\tilde{D}^*)$$

and

$$t' = \mathcal{F}t = (t_0 - t_2) - i(t_1 - t_3) \in \ker(\tilde{D}^* + iI)$$
$$t'' = \mathcal{F}^2 t = (t_0 + t_2) - (t_1 + t_3) \in \ker(D^* + iI)$$
$$t''' = \mathcal{F}^3 t = (t_0 - t_2) + i(t_1 - t_3) \in \ker(\tilde{D} - iI)$$

with

$$\begin{aligned}
D^*(t_0+t_2) &= i(t_1+t_3)\\
D^*(t_1+t_3) &= i(t_0+t_2)\\
\tilde{D}^*(t_0-t_2) &= -(t_1-t_3)\\
\tilde{D}^*(t_1-t_3) &= t_0-t_2
\end{aligned}$$

and therefore

$$\begin{aligned}
&(t_0+t_2)-(t_1+t_3) \in \ker(D^*+iI)\\
&(t_0-t_2)+i(t_1+t_3) \in \ker(\tilde{D}^*-iI)\\
&(t_0-t_2)-i(t_1+t_3) \in \ker(\tilde{D}^*+iI)
\end{aligned}$$

This will completely characterize the relations between the pieces of the domains

$$\mathcal{D}(D^*) = \mathcal{D}(D) \oplus_T \ker(D^*-iI) \oplus_T \ker(D^*+iI)$$

and

$$\mathcal{D}(\tilde{D}^*) = \mathcal{D}(\tilde{D}) \oplus_T \ker(\tilde{D}^*-iI) \oplus_T \ker(\tilde{D}^*+iI)$$

in terms of eigensections of the Fourier transform. Since the Hermite functions are eigenfunctions of both the Fourier transform and the harmonic oscillator, these relations can be more succintly stated as follows: Writing $t = \sum_{\ell\in\mathbb{N}_0} \tau_\ell$, where $\tau_\ell \in \Gamma(\mathbf{S}_\ell)$, we will have

$$\begin{aligned}
t_0+t_2 &= \sum_{\ell=2k} \tau_\ell\\
t_1+t_3 &= \sum_{\ell=2k+1} \tau_\ell
\end{aligned}$$

From the above considerations, we see that, starting with $t \in \ker(D^*-iI)$, the operator D^* takes the even-dimensional pieces τ_ℓ of t to the odds and vice-versa (times i, of course), and similarly with $\tilde{D}^*$ except for the fact that we use t_0-t_2 and t_1-t_3 instead of t_0+t_2 and t_1+t_3. The starting position of $\ker(D^*-iI)$ was arbitrary. We could have started with $t \in \ker(D^*+iI)$ or

$t \in \ker(\tilde{D}^* \pm iI)$ and gotten similar results. Indeed, by relabeling the t_j this becomes patently clear: if $t_0 + t_1 + t_2 + t_3 \in \ker(D^* - iI)$, then for example $(t_0 - t_2) + i(t_1 - t_3) \in \ker(\tilde{D}^* - iI)$, so writing $s_0 = t_0$, $s_1 = it_1$, $s_2 = -t_2$ and $s_3 = -it_3$ we have $s = s_0 + s_1 + s_2 + s_3 \in \ker(\tilde{D}^* - iI)$, and the other elements of $\ker(\tilde{D}^* + iI)$ and $\ker(D^* \pm iI)$ are completely determined by s. Thus, it suffices to consider $t \in \ker(D^* - iI)$.

Following this description of $\mathcal{D}(\tilde{D}^*)$ and $\mathcal{D}(D^*)$, we will show that the description of $t \in \mathcal{D}(D^*)$ in terms of eigensections $\tau_\ell \in \Gamma(\mathbf{Q}_\ell)$ will in fact reveal that $t = 0$, thus showing that D^* and $\tilde{D}^*$ are essentially self-adjoint.

For the remainder of this section we declare the domains of D and $\tilde{D}$ to be the compactly supported Schwartz spinors,

$$\mathcal{D}(D) = \mathcal{D}(\tilde{D}) = \Gamma_c(\mathbf{S}) \tag{5.129}$$

We do not assume at first that D and $\tilde{D}$ are closed, but only symmetric on $\Gamma_c(\mathbf{S})$. Our ultimate goal in this section is to demonstrate essential self adjointness, by use of (2) of Corollary 17, showing that $\ker(D^* \pm iI) = \ker(\tilde{D}^* \pm iI) = \{0\}$ (which will require, of course, that D and $\tilde{D}$ be closed, but we will make this necessary adjustment when it is needed). We begin with some preliminary observations.

Lemma 15 *The relations (5.100)-(5.101), which give $\mathcal{F}\tilde{D} = D\mathcal{F}$ and $\mathcal{F}D = -\tilde{D}\mathcal{F}$ on $\Gamma_c(\mathbf{S})$, extend to D^* and $\tilde{D}^*$,*

$$\mathcal{F}\tilde{D}^* = D^*\mathcal{F} \tag{5.130}$$

$$\mathcal{F}D^* = -\tilde{D}^*\mathcal{F} \tag{5.131}$$

on the domains $\mathcal{D}(D^*)$ and $\mathcal{D}(\tilde{D}^*)$, which consequently shows that these domains are isometric (via $\mathcal{F}$),

$$\mathcal{D}(D^*) \cong \mathcal{D}(\tilde{D}^*) \tag{5.132}$$

More precisely

$$\mathcal{F}(\mathcal{D}(D^*)) = \mathcal{D}(\tilde{D}^*) \tag{5.133}$$

$$\mathcal{F}(\mathcal{D}(\tilde{D}^*)) = \mathcal{D}(D^*) \tag{5.134}$$

Proof: We will prove this directly from the definitions of $\mathcal{D}(D^*)$ and $\mathcal{D}(\tilde{D}^*)$. Recall,

$$\mathcal{D}(D^*) = \{t \in L^2(\mathbf{Q}) \mid \forall s \in \mathcal{D}(D),\ \exists \tau \in L^2(\mathbf{Q}) \text{ such that } (s,\tau) = (Ds,t)\}$$

and in this case we write $D^*t = \tau$. Suppose that $t \in \mathcal{D}(D^*)$. To show that $\mathcal{F}t \in \mathcal{D}(\tilde{D}^*)$ we need to show that for all $\sigma \in \mathcal{D}(\tilde{D})$ there exists $\tau \in L^2(\mathbf{Q})$ such that $(\sigma,\tau) = (\tilde{D}\sigma, \mathcal{F}t)$. We will need the following facts:

(1) $\mathcal{D}(D) = \mathcal{D}(\tilde{D}) = \Gamma_c(\mathbf{S})$, by definiton of D and $\tilde{D}$.

(2) $\mathcal{F}\tilde{D} = D\mathcal{F}$ and $\mathcal{F}D = -\tilde{D}\mathcal{F}$ on $\mathcal{D}(D)$, which are the known relations (5.100)-(5.101). We can of course rewrite these as $\mathcal{F}^{-1}D = \tilde{D}\mathcal{F}^{-1}$ and $\mathcal{F}^{-1}\tilde{D} = -D\mathcal{F}^{-1}$.

(3) $\mathcal{F}(\Gamma_c(\mathbf{S})) = \Gamma_c(\mathbf{S})$, which is a consequence of the invertibility of $\mathcal{F}$ on $\mathcal{S}(\mathbb{R}^n)$.

(4) $\mathcal{F}$ is a unitary transformation of $L^2(\mathbf{Q})$, because it is so fiberwise on $L^2(\mathbb{R}^n)$.

Now, if $t \in \mathcal{D}(D^*)$, then for all $s \in \mathcal{D}(D)$ there exists $r = D^*t \in L^2(\mathbf{Q})$ such that $(s, D^*t) = (s,r) = (Ds,t)$. By fact (3), there is a unique $\sigma = \mathcal{F}s \in \mathcal{D}(D)$, so that we can write $(\mathcal{F}^{-1}\sigma, D^*t) = (s, D^*t)$. Combining these and using again the unitarity of $\mathcal{F}$, fact (4), we have

$$(\sigma, \mathcal{F}D^*t) = (\mathcal{F}^{-1}\sigma, D^*t) = (s, D^*t) = (Ds,t) \tag{5.135}$$

Using facts (2) and (4) again, we then get

$$(Ds,t) = (D\mathcal{F}^{-1}\sigma, t) = (-\mathcal{F}^{-1}\tilde{D}\sigma, t) \tag{5.136}$$

Now, by unitarity again, we have that $t = \mathcal{F}^{-1}t'$ for a unique $t' := \mathcal{F}t \in L^2(\mathbf{Q})$, so

$$(\mathcal{F}^{-1}\tilde{D}\sigma, t) = (\mathcal{F}^{-1}\tilde{D}\sigma, \mathcal{F}^{-1}t') = (\tilde{D}\sigma, t') = (\tilde{D}\sigma, \mathcal{F}t) \tag{5.137}$$

where the second equality follows from the fact that $\mathcal{F}^{-1}$ is unitary. Putting together equations (5.135)-(5.137) we see that for all $\sigma \in \mathcal{D}(\tilde{D})$ there exists a $\tau \in L^2(\mathbf{Q})$, namely $\tau = -\mathcal{F}D^*t$, such that

$$(\sigma, \tau) = (\sigma, -\mathcal{F}D^*t) = (\tilde{D}\sigma, t') = (\tilde{D}\sigma, \mathcal{F}t) \tag{5.138}$$

This shows that $\mathcal{F}(\mathcal{D}(D^*)) \subseteq \mathcal{D}(\tilde{D}^*)$. Since $\mathcal{F}$ is an isometry, we conclude that this is in fact an equality,

$$\mathcal{F}(\mathcal{D}(D^*)) = \mathcal{D}(\tilde{D}^*)$$

(Though this fact also follows from the following inclusion, $\mathcal{F}(\mathcal{D}(\tilde{D}^*)) \subseteq \mathcal{D}(D^*)$, noting that $\mathcal{F}^2$ is the parity operator, which is an isometry of $L^2(\mathbf{Q})$. It can also be proved by the analogous inclusions with $\mathcal{F}$ replaced by $\mathcal{F}^{-1}$—the proofs are entirely analogous.)

Next, we show the equality

$$\mathcal{F}(\mathcal{D}(\tilde{D}^*)) = \mathcal{D}(D^*)$$

by exactly the same means. If $t \in \mathcal{D}(\tilde{D}^*)$, then for all $s = \mathcal{F}^{-1}\sigma \in \mathcal{D}(D)$ there exists a $\tau = \tilde{D}^*t \in L^2(\mathbf{Q})$ such that $(s, \tilde{D}^*t) = (s, \tau) = (\tilde{D}s, t)$. Proceeding as above, we get that for all $\sigma = \mathcal{F}s \in \mathcal{D}(D)$ there exists a $\tau' := \mathcal{F}\tilde{D}^*t \in L^2(\mathbf{Q})$ such that

$$(\sigma, \tau') = (\sigma, \mathcal{F}\tilde{D}^*t) = (\mathcal{F}^{-1}\sigma, \tilde{D}^*t) = (s, \tilde{D}^*t) = (\tilde{D}s, t)$$

$$(\tilde{D}\mathcal{F}^{-1}\sigma, \mathcal{F}^{-1}t') = (\mathcal{F}^{-1}D\sigma, \mathcal{F}^{-1}t') = (D\sigma, t') = (D\sigma, \mathcal{F}t)$$

where $t' := \mathcal{F}t$.

As a byproduct of the above proofs, we also see that if $t \in \mathcal{D}(D^*)$, then $-\mathcal{F}D^*t = \tilde{D}^*\mathcal{F}t$, while if $t \in \mathcal{D}(\tilde{D}^*)$, then $\mathcal{F}\tilde{D}^*t = D^*\mathcal{F}t$, thus establishing the relations (5.130) and (5.131). ■

Remark 64 Since $\mathcal{F}$ is a unitary operator on $L^2(\mathbf{Q})$, we have immediately that the relations (5.130)-(5.131) can be restaed as

$$\tilde{D}^*\mathcal{F}^{-1} = \mathcal{F}^{-1}D^* \tag{5.139}$$

$$D^*\mathcal{F}^{-1} = -\mathcal{F}^{-1}\tilde{D}^* \tag{5.140}$$

and consequently we also have

$$\mathcal{F}^{-1}(\mathcal{D}(D^*)) = \mathcal{D}(\tilde{D}^*) \tag{5.141}$$

$$\mathcal{F}^{-1}(\mathcal{D}(\tilde{D}^*) = \mathcal{D}(D^*) \tag{5.142}$$

We remark that these statements could, of course, be proved directly using the same sort of arguments as in the proof of Lemma 15. ■

We restate here Lemma 14 for the convenience of the reader, because we will make heavy use of its results in the following pages.

Lemma 16 *The four eigenspaces of the Fourier transform $\mathcal{F}$ on $L^2(\mathbf{Q})$ defined in (5.44) above give an orthogonal decomposition of $L^2(\mathbf{Q})$,*

$$L^2(\mathbf{Q}) = \bigoplus_{j=0}^{3} \ker(\mathcal{F} - (-i)^j I) \tag{5.143}$$

and the jth orthogonal projection onto $\ker(\mathcal{F} - (-i)^j I)$ is given by

$$P_j = \frac{1}{4}\sum_{k=0}^{3}(i^j \mathcal{F})^k \tag{5.144}$$

■

Lemma 17 *If $t \in \mathcal{D}(D^*)$ and $t_j \in E_j := \ker(\mathcal{F} - (-i)^j I)$, as in the last lemma, then $t_0 + t_2,\ t_1 + t_3 \in \mathcal{D}(D^*)$ and $t_0 - t_2,\ t_1 - t_3 \in \mathcal{D}(\tilde{D}^*)$. Similarly, if $t = t_0 + t_1 + t_2 + t_3 \in \mathcal{D}(\tilde{D}^*)$, then $t_0 + t_2, t_1 + t_3 \in \mathcal{D}(\tilde{D}^*)$ and $t_0 - t_2, t_1 - t_3 \in \mathcal{D}(D^*)$.*

Proof: (1) First, let us show that if $s = s_0 + s_1 + s_2 + s_3 \in \mathcal{D}(D)$, where $s_j \in \ker(\mathcal{F} - (-i)^j)$, then

$$(Zs)_j = Z(s_{(j+1) \bmod 4}) \tag{5.145}$$

$$(Z^*s)_j = Z^*(s_{(j-1) \bmod 4}) \tag{5.146}$$

That is,

$$Zs_0 \in E_3 := \ker(\mathcal{F} - (-i)^3 I), \qquad Z^* s_0 \in E_1 := \ker(\mathcal{F} - (-i)^1 I),$$

$$Zs_1 \in E_0 := \ker(\mathcal{F} - (-i)^0 I), \qquad Z^* s_1 \in E_2 := \ker(\mathcal{F} - (-i)^2 I),$$

$$Zs_2 \in E_1 := \ker(\mathcal{F} - (-i)^1 I), \qquad Z^* s_2 \in E_3 := \ker(\mathcal{F} - (-i)^3 I),$$

$$Zs_3 \in E_2 := \ker(\mathcal{F} - (-i)^2 I), \qquad Z^* s_3 \in E_0 := \ker(\mathcal{F} - (-i)^0 I)$$

But these follow directly from the relations (5.104) and (5.105): $\mathcal{F}Z = iZ\mathcal{F}$ and $\mathcal{F}Z^* = -iZ^*\mathcal{F}$, for

$$\mathcal{F}Zs_j = iZ\mathcal{F}s_j = i(-i)^j Zs_j = -(-i)^{j+1} Zs_j = (-i)^{j+3} Zs_j = (-i)^{j-1} Zs_j$$

and

$$\mathcal{F}Z^* s_j = -iZ^*\mathcal{F}s_j = -i(-i)^j Z^* s_j = (-i)^{j+1} Z^* s_j$$

Modulo 4, these are precisely the statements (5.145) and (5.146) above.

(2) Consequently, since $D = \frac{1}{2}(Z + Z^*)$ and $\tilde{D} = \frac{1}{2i}(Z - Z^*)$, we have

$$Ds_0,\ \tilde{D}s_0 \in E_1 \oplus E_3 \tag{5.147}$$

$$Ds_1,\ \tilde{D}s_1 \in E_0 \oplus E_2 \tag{5.148}$$

$$Ds_2,\ \tilde{D}s_2 \in E_1 \oplus E_3 \tag{5.149}$$

$$Ds_3,\ \tilde{D}s_3 \in E_0 \oplus E_2 \tag{5.150}$$

More explicitly, writing $Ds = (Ds)_0 + (Ds)_1 + (Ds)_2 + (Ds)_3$ and using the above properties of Z and Z^*, we have

$$\begin{aligned}
(Ds)_0+&(Ds)_1 + (Ds)_2 + (Ds)_3 \\
&= Ds \\
&= D(s_0 + s_1 + s_2 + s_3) \\
&= Ds_0 + Ds_1 + Ds_2 + Ds_3 \\
&= \frac{1}{2}(Zs_0 + Z^*s_0) + \frac{1}{2}(Zs_1 + Z^*s_1) + \frac{1}{2}(Zs_2 + Z^*s_2) + \frac{1}{2}(Zs_3 + Z^*s_3) \\
&= \frac{1}{2}(Zs_1 + Z^*s_3) + \frac{1}{2}(Zs_2 + Z^*s_0) + \frac{1}{2}(Zs_3 + Z^*s_1) + \frac{1}{2}(Zs_0 + Z^*s_2)
\end{aligned}$$

which shows that

$$\begin{aligned}
(Ds)_0 &= \frac{1}{2}(Zs_1 + Z^*s_3) = \frac{1}{2}D(s_1 + s_3) + \frac{i}{2}\tilde{D}(s_1 - s_3) \\
(Ds)_1 &= \frac{1}{2}(Zs_2 + Z^*s_0) = \frac{1}{2}D(s_0 + s_2) - \frac{i}{2}\tilde{D}(s_0 - s_2) \\
(Ds)_2 &= \frac{1}{2}(Zs_3 + Z^*s_1) = \frac{1}{2}D(s_1 + s_3) - \frac{i}{2}\tilde{D}(s_1 - s_3) \\
(Ds)_3 &= \frac{1}{2}(Zs_0 + Z^*s_2) = \frac{1}{2}D(s_0 + s_2) + \frac{i}{2}\tilde{D}(s_0 - s_2)
\end{aligned}$$

This in turn gives us the more useful relations

$$(Ds)_0 + (Ds)_2 = Ds_1 + Ds_3 \tag{5.151}$$

$$(Ds)_1 + (Ds)_3 = Ds_0 + Ds_2 \tag{5.152}$$

Similarly,

$$\begin{aligned}
(\tilde{D}s)_0+&(\tilde{D}s)_1 + (\tilde{D}s)_2 + (\tilde{D}s)_3 \\
&= \tilde{D}s \\
&= \tilde{D}(s_0 + s_1 + s_2 + s_3) \\
&= \tilde{D}s_0 + \tilde{D}s_1 + \tilde{D}s_2 + \tilde{D}s_3 \\
&= \frac{1}{2i}(Zs_0 - Z^*s_0) + \frac{1}{2i}(Zs_1 - Z^*s_1) + \frac{1}{2i}(Zs_2 - Z^*s_2) + \frac{1}{2i}(Zs_3 - Z^*s_3) \\
&= \frac{1}{2i}(Zs_1 - Z^*s_3) + \frac{1}{2i}(Zs_2 - Z^*s_0) + \frac{1}{2i}(Zs_3 - Z^*s_1) + \frac{1}{2i}(Zs_0 - Z^*s_2)
\end{aligned}$$

which shows that

$$
\begin{aligned}
(\tilde{D}s)_0 &= \frac{1}{2}(Zs_1 - Z^* s_3) = -\frac{i}{2}D(s_1 - s_3) + \frac{1}{2}\tilde{D}(s_1 + s_3)\\
(\tilde{D}s)_1 &= \frac{1}{2}(Zs_2 - Z^* s_0) = \frac{i}{2}D(s_0 - s_2) + \frac{1}{2}\tilde{D}(s_0 + s_2)\\
(\tilde{D}s)_2 &= \frac{1}{2}(Zs_3 - Z^* s_1) = \frac{i}{2}D(s_1 - s_3) + \frac{1}{2}\tilde{D}(s_1 + s_3)\\
(\tilde{D}s)_3 &= \frac{1}{2}(Zs_0 - Z^* s_2) = -\frac{i}{2}D(s_0 - s_2) + \frac{1}{2}\tilde{D}(s_0 + s_2)
\end{aligned}
$$

and from these we get

$$(\tilde{D}s)_0 + (\tilde{D}s)_2 = \tilde{D}s_1 + \tilde{D}s_3 \tag{5.153}$$

$$(\tilde{D}s)_1 + (\tilde{D}s)_3 = \tilde{D}s_0 + \tilde{D}s_2 \tag{5.154}$$

(3) We now want to use the above equations (5.147)-(5.150) to show that if $t = t_0+t_1+t_2+t_3 \in \mathcal{D}(D^*)$, then $t_0 + t_2,\ t_1 + t_3 \in \mathcal{D}(D^*)$ and $t_0 - t_2,\ t_1 - t_3 \in \mathcal{D}(\tilde{D}^*)$. Now, if $t \in \mathcal{D}(D^*)$, then there is a $\tau =: D^*t \in L^2(\mathbf{Q})$ such that for all $s = s_0 + s_1 + s_2 + s_3$ we have $(Ds_j, t) = (s_j, \tau) = (s_j, \tau_j)$, $j \in \{0, 1, 2, 3\}$. On the other hand, from (5.147)-(5.150) we know that

$$
\begin{aligned}
(Ds_0, t) &= (Ds_0, t_1 + t_3)\\
(Ds_1, t) &= (Ds_0, t_0 + t_2)\\
(Ds_2, t) &= (Ds_0, t_1 + t_3)\\
(Ds_3, t) &= (Ds_0, t_0 + t_2)
\end{aligned}
$$

so for all $s \in \mathcal{D}(D)$ we get

$$
\begin{aligned}
(Ds, t_1 + t_3) &= (Ds_0, t_1 + t_3) + (Ds_2, t_1 + t_3)\\
&= (Ds_0, t) + (Ds_2, t)\\
&= (s_0, \tau_0) + (s_2, \tau_2)\\
&= (s, \tau_0) + (s, \tau_2)\\
&= (s, \tau_0 + \tau_2)
\end{aligned}
$$

and

$$\begin{aligned}(Ds, t_0+t_2) &= (Ds_1, t_0+t_2)+(Ds_3, t_0+t_2)\\ &= (Ds_1, t)+(Ds_3, t)\\ &= (s_1, \tau_1)+(s_3, \tau_3)\\ &= (s, \tau_1)+(s, \tau_3)\\ &= (s, \tau_1+\tau_3)\end{aligned}$$

Thus, t_0+t_2 and t_1+t_3 lie in $\mathcal{D}(D^*)$, and $D^*(t_1+t_3) = (D^*t)_0+(D^*t)_2$ and $D^*(t_0+t_2) = (D^*t)_1+(D^*t)_3$. From the relation $\mathcal{F}\tilde{D}^* = D^*\mathcal{F}$, we have that $t \in \mathcal{D}(D^*)$ iff $\mathcal{F}t \in \mathcal{D}(\tilde{D}^*)$. Therefore since $t_0+t_2,\ t_1+t_3 \in \mathcal{D}(D^*)$ we must have $t_0-t_2 = \mathcal{F}(t_0+t_2),\ t_1-t_3 = i\mathcal{F}(t_1+t_3) \in \mathcal{D}(\tilde{D}^*)$.

(4) Similarly, if $t = t_0+t_1+t_2+t_3 \in \mathcal{D}(\tilde{D}^*)$, then $t_0+t_2, t_1+t_3 \in \mathcal{D}(\tilde{D}^*)$ and $t_0-t_2, t_1-t_3 \in \mathcal{D}(D^*)$, for, since $(\tilde{D}s, t) = (s, \tau)$ for all $s \in \mathcal{D}(\tilde{D})$, we have

$$(\tilde{D}s_0, t) = (\tilde{D}s_0, t_1+t_3)$$

$$(\tilde{D}s_1, t) = (\tilde{D}s_0, t_0+t_2)$$

$$(\tilde{D}s_2, t) = (\tilde{D}s_0, t_1+t_3)$$

$$(\tilde{D}s_3, t) = (\tilde{D}s_0, t_0+t_2)$$

so for all $s \in \mathcal{D}(\tilde{D})$ we get

$$\begin{aligned}(\tilde{D}s, t_1+t_3) &= (\tilde{D}s_0, t_1+t_3)+(\tilde{D}s_2, t_1+t_3)\\ &= (\tilde{D}s_0, t)+(\tilde{D}s_2, t)\\ &= (s_0, \tau_0)+(s_2, \tau_2)\\ &= (s, \tau_0)+(s, \tau_2)\\ &= (s, \tau_0+\tau_2)\end{aligned}$$

and

$$
\begin{aligned}
(\tilde{D}s, t_0 + t_2) &= (\tilde{D}s_1, t_0 + t_2) + (\tilde{D}s_3, t_0 + t_2) \\
&= (\tilde{D}s_1, t) + (\tilde{D}s_3, t) \\
&= (s_1, \tau_1) + (s_3, \tau_3) \\
&= (s, \tau_1) + (s, \tau_3) \\
&= (s, \tau_1 + \tau_3)
\end{aligned}
$$

This completes the proof. ■

Lemma 18 *Let $t = t_0 + t_1 + t_2 + t_3 \in \mathcal{D}(D^*)$, where $t_j \in E_j := \ker(\mathcal{F} - (-i)^j I)$. By the previous lemma we know that $t_0 + t_2,\ t_1 + t_3 \in \mathcal{D}(D^*)$ and $t_0 - t_2,\ t_1 - t_3 \in \mathcal{D}(\tilde{D}^*)$, and we claim that if $t \in \ker(D^* - iI)$, then*

$$D^*(t_0 + t_2) = i(t_1 + t_3) \tag{5.155}$$

$$D^*(t_1 + t_3) = i(t_0 + t_2) \tag{5.156}$$

$$\tilde{D}^*(t_0 - t_2) = -(t_1 - t_3) \tag{5.157}$$

$$\tilde{D}^*(t_1 - t_3) = t_0 - t_2 \tag{5.158}$$

Proof: The first two equalities follow directly from part (3) of the proof of the previous lemma, which shows that $D^*(t_1 + t_3) = (D^*t)_0 + (D^*t)_2$ and $D^*(t_0 + t_2) = (D^*t)_1 + (D^*t)_3$, combined with the fact that $D^*t = t$. To see (5.157), note that by Lemma 15 we have $\mathcal{F}\tilde{D}^* = D\mathcal{F}$ and $\mathcal{F}D = -\tilde{D}\mathcal{F}$, and consequently $\mathcal{F}(\mathcal{D}(D^*)) = \mathcal{D}(\tilde{D}^*)$ and $\mathcal{F}(\mathcal{D}(\tilde{D}^*)) = \mathcal{D}(D^*)$. Combined with equations (5.155) and (5.156) we have

$$
\begin{aligned}
\tilde{D}^*(t_0 - t_2) &= \tilde{D}^*\mathcal{F}(t_0 + t_2) \\
&= -\mathcal{F}D^*(t_0 + t_2) \\
&= -i\mathcal{F}(t_1 + t_3) \\
&= -i(-it_1 + it_3) \\
&= -(t_1 - t_3)
\end{aligned}
$$

Similarly, (5.158) follows from

$$\begin{aligned}\tilde{D}^*(t_1 - t_3) &= i\tilde{D}^*(-it_1 + it_3)\\ &= i\tilde{D}\mathcal{F}(t_1 + t_3)\\ &= -i\mathcal{F}D^*(t_1 + t_3)\\ &= (-i)i\mathcal{F}(t_0 + t_2)\\ &= t_0 - t_2\end{aligned}$$

which completes the proof. ∎

Remark 65 We may now consider the *closures* $\overline{D}$ and $\overline{\tilde{D}}$ of D and $\tilde{D}$, respectively. Since $D \subseteq \overline{D} \subseteq D^*$ and $\tilde{D} \subseteq \overline{\tilde{D}} \subseteq \tilde{D}^*$, the relations

$$\mathcal{F}\overline{\tilde{D}} = \overline{D}\mathcal{F} \tag{5.159}$$

$$\mathcal{F}\overline{D} = -\overline{\tilde{D}}\mathcal{F} \tag{5.160}$$

and consequently

$$\mathcal{F}(\mathcal{D}(\overline{\tilde{D}})) = \mathcal{D}(\overline{D}) \tag{5.161}$$

$$\mathcal{F}(\mathcal{D}(\overline{D})) = \mathcal{D}(\overline{\tilde{D}}) \tag{5.162}$$

hold by our previous considerations. Hence, in what follows we may assume that D and $\tilde{D}$ are closed symmetric operators. This will be needed in order to use the decompositions of $\mathcal{D}(D^*)$ and $\mathcal{D}(\tilde{D}^*)$ provided by Proposition 66,

$$\mathcal{D}(D^*) = \mathcal{D}(D) \oplus_D \ker(D^* - iI) \oplus_D \ker(D^* + iI)$$

$$\mathcal{D}(\tilde{D}^*) = \mathcal{D}(\tilde{D}) \oplus_{\tilde{D}} \ker(\tilde{D}^* - iI) \oplus_{\tilde{D}} \ker(\tilde{D}^* + iI)$$

We note, therefore, that if $t \in \mathcal{D}(D^*)$, then by Lemma 17 we have that $t_0 + t_2,\ t_1 + t_3 \in \mathcal{D}(D^*)$ and $t_0 - t_2,\ t_1 - t_3 \in \mathcal{D}(\tilde{D}^*)$, so

$$t_0 + t_2 = a + b + c \in \mathcal{D}(D^*) \tag{5.163}$$

$$t_1 + t_3 = d + e + f \in \mathcal{D}(D^*) \tag{5.164}$$

for some $a, d \in \mathcal{D}(D)$, $b, e \in \ker(D^* - iI)$ and $c, f \in \ker(D^* + iI)$, and similarly

$$t_0 - t_2 = \alpha + \beta + \gamma \tag{5.165}$$

$$t_1 - t_3 = \delta + \epsilon + \eta \tag{5.166}$$

for some $\alpha, \delta \in \mathcal{D}(\tilde{D})$, $\beta, \epsilon \in \ker(\tilde{D}^* - iI)$ and $\gamma, \eta \in \ker(D^* + iI)$. ■

When additionally, $t \in \ker(D^* - iI)$, we have:

Proposition 74 *If $t \in \ker(D^* - iI)$, then in the decompositions (5.163)-(5.166) above we have $a = 0$, $b = e$ and $c = -f$, and analogously $\alpha = 0$, $\delta = -\tilde{D}\alpha$, $\epsilon = -i\beta$ and $\eta = i\gamma$, so that $t_0 + t_2 = b + c$, $t_1 + t_3 = b - c$, and $t_0 - t_2 = \beta + \gamma$, $t_1 - t_3 = -i\beta + i\gamma$. Moreover, $b = \frac{1}{2}t \in \ker(D^* - iI)$, $c = \frac{1}{2}(t_0 + t_2 - (t_1 + t_3)) \in \ker(D^* + iI)$, and $\beta = \frac{1}{2}(t_0 - t_2 + i(t_1 - t_3)) \in \ker(\tilde{D}^* - iI)$, $\gamma = \frac{1}{2}(t_0 - t_2 - i(t_1 - t_3)) \in \ker(\tilde{D}^* + iI)$:*

$$t_0 + t_2 = b + c = \frac{1}{2}\big(t_0 + t_2 + t_1 + t_3\big) + \frac{1}{2}\big(t_0 + t_2 - (t_1 + t_3)\big) \tag{5.167}$$

$$t_1 + t_3 = b - c = \frac{1}{2}\big(t_0 + t_2 + t_1 + t_3\big) - \frac{1}{2}\big(t_0 + t_2 - (t_1 + t_3)\big) \tag{5.168}$$

$$t_0 - t_2 = \beta + \gamma = \frac{1}{2}\big(t_0 - t_2 + i(t_1 - t_3)\big) + \frac{1}{2}\big(t_0 - t_2 - i(t_1 - t_3)\big) \tag{5.169}$$

$$t_1 - t_3 = -i\beta + i\gamma = -\frac{i}{2}\big(t_0 - t_2 + i(t_1 - t_3)\big) + \frac{i}{2}\big(t_0 - t_2 - i(t_1 - t_3)\big) \tag{5.170}$$

Proof: Assume the decompositions (5.163)-(5.166) above. By Lemma 18 we know that

$$i(d + e + f) = i(t_1 + t_3) = D^*(t_0 + t_2) = D^*(a + b + c) = Da + ib - ic$$

so that by uniqueness $d = -iDa$, $e = b$ and $f = -c$, that is

$$t_0 + t_2 = a + b + c$$

$$t_1 + t_3 = -iDa + b - c$$

Now, we know that $D^*t = it$. If we write $\tilde{t} = t_0 + t_2 - (t_1 + t_3)$, then we also have, by the previous lemma,

$$\begin{aligned} D^*\tilde{t} &= D^*(t_0 + t_2) - D^*(t_1 + t_3) \\ &= i(t_1 + t_3) - i(t_0 + t_2) \\ &= -i\big(t_0 + t_2 - (t_1 + t_3)\big) \\ &= -i\tilde{t} \end{aligned}$$

Now, writing $t_0 + t_2 = a + b + c$ and $t_1 + t_3 = -iDa + b - c$, we can see that

$$\begin{aligned} (ia + Da) + 2ib &= it \\ &= D^*t \\ &= D^*(t_0 + t_2) + D^*(t_1 + t_3) \\ &= D^*(a + b + c) + D^*(-iDa + b - c) \\ &= (Da - iD^2a) + 2ib \end{aligned}$$

and

$$\begin{aligned} (-ia + Da) - 2ic &= -i\tilde{t} \\ &= D^*\tilde{t} \\ &= D^*(t_0 + t_2) - D^*(t_1 + t_3) \\ &= D^*(a + b + c) - D^*(-iDa + b - c) \\ &= (Da + iD^2a) - 2ic \end{aligned}$$

so that $ia + Da = Da - iD^2a$ and $-ia + Da = Da + iD^2a$, which show that $a = -D^2a$. But since $t \in \ker(D^* - iI)$ and $\tilde{t} \in \ker(D^* + iI)$, we must have $ia + Da = 0 = -ia + Da$, so that $ia = -Da = -ia$, or $a = 0$. We thus have

$$\begin{aligned} t_0 + t_2 &= b + c \\ t_1 + t_3 &= b - c \end{aligned}$$

for unique $b \in \ker(D^* - iI)$ and $c \in \ker(D^* + iI)$. But we have $t \in \ker(D^* - iI)$ and $\tilde{t} \in \ker(D^* + iI)$, and

$$\frac{1}{2}t + \frac{1}{2}\tilde{t} = \frac{1}{2}\Big((t_0 + t_2 + t_1 + t_3) + (t_0 + t_2 - (t_1 + t_3)\Big) = t_0 + t_2$$
$$\frac{1}{2}t - \frac{1}{2}\tilde{t} = \frac{1}{2}\Big((t_0 + t_2 + t_1 + t_3) - (t_0 + t_2 - (t_1 + t_3)\Big) = t_1 + t_3$$

so we see that in fact $b = \frac{1}{2}t$ and $c = \frac{1}{2}\tilde{t}$.

Now consider $t_0 - t_2 = \alpha + \beta + \gamma$ and $t_1 - t_3 = \delta + \epsilon + \eta$, and note that by Lemma 17 we have

$$\alpha + \beta + \gamma = t_0 - t_2 = \tilde{D}^*(t_1 - t_3) = \tilde{D}^*(\delta + \epsilon + \eta) = D\delta + i\epsilon - i\eta$$

and similarly

$$\delta + \epsilon + \eta = t_1 - t_3 = -\tilde{D}^*(t_0 - t_2) = -\tilde{D}^*(\alpha + \beta + \gamma) = -\tilde{D}\alpha - i\beta + i\gamma$$

which shows that $\delta = -\tilde{D}\alpha$, $\epsilon = -i\beta$ and $\eta = i\gamma$, and also that $\tilde{D}^2\alpha = -\tilde{D}\delta = -\alpha$. Consequently,

$$t_0 - t_2 = \alpha + \beta + \gamma$$

$$t_1 - t_3 = -\tilde{D}\alpha - i\beta + i\gamma$$

Now, on the one hand

$$\mathcal{F}(\alpha + \beta + \gamma) = \mathcal{F}(t_0 - t_2) = t_0 + t_2 = \mathcal{F}^{-1}(t_0 - t_2) = \mathcal{F}^{-1}(\alpha + \beta + \gamma)$$

and on the other, since $t_0 + t_2 = b + c$ and $\mathcal{F}$ and $\mathcal{F}^{-1}$ provide isomorphisms between $\ker(D^* \pm iI)$ and $\ker(\tilde{D} \pm iI)$ (by Theorem 51, or alternatively via Lemma 15), we see that $\mathcal{F}\alpha = \mathcal{F}^{-1}\alpha = a = 0$, and therefore $\alpha = 0$. Thus,

$$t_0 - t_2 = \beta + \gamma$$

$$t_1 - t_3 = -i\beta + i\gamma$$

Finally, note that if we take $\beta = \frac{1}{2}(t_0 - t_2 + i(t_1 - t_3))$, then $\beta \in \ker(\tilde{D}^* - iI)$, for by Lemma 18

$$\begin{aligned}
\tilde{D}^*\beta &= \frac{1}{2}\tilde{D}^*\big(t_0 - t_2 + i(t_1 - t_3)\big) \\
&= \frac{1}{2}\big(-(t_1 - t_3) + i(t_0 + t_2)\big) \\
&= \frac{i}{2}\big(t_0 - t_2 + i(t_1 - t_3)\big) \\
&= i\beta
\end{aligned}$$

Similarly, if we take $\gamma = \frac{1}{2}\big(t_0 - t_2 - i(t_1 - t_3)\big)$, then $\gamma \in \ker(\tilde{D}^* + iI)$, for by Lemma 18

$$\begin{aligned}
\tilde{D}^*\gamma &= \frac{1}{2}\tilde{D}^*\big(t_0 - t_2 - i(t_1 - t_3)\big) \\
&= \frac{1}{2}\big(-(t_1 - t_3) - i(t_0 - t_2)\big) \\
&= -\frac{i}{2}\big(t_0 - t_2 - i(t_1 - t_3)\big) \\
&= -i\gamma
\end{aligned}$$

It is clear, too, that $\beta + \gamma = t_0 - t_2$ and $\beta - \gamma = t_1 - t_3$. This completes the proof. ■

5.5.0.1 Uniqueness of the Self-Adjoint Extensions of D and $\tilde{D}$: Essential Self-Adjointness

Consider now the operators D, $\tilde{D}$, Z and Z^* acting on $L^2(\mathbf{Q})$ weakly, that is by their adjoints. For example we say that $Zs := \tau$ *weakly* if for all $\sigma \in \Gamma_c(\mathbf{S})$ we have $(\tau, \sigma) = (s, Z\sigma)$. Note that if we write Z' for the adjoint of Z and $Z^{*\prime}$ for the adjoint of Z^*, then the fact that D and $\tilde{D}$ are symmetric shows that

$$\mathcal{D}(Z) \subseteq \mathcal{D}(Z^{*\prime}) \qquad \text{and} \qquad \mathcal{D}(Z^*) \subseteq \mathcal{D}(Z')$$

since for all $\sigma, \tau \in \Gamma_c(\mathbf{S})$ we have

$$(Z\sigma, \tau) = (D\sigma, \tau) + i(\tilde{D}\sigma, \tau) = (\sigma, D\tau) - (\sigma, i\tilde{D}\tau) = (\sigma, Z^*\tau)$$

Of course we know that

$$\mathcal{D}(D^*) \cap \mathcal{D}(\tilde{D}^*) = \mathcal{D}(D^* + \tilde{D}^*) \subseteq \mathcal{D}\big((D + \tilde{D})^*\big) = \mathcal{D}(Z') = \mathcal{D}(Z^{*\prime})$$

so for any $t \in \mathcal{D}(D^*)$ we know that $Z't$ and $Z^{*\prime}t$ are defined. Moreover, once we show that $Z : L^2(\mathbf{Q}_\ell) \to L^2(\mathbf{Q}_{\ell-1})$ and $Z^* : L^2(\mathbf{Q}_\ell) \to L^2(\mathbf{Q}_{\ell+1})$ in the weak sense (applied to elements in the domains of Z' and $Z^{*\prime}$), if we write $t = \sum_{\ell=0}^{\infty} \tau_\ell$ for $t \in \ker(D^* - iI) \subseteq \mathcal{D}(Z') = \mathcal{D}(Z^{*\prime})$, where $\tau_\ell \in L^2(\mathbf{Q}_\ell)$, then $Z\tau_\ell \in L^2(\mathbf{Q}_{\ell-1})$ and $Z^*\tau_\ell \in L^2(\mathbf{Q}_{\ell+1})$ in the weak sense.

We will show in the next lemma that it is possible to extend some of the relations between D, $\tilde{D}$, Z, Z^*, and other operators, particularly $\mathcal{H}$ and P, to relations between their weak extensions to $L^2(\mathbf{Q})$. There is reason to believe that such an extension is possible–we have already seen an example (Lemma 15) in which we extended the relations $\mathcal{F}D = -\tilde{D}\mathcal{F}$ and $\mathcal{F}\tilde{D} = D\mathcal{F}$ to D^* and $\tilde{D}^*$. The particular relations we are interested in extending are (5.50)-(5.51) and those of Theorem 47, Corollary 22, and Theorem 48. It will follow from this that $\ker Z' = \ker Z^{*\prime} \subseteq \ker P$, and that therefore, since P is elliptic, these kernels consist of smooth sections.

Lemma 19 *The adjoint operators Z' and $Z^{*\prime}$ also satisfy*

$$[\mathcal{H}, Z'] = Z' \tag{5.171}$$

$$[\mathcal{H}, Z^{*\prime}] = -Z^{*\prime} \tag{5.172}$$

which means that Z' and $Z^{*\prime}$ move eigenbundles in the same way as Z and Z^*,

$$Z : L^2(\mathbf{Q}_\ell) \to L^2(\mathbf{Q}_{\ell-1}) \tag{5.173}$$

$$Z^{*\prime} : L^2(\mathbf{Q}_\ell) \to L^2(\mathbf{Q}_{\ell+1}) \tag{5.174}$$

in the weak sense. Moreover, if $t = Z's = Z^*\tau \in \operatorname{im} Z' \cap \operatorname{im} Z^{*\prime}$ in the weak sense, then there are unique symplectic spinor fields $s', \tau' \in L^2(\mathbf{Q})$ such that $D^*s' = \tilde{D}^*\tau' = t' \in \operatorname{im} D \cap \operatorname{im} \tilde{D}$ weakly, and the pair (s, τ) is related to the pair (s', τ') by an invertible matrix

$$\begin{pmatrix} s' \\ \tau' \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -i & -i \end{pmatrix} \begin{pmatrix} s \\ \tau \end{pmatrix}$$

That is, $s' = s - \tau$, $\tau' = -is - i\tau$, and $s = \frac{1}{2}s' + \frac{i}{2}\tau'$, $\tau = -\frac{1}{2}s' + \frac{i}{2}\tau'$. As a result, we will have

$$\operatorname{im} D^* \cap \operatorname{im} \tilde{D}^* = \{0\} \tag{5.175}$$

Proof: (1) For all $\sigma \in \Gamma_c(\mathbf{S})$ and all $t \in \mathcal{D}(Z')$ such that $Z't \in \mathcal{D}(\mathcal{H})$ we have by the self-adjointness of $\mathcal{H}$

$$(\sigma, \mathcal{H}Z't) = (\mathcal{H}\sigma, Z't) = (Z\mathcal{H}\sigma, t) = ((\mathcal{H}Z + Z)\sigma, t) = (\sigma, Z'\mathcal{H}t) + (\sigma, Z't)$$

which shows that $[\mathcal{H}, Z'] = Z'$. $[\mathcal{H}, Z^{*\prime}] = -Z^{*\prime}$ follows similarly.

(2) Let $t = Zs = Z^*\tau$ weakly and define $s', \tau' \in L^2(\mathbf{Q})$ by $s' = s - \tau$, $\tau' = -is - i\tau$. Then, for all $\sigma \in \Gamma_c(\mathbf{S})$ we have

$$(t, \sigma) = (s, Z^*\sigma) = (\tau, Z\sigma)$$

which means $(s, D\sigma) + i(s, \tilde{D}\sigma) = (\tau, D\sigma) - i(\tau, \tilde{D}\sigma)$, or

$$(s', D\sigma) = (s - \tau, D\sigma) = (-is - i\tau), \tilde{D}\sigma)$$

for all $\sigma \in \Gamma_c(\mathbf{S})$, proving that $t' := Ds' = \tilde{D}\tau'$ weakly. Conversely, if $t' := Ds' = \tilde{D}\tau'$ weakly, then letting $s = \frac{1}{2}s' + \frac{i}{2}\tau'$ and $\tau = -\frac{1}{2}s' + \frac{i}{2}\tau'$, we have for all $\sigma \in \Gamma_c(\mathbf{S})$ that

$$\begin{aligned}
(s, Z^*\sigma) &= \frac{1}{2}(s', Z^*\sigma) + \frac{i}{2}(\tau', Z^*\sigma) \\
&= \frac{1}{2}(s', D\sigma) + \frac{i}{2}(s', \tilde{D}\sigma) + \frac{i}{2}(\tau', D\sigma) - \frac{1}{2}(\tau', \tilde{D}\sigma) \\
&= \frac{1}{2}(\tau', \tilde{D}\sigma) + \frac{i}{2}(s', \tilde{D}\sigma) + \frac{i}{2}(\tau', D\sigma) - \frac{1}{2}(s', D\sigma) \\
&= -\frac{1}{2}\big(s', (D - i\tilde{D})\sigma\big) + \frac{i}{2}\big(\tau', (D - i\tilde{D})\sigma\big) \\
&= -\frac{1}{2}(s', Z^*\sigma) + \frac{i}{2}(\tau', Z^*\sigma) \\
&= (\tau, Z^*\sigma)
\end{aligned}$$

so that $t := Zs = Z^*\tau$ weakly.

(3) Now, looking at the proofs of Corollary 22 and Theorem 48, which establish the identities $t' = Ds' = \tilde{D}\tau' = \frac{1}{2}\big(Z\tau - Z^*s\big)$ and $\operatorname{im} D \cap \operatorname{im} \tilde{D} = \{0\}$ (along with some others, which need not concern us here), we see that these follow strictly from the values of D, $\tilde{D}$, Z and Z^* along with the relation of the previous Theorem 47, whose results we have extended to $L^2(\mathbf{Q})$ in (1) and (2) above. Therefore, the same proofs of $t' = Ds' = \tilde{D}\tau' = \frac{1}{2}\big(Z\tau - Z^*s\big)$ and $\operatorname{im} D \cap \operatorname{im} \tilde{D} = \{0\}$ extend in a purely formal way to the extended operators D, $\tilde{D}$, Z and Z^* on $L^2(\mathbf{Q})$. ■

Consider the adjoints of D, $\tilde{D}$, Z, Z^*, and P. Since $P = i[\tilde{D}, D]$, we have the following relations among the domains of their adjoints:

$$\begin{aligned} \mathcal{D}(D^*) \cap \mathcal{D}(\tilde{D}^*) &\subseteq \mathcal{D}(D^* \tilde{D}^*) \cap \mathcal{D}(\tilde{D}^* D^*) \\ &= \mathcal{D}(\tilde{D}^* D^* - D^* \tilde{D}^*) \\ &\subseteq \mathcal{D}\big((\tilde{D}D - D\tilde{D})^*\big) \\ &= \mathcal{D}(P^*) \end{aligned}$$

By the previous lemma we conclude that:

Corollary 23 $\ker Z' = \ker Z^{*\prime} \subseteq \ker P^*$, *so since P is elliptic, elliptic regularity guarantees that weak solutions of $Zs = 0$ and $Z^*s = 0$ are in fact smooth. When (M, ω) is compact, these solutions are also compactly supported, and therefore lie in the domains of D, $\tilde{D}$, Z and Z^*.*

Proof: Let s be a weak solution of $Zs = 0$, i.e. $0 = (0, \sigma) = (s, Z\sigma)$ for all $\sigma \in \Gamma_c(\mathbf{S})$. Then $0 = (s, D\sigma) + (s, i\tilde{D}\sigma)$, or $(s, D\sigma) = (s, -i\tilde{D}\sigma)$, for all $\sigma \in \Gamma_c(\mathbf{S})$, which means $Ds = \tilde{D}(is)$ weakly. By the previous lemma, this must equal 0, so $s \in \ker D^* \cap \ker \tilde{D}^*$. Since $\mathcal{D}(\tilde{D}^* D^* - D^* \tilde{D}^*) \subseteq \mathcal{D}\big((\tilde{D}D - D\tilde{D})^*\big) = \mathcal{D}(P^*)$ and

$$i\tilde{D}^* D^* s - iD^* \tilde{D}^* s = 0$$

we conclude that $P^* s = 0$. ■

Theorem 52 *Let (M, ω) be a compact symplectic manifold. Then the symplectic Dirac operators D and $\tilde{D}$ are essentially self-adjoint on $L^2(\mathbf{Q})$.*

Proof: We will use Corollary 17, and demonstrate that $\ker(D^* \pm iI) = \ker(\tilde{D}^* \pm iI) = \{0\}$. By Theorem 51 and Proposition 74 it is enough to prove that $\ker(D^* - iI) = \{0\}$. We will use analogs of the results of Lemma 17, particularly (5.145)-(5.154) in that Lemma's proof.

Let $t \in \ker(D^* - iI)$ and note first that if we write $t = \sum_{\ell \in \mathbb{N}_0} \tau_\ell$, where each $\tau_\ell \in L^2(\mathbf{Q}_\ell)$, then in fact each τ_ℓ lies in $L^2(\mathbf{S}_\ell)$, since $\mathbf{S}_\ell$ is a finite-dimensional vector bundle of dimension $\binom{n+\ell-1}{\ell}$, spanned by $\binom{n+\ell-1}{\ell}$ basis Schwartz spinors (the Hermite-valued spinors). As we saw in the previous lemma, the operators Z and Z^* move eigensections of the harmonic oscillator $\mathcal{H}$ down and up by one, respectively,

$$Z : L^2(\mathbf{Q}_\ell) \to L^2(\mathbf{Q}_{\ell-1}), \qquad Z^* : L^2(\mathbf{Q}_\ell) \to L^2(\mathbf{Q}_{\ell+1})$$

Consequently, each τ_ℓ is moved by Z and Z^* accordingly,[2] so we have

$$\begin{aligned}
i\tau_0 &= (D^*\tau_1)_0 &&= \frac{1}{2}Z\tau_1 \\
i\tau_1 &= (D^*\tau_0)_1 + (D^*\tau_2)_1 &&= \frac{1}{2}Z^*\tau_0 + \frac{1}{2}Z\tau_2 \\
i\tau_2 &= (D^*\tau_1)_2 + (D^*\tau_3)_2 &&= \frac{1}{2}Z^*\tau_1 + \frac{1}{2}Z\tau_3 \\
i\tau_3 &= (D^*\tau_2)_3 + (D^*\tau_4)_3 &&= \frac{1}{2}Z^*\tau_2 + \frac{1}{2}Z\tau_4 \\
&\vdots
\end{aligned}$$

Now, by Lemma 19 we have $Z\tau_0 = 0$ weakly, because Z lowers indices. By Corollary 23 we conclude that $\tau_0 \in \Gamma(\mathbf{S}_0)$, i.e. is a *smooth* section, and moreover $\tau_0 \in \ker Z = \ker Z^* = \ker D \cap \ker \tilde{D}$. Since $i\tau_0 = \frac{1}{2}Z\tau_1$, applying D and $\tilde{D}$ to both sides we get $DZ\tau_1 = \tilde{D}Z\tau_1 = D\tau_0 = \tilde{D}\tau_0 = 0$ weakly. We conclude that $D^2\tau_1 = D(-i\tilde{D}\tau_1)$ and $\tilde{D}^2(-i\tau_1) = \tilde{D}D\tau_1$ weakly, which therefore equals 0 by the Lemma 19. But then

$$(D^2 \pm \tilde{D}^2)\tau_1 = P\tau_1 = i\tilde{D}D\tau_1 - iD\tilde{D}\tau_1 = 0$$

as well, and we consequently have $\tau_1 \in \ker P$, so τ_1 is smooth, and $\tau_1 \in \ker D^2 \cap \ker \tilde{D}^2$, so by Theorem 49,

$$\tau_1 \in \ker D^2 \cap \ker \tilde{D}^2 = \ker D \cap \ker \tilde{D} = \ker Z = \ker Z^*$$

We finally see, therefore, that

$$\tau_0 = -\frac{i}{2}Z\tau_1 = 0$$

Now, we also have $i\tau_1 = \frac{1}{2}Z\tau_2$, so applying D and $\tilde{D}$ to $i\tau_1 = \frac{1}{2}Z\tau_2$ gives $0 = iD\tau_1 = \frac{1}{2}DZ\tau_2 = i\tilde{D}\tau_1 = \frac{1}{2}\tilde{D}Z\tau_2$, which shows that $D^2\tau_2 = D(-i\tilde{D}\tau_2)$ and $\tilde{D}^2(-i\tau_2) = \tilde{D}D\tau_2$, and this equals 0 by

[2] We remark that this also follows from local coordinate considerations (5.35)-(5.38), taking into account that ∇ preserves each $\Gamma(\mathbf{Q}_\ell)$ and that symplectic Clifford multiplication by e_j and f_j are known to move Hermite-valued spinors h_α as $e_j \cdot h_\alpha = -\alpha_j i h_{\alpha-\varepsilon(j)} - \frac{i}{2}h_{\alpha+\varepsilon(j)}$ and $f_j \cdot h_\alpha = -\alpha_j i h_{\alpha-\varepsilon(j)} + \frac{i}{2}h_{\alpha+\varepsilon(j)}$, where $\varepsilon(j) \in \mathbb{N}_0^n$ has a 1 in the jth position and 0 elsewhere; see below.

Lemma 19. Therefore, $P\tau_2 = 0$, so τ_2 is smooth, and $\tau_2 \in \ker D^2 \cap \ker \tilde{D}^2$, so by Theorem 49 we again have $Z\tau_2 = 0$, showing that $i\tau_1 = \frac{1}{2}Z\tau_2 = 0$.

This indicates an inductive argument: The base case was given above, showing that $\tau_0 = 0$, while the inductive step follows as in the case of τ_1: suppose that for some $k \geq 0$ we have $\tau_0 = \tau_1 = \cdots = \tau_k = 0$, then we want to show that $\tau_{k+1} = 0$. Because all $\tau_j = 0$ for $j = 0, \ldots, k$, the expression $0 = i\tau_k = \frac{1}{2}Z^*\tau_{k-1} + \frac{1}{2}Z\tau_{k+1}$ shows that $Z\tau_{k+1} = 0$ weakly. By Corollary 23 we have that τ_{k+1} is smooth, and hence also in the kernels of D and $\tilde{D}$, while the relation

$$i\tau_{k+1} = \frac{1}{2}Z^*\tau_k + \frac{1}{2}Z\tau_{k+2} = \frac{1}{2}Z\tau_{k+2}$$

shows that, by applying D and $\tilde{D}$ to both sides, we have

$$\frac{1}{2}\tilde{D}Z\tau_{k+2} = i\tilde{D}\tau_{k+1} = 0 = iD\tau_{k+1} = \frac{1}{2}DZ\tau_{k+2}$$

weakly. But then we have $D^2\tau_{k+2} = D\tilde{D}(-i\tau_{k+2})$ and $\tilde{D}D\tau_{k+2} = \tilde{D}^2(-i\tau_{k+2})$ weakly. By Lemma 19 these must equal 0, weakly, and therefore

$$P\tau_{k+2} = i\tilde{D}D\tau_{k+2} - iD\tilde{D}\tau_{k+2} = 0$$

so $\tau_{k+2} \in \ker P$ is smooth by Corollary 23, while $\tau_{k+2} \in \ker D^2 \cap \ker \tilde{D}^2$ shows that

$$\tau_{k+2} \in \ker D \cap \ker \tilde{D} = \ker Z = \ker Z^*$$

We conclude that

$$0 = \frac{1}{2}Z\tau_{k+2} = i\tau_{k+1}$$

which completes the inductive argument. We conclude that $t = 0$, and therefore $\ker(D^* \pm iI) = \ker(\tilde{D}^* \pm iI) = \{0\}$, so the operators D and $\tilde{D}$ are essentially self-adjoint. ■

5.6 Local Form of the Symplectic Dirac Operators

Recall that $L^2(\mathbb{R}^n)$ decomposes into a countable direct sum $\bigoplus_{\ell=0}^{\infty} \mathfrak{W}_\ell$ of finite-dimensional eigenspaces $\mathfrak{W}_\ell$ of the Harmonic oscillator H_0, corresponding to the eigenvalues $\lambda = -(\ell + n/2)$

and associated Hermite eigenfunctions h_α. Since the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ is a dense subspace of $L^2(\mathbb{R}^n)$ it, too, decomposes as $\bigoplus_{\ell=0}^{\infty}\mathfrak{S}_\ell$. We now consider the symplectic connection ∇ acting on the smooth, compactly supported Schwartz space spinors, $\nabla:\Gamma_c(\mathbf{S})\to\Gamma_c(T^*M\otimes\mathbf{S})$. Locally, $\nabla=d+\omega$, where d is the exterior derivative and $\omega=d\mathfrak{m}(\tilde{\alpha})\in\Gamma_c(U,\mathrm{End}(\mathbf{S}))$ is the connection 1-form of ∇.[3] In view of the fact that the Hermite basis for $L^2(\mathbb{R}^n)$ gives rise to a local orthonormal frame $(h_\alpha)_{\alpha\in\mathbb{N}_0^n}\in\Gamma(P_{\mathrm{Mp}(n,\mathbb{R})})$ for the Schwartz spinors $\Gamma(\mathbf{S})$, which for convenience we enumerate now as $(h_j)_{j\in\mathbb{N}}$, we observe that ∇ acts on local sections $s\in\Gamma(\mathbf{S})$ as

$$\begin{aligned}\nabla s &= \nabla(s^i h_i)\\ &= ds^i\otimes h_i+s^i\nabla h_i\\ &= \frac{\partial s^i}{\partial x_j}dx_j\otimes h_i+s^i\omega_i^k\otimes h_k\end{aligned}$$

where we used the Einstein summation convention, with j running from 1 to $2n$ and i,k running over $\mathbb{N}_0$, and with $\omega_i^j\in\Omega^1(M)$.[4] That is, if we write $\omega=(\omega_i^j)_{i,j\in\mathbb{N}_0}$ as an infinite matrix (acting on $\mathbf{S}$) and $s=s^ih_i$ as a column vector $(s^0,s^1,\dots)^T$, and similarly with $\partial s/\partial x_j=(\partial s^0/\partial x_j,\partial s^1/\partial x_j,\dots)^T$, then the above expression can be written in matrix form as

$$\begin{aligned}\nabla s &= \sum_{j=1}^{2n}\sum_{i,k=0}^{\infty}\frac{\partial s^i}{\partial x_j}dx_j\otimes h_i+s^i\omega_i^k\otimes h_k\\ &= \sum_{j=1}^{2n}\begin{pmatrix} dx^j\otimes h_0 & 0 & \cdots\\ 0 & dx^j\otimes h_1 & \cdots\\ \vdots & \vdots & \ddots\end{pmatrix}\begin{pmatrix}\frac{\partial s^0}{\partial x_j}\\ \frac{\partial s^1}{\partial x_j}\\ \vdots\end{pmatrix}+\begin{pmatrix}\omega_0^0\otimes h_0 & \omega_0^1\otimes h_1 & \cdots\\ \omega_1^0\otimes h_0 & \omega_1^1\otimes h_1 & \cdots\\ \vdots & \vdots & \ddots\end{pmatrix}\begin{pmatrix}s^0\\ s^1\\ \vdots\end{pmatrix}\\ &= \sum_{j=1}^{2n}\mathrm{diag}(dx^j\otimes h_0,dx^j\otimes h_1,\dots)\frac{\partial s}{\partial x_j}+\omega s\\ &= \sum_{j=1}^{2n}A_j\frac{\partial s}{\partial x_j}+Bs\end{aligned}$$

where $A_j=\mathrm{diag}(dx^j\otimes h_0,dx^j\otimes h_1,\dots)$, $B=\omega\in\Omega^1(U,\mathbf{S})$.

[3] Our previous notation, in Section 5.2, for ω was $d\mathfrak{m}(\alpha)$, where $\alpha:=\varphi^*\tilde{\omega}\in\Omega^1(U;\mathfrak{mp}(n,\mathbb{R}))$ is the pullback of the $\mathfrak{mp}(n,\mathbb{R})$-valued 1-form $\tilde{\omega}:=d\rho^{-1}\circ F^*\omega\in\Omega^1(P_{\mathrm{Mp}(n,\mathbb{R})};\mathfrak{mp}(n,\mathbb{R}))$ via a local symplectic frame $\varphi=(e,f)\in\Gamma(U,P_{\mathrm{Sp}(n,\mathbb{R})})$, with ω in this setting the connection 1-form associated to the symplectic connection ∇ on M.

[4] We recall, from Section 5.2, the expression of ∇ in terms of a local symplectic frame $\varphi=(e,f)$, namely $(\nabla_X s)_p=ds_p(\tilde{X})-\frac{i}{2}\sum_{j=1}^{n}(f_j\cdot\nabla_{\tilde{X}}e_j-e_j\cdot\nabla_{\tilde{X}}f_j)\cdot s$. This expression is not the same as the one above, which is with respect to a local section $s\in\Gamma(U,\mathbf{S})$ expanded in the local $\mathbf{S}$-frame $(h_j)_{j\in\mathbb{N}_0}$.

Remark 66 Using (5.24) we can give a more explicit description of $B = \omega = d\mathfrak{m}(\alpha)$, namely $\omega(X)s = d\mathfrak{m}(\alpha(X))s = -\alpha(X)\cdot s = -\sigma(\alpha(X))s$, which is a linear combination of multiplication-by-y_j maps and partial derivative $\partial/\partial y_j$ operators at any point $x \in M$, with y_j the standard coordinates for $\mathbb{R}^n$. ■

If we now add in the actions of the sharpening operators $\omega^\sharp, g^\sharp : \Gamma(T^*M) \to \Gamma(TM)$ and of Clifford multiplication $\mu : \Gamma(TM\otimes\mathbf{S}) \to \Gamma(\mathbf{S})$, then we get local expressions for the Dirac operators $D = \mu\circ\omega^\sharp\circ\nabla$ and $\tilde{D} := \mu\circ g^\sharp\circ\nabla$, as maps from $\Gamma(\mathbf{S})$ to $\Gamma(\mathbf{S})$. For we recall the action of Clifford multiplication is $\mu(X\otimes s) := X\cdot s := \sigma(X)(s)$, where $\sigma : TM \subseteq \mathrm{Cl}(TM) \to \mathrm{End}(\mathbf{S})$ is the (induced) quantization map (Section **??**, equation (4.28)), and this shows that locally

$$\begin{aligned} Ds &= \mu\circ\omega^\sharp\circ\nabla s \\ &= \mu\circ\omega^\sharp\left(\frac{\partial s^i}{\partial x_j}dx_j\otimes h_i + s^i\omega_i^k\otimes h_k\right) \\ &= \frac{\partial s^i}{\partial x_j}x^j\cdot h_i + s^i\beta_i^k\cdot h_k \\ &= \sigma(X_{x^j})(h_i)\frac{\partial s^i}{\partial x_j} + \sigma(\beta_i^k)(h_k)s^i \\ &= \sum_{j=1}^{2n} A'_j\frac{\partial s}{\partial x_j} + B's \end{aligned} \tag{5.176}$$

where $A'_j = \mathrm{diag}(\sigma(X_{x^j})(h_0), \sigma(X_{x^j})(h_1), \dots)$ and $B' = \big(\sigma(\beta_i^k)(h_k)\big)_{i,k\in\mathbb{N}_0} \in \Gamma\big(U, \mathrm{End}(\mathbf{S})\big)$, with $\beta_i^k = \omega^\sharp(\omega_i^k) \in \Gamma(U, TM)$, and where X_{x^j} is the Hamiltonian vector field associated to x^j, the dual of the local basis vector x_j (see the remark below). Similarly,

$$\tilde{D}s = \sum_{j=1}^{2n} A''_j\frac{\partial s}{\partial x_j} + B''s \tag{5.177}$$

as above, except $\beta = g^\sharp(\omega_i^k)$. Thus, we have similar local expressions for D and $\tilde{D}$, as sums of the form

$$Ls = \sum_{j=1}^{2n} A_j(x)\frac{\partial s}{\partial x_j} + B(x)s \tag{5.178}$$

where $x \in U \subseteq M$ and $A_j, B \in \Gamma\big(U, \mathrm{End}(\mathbf{S})\big)$, $j = 1, \dots, 2n$.

We can be yet more specific about these local operators:

Proposition 75 *At a given point $x \in U \subseteq M$ the operators $A'_j(x)$, $A''_j(x)$, $B'(x)$ and $B''(x)$ in (5.176) and (5.177), as endomorphisms of $\mathcal{S}(\mathbb{R}^n) = \bigoplus_{\ell=0}^{\infty} \mathfrak{S}_\ell$, are given explicitly by*

$$A'_j(x)s = \begin{cases} -\dfrac{\partial s}{\partial x_j}, & j \leq n \\ ix_j s, & j > n \end{cases} \tag{5.179}$$

$$A''_j(x)s = \begin{cases} ix_j s, & j \leq n \\ \dfrac{\partial s}{\partial x_j}, & j > n \end{cases} \tag{5.180}$$

where the index j runs over the entire enumeration of the Hermite functions h_j. In the case of B' and B'', we let the index ℓ run over the enumeration of the eigenspaces $\mathfrak{S}_\ell$, where we recall $\dim \mathfrak{S}_\ell = \binom{n+\ell-1}{\ell}$ *and are therefore spanned by $\binom{n+\ell-1}{\ell}$ Hermite functions h_α with $|\alpha| = \ell$:*

$$B'(x)s = \begin{pmatrix} 0 & -M_0^* & 0 & 0 & 0 & \dots \\ M_0 & 0 & -M_1^* & 0 & 0 & \dots \\ 0 & M_1 & 0 & -M_2^* & 0 & \dots \\ 0 & 0 & M_2 & 0 & -M_3^* & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \\ \vdots \end{pmatrix} \tag{5.181}$$

$$B''(x)s = \begin{pmatrix} 0 & -N_0^* & 0 & 0 & 0 & \dots \\ N_0 & 0 & -N_1^* & 0 & 0 & \dots \\ 0 & N_1 & 0 & -N_2^* & 0 & \dots \\ 0 & 0 & N_2 & 0 & -N_3^* & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \\ \vdots \end{pmatrix} \tag{5.182}$$

Here,

$$M_\ell, N_\ell \in \mathcal{L}(\mathfrak{S}_\ell, \mathfrak{S}_{\ell+1}) \cong M_{\binom{n+\ell-1}{\ell}, \binom{n+\ell}{\ell+1}}(\mathbb{C})$$

and their elements, as we have seen, are of the form $\sigma(\beta_i^\alpha)(h_\alpha)$, which are linear combinations of the n multiplication operators ix_j and the n differential operators $\partial/\partial x_j$ applied to the $\binom{n+\ell-1}{\ell}$ Hermite functions h_α spanning $\mathfrak{S}_\ell$. As a result of Theorems 36 and 37 and the remark following, *we know that the operators ix_j and $\partial/\partial x_j$ are skew-symmetric on $\mathcal{S}(\mathbb{R}^n)$ and skew-adjoint on their respective domains, specified there. Therefore, we conclude that*

$$A'_j(x)^* = -A'_j(x)$$

$$A''_j(x)^* = -A''_j(x)$$

$$B'(x)^* = -B'(x)$$

$$B''(x)^* = -B''(x)$$

on their respective domains.

Proof: Since for any $f \in C^\infty(M)$ we have $df = \omega^\flat(X_f)$, where X_f is the Hamiltonian vector field of f (i.e. $df(Y) = \omega(X_f, Y) = \omega^\flat(X_f)(Y)$), and since $g(X,Y) = \omega(X, JY)$, and moreover in a local symplectic frame (e,f) we have $X_f = -Jdf = \frac{\partial f}{\partial e_j}\frac{\partial}{\partial f_j} - \frac{\partial f}{\partial f_j}\frac{\partial}{\partial e_j}$, we see that $dx^j = \omega^\flat(X_{x^j})$, where $x^j = x_j^*$ is the dual basis vector of one of e_k or f_ℓ, so

$$\omega^\sharp(dx^j) = \omega^\sharp\omega^\flat(X_{x^j}) = X_{x^j} = \begin{cases} -\dfrac{\partial}{\partial f_j}, & j \le n \\ \dfrac{\partial}{\partial e_j}, & j > n \end{cases}$$

$$g^\sharp(dx^j) = g^\sharp\omega^\flat(X_{x^j}) = g^\sharp g^\flat(JX_{x^j}) = JX_{x^j} = \begin{cases} \dfrac{\partial}{\partial e_j}, & j \le n \\ \dfrac{\partial}{\partial f_j}, & j > n \end{cases}$$

and therefore, identifying $\partial/\partial x_j$ with e_j if $j \le n$ and with f_j if $j \ge n$, we get that

$$\sigma(\omega^\sharp(dx^j))(h_i) = \begin{cases} -\dfrac{\partial h_i}{\partial x_j}, & j \le n \\ ix_j h_i & j > n \end{cases}$$

$$\sigma(g^\sharp(dx^j))(h_i) = \begin{cases} ix_j h_i, & j \le n \\ \dfrac{\partial h_i}{\partial x_j}, & j > n \end{cases}$$

Consequently, at a given $x \in U$,

$$A_j'(x) = \begin{cases} -\operatorname{diag}\Big(\dfrac{\partial h_0}{\partial x_j}, \dfrac{\partial h_1}{\partial x_j}, \dots\Big) = \dfrac{\partial}{\partial x_j}\operatorname{diag}\big(h_0, h_1, \dots\big), & j \le n \\ \operatorname{diag}\big(-ix_j h_0, -ix_j h_1, \dots\big) = ix_j \operatorname{diag}\big(h_0, h_1, \dots\big), & j > n \end{cases}$$

$$A_j''(x) = \begin{cases} \operatorname{diag}\big(ix_j h_0, ix_j h_1, \dots\big) = ix_j \operatorname{diag}\big(h_0, h_1, \dots\big), & j > n \\ \operatorname{diag}\Big(\dfrac{\partial h_0}{\partial x_j}, \dfrac{\partial h_1}{\partial x_j}, \dots\Big) = \dfrac{\partial}{\partial x_j}\operatorname{diag}\big(h_0, h_1, \dots\big), & j \le n \end{cases}$$

That is, $A_j'(x)$ and $A_j''(x)$ are either multiplication-by-x_j operators or a shift (involving partial differentiation with respect to x_j of the basis vectors h_i, since $\partial h_i/\partial h_j = h_k$ for some k) operators, i.e. differentiation of s, since $\partial s/\partial x_j = \sum_{i=0}^\infty \partial(s^i h_i)/\partial x_j = \sum_{i=0}^\infty s^i \partial h_i/\partial x_j = \partial/\partial x_j \operatorname{diag}(h_0, h_1, \dots)(s^0, s^1, \dots)^T$:

$$A_j'(x)s = \begin{cases} -\dfrac{\partial s}{\partial x_j}, & j \le n \\ ix_j s, & j > n \end{cases}$$

$$A_j''(x) = \begin{cases} ix_j s, & j \le n \\ \dfrac{\partial s}{\partial x_j}, & j > n \end{cases}$$

where, of course, $s = \sum_{i=0}^\infty s^i h_i$.

Let us now describe the operators $B'(x)$ and $B''(x)$. The observations (5.50)-(5.51) showed us that the operators Z and Z^* decrease and increase, respectively, the degree of the eigenbundles, $Z : \Gamma(\mathbf{Q}_\ell) \to \Gamma(\mathbf{Q}_{\ell-1})$ and $Z^* : \Gamma(\mathbf{Q}_\ell) \to \Gamma(\mathbf{Q}_{\ell+1})$, which means $D = \frac{1}{2}(Z + Z^*)$ and $\tilde{D} = \frac{1}{2i}(Z - Z^*)$ move eigensections $s_\ell \in \Gamma(\mathbf{Q}_\ell)$ to $\Gamma(\mathbf{Q}_{\ell+1})$ and $\Gamma(\mathbf{Q}_{\ell-1})$. The remainder of the claims follow from the skew-symmetry of ix_j and $\partial/\partial x_j$, as above. $\blacksquare$

Bibliography

[1] Adams, R. T., *Sobolev Spaces*, Academic Press (1975)

[2] Angella, D., *Cohomological Aspects in Complex Non-Kähler Geometry*, Springer (2014)

[3] Arnol'd, V. I., On a characteristic class entering into conditions of quantisation. English translation. *Functional Analysis and Its Applications* 1, 1-14, (1967)

[4] Arnol'd, V. I., Givental, A.B., *Symplectic geometry*, Translated from the 1985 Russian original, in *Dynamical Systems* IV, *Encycl. of Math. Sciences* 4, Springer, 1-136, (1990)

[5] Atiyah, M. F., Singer, I. M., The Index of Elliptic Operators on Compact Manifolds, *Bull. Amer. Math. Soc.* 69, 422433 (1963)

[6] Atiyah, M. F., Singer, I. M., The Index of Elliptic Operators. I, *Ann. of Math. (2)* 87, 484-530 (1968)

[7] Atiyah, M. F.; Segal, G. B., The Index of Elliptic Operators. II. *Ann. of Math. (2)* 87, 531-545. (1968)

[8] Atiyah, M. F.; Singer, I. M., The Index of Elliptic Operators. III. *Ann. of Math. (2)* 87, 546-604 (1968)

[9] Atiyah, M. F., Singer, I. M. The Index of Elliptic Operators. IV. *Ann. of Math. (2)* 93, 119-138. (1971)

[10] Atiyah, M. F.; Singer, I. M. The Index of Elliptic Operators. V. *Ann. of Math. (2)* 93, 139-149. (1971)

[11] Atiyah, M. F., Singer, I. M., Patodi, V. K., On the Heat Equation and the Index Theorem, *Invent. Math.* 19 (4), 279330 (1973)

[12] Audin, M., Symplectic and Almost Complex Manifolds, in Audin, M., Lafontaine, J. (Eds.), *Holomorphic Curves in Symplectic Geometry*, Birkhäuser (1994)

[13] Baum, P., van Erp, E., K-Homology and Fredholm Operators I: Dirac Operators, eprint arXiv:1604.03502 (2016)

[14] Baum, P., van Erp, E., K-Homology and Fredholm Operators II: Elliptic Operators, eprint arXiv:1604.03535 (2016)

[15] Berline, N., Getzler, E., Vergne, M., *Heat Kernels and Dirac Operators*, Springer (2004)

[16] Berndt, R., *An Introduction to Symplectic Geometry*, American Mathematical Society (2000)

[17] Bieliavsky, P., Cahen, M., Gutt, S., Rawnsley, J., Schwachöfer, L., Symplectic Connections, *Int. J. Geom. Methods Mod. Phys.* 3, no. 3, 375420. (2006)

[18] Borel, A., Hirzebruch, F., Characteristic Classes and Homogeneous Spaces, I, *American Journal of Mathematics*, Vol. 80, No. 2, pp. 458-538. (1958)

[19] Borel, A.; Serre, J.-P., Groupes de Lie et Puissances Réduites de Steenrod. *Amer. J. Math.* 75, 409-448. (1953)

[20] Brauer, R., Weyl, H., Spinors in n Dimensions, *American Journal of Mathematics* Vol. 57, No. 2, pp. 425-449. (1935)

[21] Bump, D., *Lie Groups*, Springer (2004)

[22] de Gosson, M., *Symplectic Geometry and Quantum Mechanics*, Birkhäuser (2006)

[23] Calabi, E., Construction and Properties of some 6-Dimensional Almost Complex Manifolds, *Trans. Amer. Math. Soc.* 87, 407438. (1958)

[24] Cartan, E., Sur certaines formes Riemanniennes Remarquables des Géométries à Groupe Fondamental Simple, *Ann. Sci. cole Norm. Sup.* (3) 44, 345467. (1927)

[25] Chernoff, R.P., Essential Self-Adjointness of Powers of Generators of Hyperbolic Equations, *J. Functional Analysis* **12**, 401-414. (1973)

[26] Cycon, H.L, Froese, R.G., Kirsch, W., Simon, B., *Schr odinger Operators with Applications to Quantum Mechanics and Global Geometry*, Springer (1987)

[27] Donaldson, S.K, Kronheimer, P., *The Geometry of Four-Manifolds*, Oxford (1990)

[28] Duistermaat, J.J., Kolk, J.A.C., *Lie Groups*, Springer (2000)

[29] Duistermaat, J.J., Kolk, J.A.C., *Distributions Theory and Applications*, Birkhäuser (2010)

[30] Duford, N., Schwartz, J., *Linear Operators*, Wiley (1963)

[31] Etesi, G., Complex Structure on the Six Dimensional Sphere from a Spontaneous Symmetry Breaking, *J. Math. Phys.* 56, no. 9, 099901, 1 p. (2015)

[32] Evans, L.C., *Partial Differential Equations*, 2nd Ed., American Mathematical Society (2010)

[33] Fefferman, C., Whitney's Extension Problem for C^m, Ann. of Math. (2) 164, no. 1, 313-359. (2006)

[34] Folland, G.B., *Harmonic Analysis in Phase Space*, Princeton University Press (1989)

[35] Folland, G.B., *Real Analysis: Modern Techniques and Their Applications*, 2nd Ed., Wiley (1999)

[36] Gel'fand, I. M., On Elliptic Equations, *Russ. Math.Surv.* 15 (3), 113123. (1960)

[37] Gel'fand, I., Retakh, V., Shubin, M., Fedosov Manifolds, *Adv. Math.* 136, no. 1, 104140. (1998)

[38] Getzler, E., A Short Proof of the Local AtiyahSinger Index Theorem, *Topology* 25: 111117. (1986)

[39] Gil-Medrano, O., Michor, P.W., The Riemannian Manifold of all Riemannian Metrics, *Quarterly J. Math.* Oxford (2) 42, 183–202. (1991)

[40] Gilkey, P., Curvature and the Eigenvalues of the Laplacian for Elliptic Complexes, *Advances in Math.* 10, 344-382. (1973)

[41] Gompf, R.E., A New Construction of Symplectic Manifolds, *Ann. of Math.*, 142, No. 3 , 527-595. (1995)

[42] Greub, W., *Multilinear Algebra*, 2nd Ed., Springer (1978)

[43] Greub, W., Petry H.R., On the Lifting of Structure Groups, in Bleuer, Petry, and Reetz, editors, *Differential Geometric Methods in Mathematical Physics II*, pp. 217-246, Springer (1978)

[44] Grove, L. C., *Classical Groups and Geometric Algebra*, American Mathematical Society (2002)

[45] Gromov, M., Pseudo-Holomorphic Curves in Symplectic Manifolds, Invent. Math. (1985)

[46] Gromov, M., Stable Mappings of Foliations into Manifolds, *Izv. Akad. Nauk SSSR Ser. Mat.*, 33, 707-734. (1969)

[47] Guillemin, V., Sternberg, S., *Symplectic Techniques in Physics*, Cambridge (1984)

[48] Guillemin, V., Sternberg, S., *Geometric Asymptotics*, American Mathematical Society (1991)

[49] Guan, D., Examples of Compact Holomorphic Symplectic Manifolds which are not K¨ahlerian. II. *Invent. Math.*, 121(1):135145. (1995)

[50] Habermann, K., The Dirac Operator on Symplectic Spinors. *Ann. Global Anal. Geom.* 13, 155-168 (1995)

[51] Habermann, K., Basic Properties of Symplectic Dirac Operators. *Commun. Math. Phys.* 184, 629-652 (1997)

[52] Habermann, K., Harmonic Symplectic Spinors on Riemann Surfaces. *Manuscr. Math.* 94, 465-484 (1997)

[53] Habermann, K., Symplectic Dirac Operators on Kähler Manifolds. *Math. Nachr.* 211, 37-62 (2000)

[54] Habermann, K., Habermann, L., *Introduction to Symplectic Dirac Operators* (Lecture Notes in Mathematics, Vol. 1887), Springer (2006)

[55] Habermann, K., Habermann, L., Rosenthal, P., Symplectic Yang-Mills theory, Ricci tensor, and connections, *Calc. Var. Partial Differential Equations* 30, no. 2, 137152. (2007)

[56] Hatcher, A., *Algebraic Topology*, Cambridge, 2001

[57] Hörmander, L, *The Analysis of Linear Partial Differential Operators I, Distribution Theory and Fourier Analysis*, 2nd Ed., Springer (2003)

[58] Husemoller, D. *Fibre Bundles*, 3rd Ed., Springer (1994)

[59] Huybrechts, Daniel *Complex Geometry: An Introduction*, Springer (2000)

[60] Iwasawa, K., On Some Types of Topological Groups, *Ann. of Math. (2)*, 50, 507558. (1949)

[61] Jost, J., *Riemannian Geometry and Geometric Analysis*, 5th Ed., Springer (2008)

[62] Kadison, R.V. and Ringrose, J.R., *Fundamentals of the Theory of Operator Algebras, Volume I: Elementary Theory*, American Mathematical Society (1997) (original printing Academic Press (1983))

[63] Kato, T., Fundamental Properties of Hamiltonian Operators of Schrdinger Type, *Trans. Amer. Math. Soc.* 70, 195211. (1951)

[64] Kirillov Jr., A., *An Introduction to Lie Groups and Lie Algebras*, Cambridge (2008)

[65] Kostant, B., Symplectic Spinors. *Symp. math.* 14 (Academic Press) XIV: 139152. (1974)

[66] Knapp, A. W., *Lie Groups: Beyond an Introduction*, 2nd Ed., Birkhäuser (2002)

[67] Knapp, A. W., *Advanced Real Analysis*, Birkhäuser (2005)

[68] Kolář, I., Michor, P.W., Slovák, J. *Natural Operations in Differential Geometry*, Springer (1993)

[69] Lang, S., *Fundamentals of Differential Geometry*, Springer (1991)

[70] Lang, S., *Algebra*, Springer (2002)

[71] Lawson, H. B. Jr., Michelsohn, M-L., *Spin Geometry*, Princenton University Press (1989)

[72] Lee, Jeffrey M., *Manifolds and Differential Geometry*, American Mathematical Society (2009)

[73] Lee, John M., *Introduction to Smooth Manifolds*, 2nd Ed., Springer (2013)

[74] Leoni, G., *A First Course in Sobolev Spaces*, American Mathematical Society (2009)

[75] Marathe, K. B., Martucci, G. *Mathematical Foundations of Gauge Theories*, Elsevier (1992)

[76] McDuff, D., Salamon, D., *Introduction to Symplectic Topology*, Oxford (1999)

[77] McDuff, D., Salamon, D., *J-holomorphic Curves and Symplectic Toplogy*, 2nd Ed., American Mathematical Society (2012)

[78] Meyers, N.G., Serrin, J., H=W, *Proc. Natl. Acad. Sci. USA* 51(6): 10551056. (1964)

[79] Milnor. J.W., Stasheff, J.D., *Characteristic Classes*, Annals of Mathematics Studies, Princeton (1974)

[80] McKean, H. P. Jr., Singer, I. M., Curvature and the Eigenvalues of the Laplacian, *J. Differential Geometry*, 1, 43-69. (1967)

[81] Melrose, R. B., *The AtiyahPatodiSinger Index Theorem*, CRC Press (1993)

[82] Moretti, V., *Spectral Theory and Quantum Mechanics, With an Introduction to the Algebraic Formulation*, Springer (2013)

[83] Mostow, G.D., A New Proof of E. Cartan's Theorem on the Topology of Semi-Simple Groups, *Bull. Amer. Math. Soc.* 55, 969980. (1949)

[84] Newlander, A., Nirenberg, L., Complex Analytic Coordinates in Almost Complex Manifolds. *Ann. Math.* 65, 391-404. (1957)

[85] Nicolaescu, Liviu I., *Lectures on the Geometry of Manifolds*, 2nd Ed., World Scientific (2007)

[86] Palais, R. S., *Seminar on the AtiyahSinger Index Theorem*, Annals of Mathematics Studies 57, Princeton (1965)

[87] Patodi, V. K., Curvature and the Eigenforms of the Laplace Operator. *J. Difl Grom.* 5, 233-249. (1971)

[88] Peetre, J., Une caractérisation abstraite des opérateurs différentiels, *Math. Scand.* 7, 211-218. (1959)

[89] Reed, M., Simon, B., *Methods of Modern Mathematical Physics I: Functional Analysis, Revised and Enlarged Edition*, Academic Press, 1980.

[90] Reed, M., Simon, B., *Methods of Modern Mathematical Physics II: Fourier Analysis, Self-Adjointness*, Academic Press, 1975.

[91] Roe, J., *Elliptic Operators, Topology and Asymptotic Methods*, 2nd Ed., Chapman and Hall/CRC (1998)

[92] Roman, S., *Advanced Linear Algebra*, 3rd Ed., Springer (2008)

[93] Rudin, W., *Real and Complex Analysis*, McGraw-Hill (1970)

[94] Schmüdgen, K., *Unbounded Self-adjoint Operators on Hilbert Space*, Springer (2012)

[95] Shubin, M., *Pseudodifferential Operators and Spectral Theory*, 2nd Ed., Springer (2001)

[96] Steenrod, N., *The Topology of Fiber Bundles*, Princeton University Press (1951)

[97] Szymiczek, Kazimierz, *Bilinear Algebra: An Introduction ot the Algebraic Theory of Quadratic Forms*, Gordon and Breach Science Publishers (1997)

[98] Taubes, C., The Seiberg-Witten invariants and symplectic forms, *Math. Res. Letters 1*, 809-822. (1994)

[99] Taubes, C., More constraints on symplectic forms from Seiberg-Witten equations, *Math. Res. Letters* 2, 9-14. (1995)

[100] Taylor, M., *Partial Differential Equations*, Volumes I-III, 2nd Ed., Springer (2010-2011)

[101] Thurston, W. P., Some simple examples of symplectic manifolds, *Proc. Amer. Math. Soc.* 55, no. 2, 467468. (1976)

[102] Treves, F., *Topological Vector Spaces, Distributions and Kernels*, Academic Press (1967)

[103] Vaisman, I., Symplectic Curvature Tensors, *Monatsh. Math.*, 100, no. 4, 299327. (1985)

[104] van den Ban, E., Crainic, M., *Analysis on Manifolds Lecture Notes for the 2012/2013 Mastermath Course*, http://www.staff.science.uu.nl/~crain101/AS-2013/chapters-1-2-3.pdf (2013)

[105] Varadarajan, V. S., *Lie Groups, Lie Algebras, and their Representations*, Springer, 1984

[106] Voisin, C, (Translated by L. Schneps), *Hodge Theory and Complex Algebraic Geometry* Vol. I, Cambridge (2002)

[107] von Neumann, J., Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren, *Math. Ann.* 102, no. 1, 49131. (1930)

[108] Webster, S. M., A new proof of the Newlander-Nirenberg theorem, *Mathematische Zeitschrift*, 201, no. 3, 303316. (1989)

[109] Whitney, H., Analytic extensions of differentiable functions defined in closed sets, *Trans. Amer. Math. Soc.* 36, no. , 63-89. (1934)

[110] Wolf, J.A., Essential Self Adjointness for the Dirac Operator and its Square. *Indiana Univ. Math. Journal*, Vol. 22, No. 7, 611640. (1973)

[111] Wu, W.-T.., Sur les classes caractéristiques des structures fibrées sphériques. *Publ. Inst. Math. Univ. Strasbourg* 11, pp. 5-89, 155-156 (1952)