# Review Notes for Qualifying Exam

Gabriel Beiner

*Email address*: gabriel.beiner@berkeley.edu

935 Evans Hall, University of California Berkeley

# Table of Contents


Chapter 1. Symplectic Geometry . . . . . 6
1. Symplectic Vector Spaces . . . . . 6
1.1. Basic Definitions . . . . . 6
1.2. The Symplectic Group . . . . . 7
1.3. Maslov Index . . . . . 8
1.4. Compatible Triples . . . . . 16
1.5. Symplectic Vector Bundles . . . . . 20
2. Symplectic Manifolds . . . . . 24
2.1. Basic Definitions . . . . . 24
2.2. Examples . . . . . 26
2.3. Moser's Trick . . . . . 29
2.4. Tubular Neighbourhood Theorems . . . . . 32
2.5. Almost-Complex Structures . . . . . 34
3. Classical Mechanics . . . . . 38
3.1. Hamilton's Equations . . . . . 38
3.2. Poisson Manifolds . . . . . 40
3.3. Integrable Systems and the Arnold–Liouville Theorem . . . . 44
4. Contact Manifolds . . . . . 52
4.1. Contact Structures . . . . . 52
4.2. Examples . . . . . 53
4.3. Structure Theorems . . . . . 57
4.4. Contact and Symplectic Geometry . . . . . 58
5. Symplectic and Hamiltonian Actions . . . . . 63
5.1. Circle Actions . . . . . 63
5.2. Hamiltonian Actions and Momentum Maps . . . . . 66
5.3. Examples of Momentum Maps . . . . . 69
5.4. Marsden–Weinstein–Meyer Theorem . . . . . 74
5.5. Torus Actions and Convexity . . . . . 79
5.6. Duistermaat–Heckman Localization . . . . . 85
6. Symplectic Constructions . . . . . 89
6.1. Complex Blowup . . . . . 89
6.2. Symplectic Blowups and Blowdowns . . . . . 91
6.3. Blowups and Momentum Maps . . . . . 93

6.4. Fibre Connected Sums . . . 95
6.5. Some More Results in Symplectic Topology . . . 98
6.6. Weinstein Domains . . . 100
7. Pseudoholomorphic Curves and Their Moduli Spaces . . . 104
7.1. Basic Properties of Pseudoholomorphic Curves . . . 105
7.2. Somewhere Injectivity . . . 107
7.3. Properties of Four-Dimensional Curves . . . 109
7.4. Moduli Spaces of Curves, Transversality, and Regularity . . . 112
7.5. Examples from the Sphere and Automatic Transversality . . . 123
7.6. Gromov Compactness and Bubbling . . . 125
7.7. Gromov Nonsqueezing and Results in Symplectic Topology . . . 135

Chapter 2. Algebraic Topology . . . 143
1. Basic Theory . . . 143
1.1. CW Complexes, Elementary Spaces, and Operations . . . 143
1.2. Homotopy . . . 147
1.3. Cofibrations and Fibrations . . . 148
1.4. Cellular Approximation . . . 151
2. The Fundamental Group and Coverings . . . 152
2.1. Fundamental Group . . . 152
2.2. Covering Spaces . . . 152
2.3. Seifert–Van Kampen Theorem and $\pi_1$ of CW Complexes . . . 155
3. Homology and Cohomology . . . 157
3.1. Singular Homology and Its Many Imitators . . . 157
3.2. Geometric Constructions: Reduced Homology, Homology of a Pair, Excision, and Mayer–Vietoris . . . 164
3.3. Algebraic Constructions: The Universal Coefficient Theorem, Künneth Formula, and Products on Cohomology . . . 172
3.4. Homology on Manifolds . . . 181
3.5. Cohomology of Some Common Spaces . . . 196
4. Homotopy Theory . . . 198
4.1. Homotopy Groups . . . 198
4.2. Long Exact Sequences in Homotopy . . . 201
4.3. The Freudenthal Suspension Theorem and Stable Homotopy Theory . . . 205
4.4. Homotopy and CW Complexes . . . 209
4.5. Obstruction Theory . . . 217
4.6. Eilenberg–MacLane Spaces, Messing Around with Homotopy Groups, and Applications to Cohomology . . . 225
5. Spectral Sequences . . . 242
5.1. Spectral Sequences in Homological Algebra . . . 243
5.2. Basic Examples of Spectral Sequences . . . 246

5.3. The Serre Spectral Sequence . . . 249
5.4. Computations with the Serre Spectral Sequence . . . 263
5.5. Generalized Cohomologies and the Atiyah–Hirzebruch Spectral Sequence . . . 280
6. Characteristic Classes . . . 298
6.1. Vector Bundles and Classifying Spaces . . . 298
6.2. Abstract Characteristic Classes and the Cohomology of Classifying Spaces . . . 311
6.3. Characteristic Classes and First Obstructions . . . 321
6.4. Properties of Characteristic Classes and Computations . . . 329
6.5. Thom Spectra and the Pontryagin–Thom Construction . . . 345

Chapter 3. Partial Differential Equations . . . 354
1. Functional Analysis . . . 354
1.1. Important Facts . . . 354
1.2. Riesz Representation and Lax–Milgram . . . 357
1.3. Compact Operators . . . 358
1.4. Fredholm Operators . . . 362
2. Sobolev Spaces . . . 362
2.1. Basic Properties . . . 363
2.2. Extension Operators . . . 364
2.3. Traces . . . 367
2.4. Sobolev Embeddings . . . 370
2.5. Compact Embeddings . . . 378
2.6. The Dual Space of $H_0^1$ . . . 382
3. Second Order Elliptic PDEs . . . 384
3.1. Basic Definitions . . . 384
3.2. Existence and Uniqueness . . . 386
3.3. Elliptic Regularity . . . 390
3.4. Maximum Principles . . . 398
4. The Calculus of Variations . . . 403
4.1. Variational Problems and the Euler–Lagrange Equations . . . 404
4.2. Existence . . . 409
4.3. Regularity . . . 416
4.4. Constraints and Lagrange Multipliers . . . 419

Bibliography . . . 426

CHAPTER 1

# Symplectic Geometry

## 1. Symplectic Vector Spaces

As inner product spaces are the local model for Riemannian geometry and complex vector spaces are the local model for complex geometry, so shall symplectic vector spaces describe the local structure of symplectic geometry.

**1.1. Basic Definitions.** We begin with the basic notions in symplectic linear algebra which will all have analogues in the manifold setting.

**Definition 1.1**

A *symplectic vector space* $(V, \omega)$ is a $2n$-dimensional vector space $V$ equipped with a *non-degenerate* skew–symmetric bilinear form $\omega$ called the *symplectic form*. I.e. $\iota_v\omega \in V^*$ is non-zero for each $v \neq 0$ or equivalently $\bigwedge^n \omega \neq 0$.

A *linear symplectomorphism* $\varphi : (V, \omega) \to (W, \omega')$ of symplectic vector spaces is a linear isomorphism so that $\varphi^*\omega' = \omega$. The group of symplectic isomorphisms of $V$ is denoted $Sp(V, \omega)$.

EXAMPLE 1.2. Consider $\mathbb{R}^{2n}$ with coordinates $x_1, \ldots, x_n, y_1, \ldots, y_n$. There is a *standard symplectic form*,

$$\omega_{\text{std}} := \sum_{i=1}^{n} \mathrm{d}x_i \wedge \mathrm{d}y_i \,.$$

Let $J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$, a block matrix on $\mathbb{R}^{2n}$; note $\omega(v, w) = w^\intercal J v$. Then $Sp(2n) := Sp(\mathbb{R}^{2n}, \omega_{\text{std}})$ may equivalently be regarded as the group of matrices $A \in GL(\mathbb{R}^{2n})$ so that $A^\intercal J A = J$. ▲

**Definition 1.3**

Given a symplectic vector space $(V, \omega)$ and subspace $W$, the *symplectic complement* of $W$ is,

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

We say a subspace $W$ is: *isotropic* if $W \subset W^\omega$, *coisotropic* if $W^\omega \subset W$, *symplectic* if $W \cap W^\omega = \{0\}$, and *Lagrangian* if $W = W^\omega$.

EXERCISE 1.4. Show $\dim W + \dim W^\omega = \dim V$ and $W^{\omega\omega} = W$. Show $W$ is Lagrangian if and only if it is co/isotropic with half the dimension of $V$. ▲

LEMMA 1.5. *Every symplectic vector space $(V, \omega)$ of dimension $2n$ has a* symplectic basis $x_1, \ldots, x_n, y_1, \ldots, y_n$ *so that,*

$$\omega(x_i, x_j) = \omega(y_i, y_j) = 0 \quad \textit{and} \quad \omega(x_i, y_j) = \delta_{i,j}.$$

*Equivalently, $V$ is symplectomorphic to $(\mathbb{R}^{2n}, \omega_{\text{std}})$.*

PROOF. Standard inductive exercise. Pick a pair of complementary vectors and apply induction hypothesis to the symplectic complement. □

COROLLARY 1.6. *Every isotropic subspace of $V$ extends to a Lagrangian subspace $L$. Any basis of $L$ extends to a symplectic basis of $V$.*

PROPOSITION 1.7 (Linear Symplectic Reduction). *Given a symplectic vector space $(V, \omega)$ with coisotropic subspace $W$, the quotient $W/W^\omega$ is a symplectic vector space with a unique symplectic structure pulling back to $\omega|_W$ under the projection.*

**1.2. The Symplectic Group.** We discuss some important properties

The Lie algebra of $Sp(V, \omega)$ can be described as follows.

$$\mathfrak{sp}(V, \omega) = \{A \in \text{End}(V) : \omega(A\cdot, \cdot) = -\omega(\cdot, A\cdot)\}.$$

In particular $\mathfrak{sp}(2n)$ are matrices $A$ so that $JA = -A^\intercal J$.

PROPOSITION 1.8 (2-out-of-3 Property). *One has equalities of Lie groups,*

$$U(n) = Sp(2n) \cap O(2n) = Sp(2n) \cap GL(n, \mathbb{C}) = O(2n) \cap GL(n, \mathbb{C}).$$

*Every symplectic matrix has determinant one.*

PROOF. Consider an invertible matrix $X$. We have $X \in GL(n,\mathbb{C})$ iff $XJ = JX$, $X \in Sp(2n)$ iff $X^\intercal JX = J$ and $X \in O(2n)$ iff $X^\intercal X = I$. It is easy enough to see any two conditions imply the third and in this case $X$ must be unitary.

If $X \in Sp(2n)$, we must have $X^*\omega_{\text{std}}^n = \omega_{\text{std}}^n$. But $\omega_{\text{std}}^n$ is the volume form on $\mathbb{R}^{2n}$ and so $X$ must be volume and orientation preserving so $\det X = 1$. □

LEMMA 1.9. *If $X \in Sp(2n)$ and $\lambda \in \text{Spec}(X)$, then $\lambda^{-1} \in \text{Spec}(X)$ with an equal multiplicity. The multiplicity of $\pm 1$ is even.*

PROOF. Note $X^\intercal = JX^{-1}J$. So, $X, X^\intercal, X^{-1}$ have the same eigenvalues. Since the eigenvalues of $X^{-1}$ are the inverse of the eigenvalues of $X$, the first statement follows. Moreover, the eigenvalues $\neq \pm 1$ counted with multiplicity can be paired off, so the combined eigenspace of $\pm 1$ is even dimensional. Since $\det X = 1$, the eigenvalue $-1$ (and hence 1) must have even multiplicity. □

PROPOSITION 1.10. *$U(n)$ is the maximal compact subgroup of $Sp(2n)$. In particular, $U(n)$ is a deformation retract of $Sp(2n)$.*

PROOF. The map, $F : I \times Sp(2n) \to Sp(2n)$ given by,

$$F_t(X) = X(X^\intercal X)^{-t/2},$$

is an explicit deformation retract. This is just polar decomposition for symplectic matrices. Recall the polar decomposition $X = PO$ with $O$ orthogonal and $P = \sqrt{X^\intercal X}$ positive definite. But $X^\intercal X$ is symplectic and so diagonalizing, $(X^\intercal X)^\lambda$ with be symplectic for any $\lambda \neq 0$. This implies $O$ is symplectic and since $O(2n) \cap Sp(2n) = U(n)$, it is unitary. $F_t(X)$ is then a symplectic path between $X$ and its unitary part $O$.

Given a larger compact subgroup of $Sp(2n)$, it must by polar decomposition contain some non-trivial positive definite matrix $P$. Since $\det P = 1$, such a thing must have an eigenvalue $\lambda > 1$. But then $\{P^k\}_{k\in\mathbb{N}}$ is a sequence with no convergent subsequence in our supposedly compact subgroup. □

**1.3. Maslov Index.** The Maslov index is a construction that assigns a homotopy invariant integer to paths of symplectic matrices and paths of Lagrangian subspaces. This construction is very important in symplectic Floer theories as it is crucial to defining a grading on various Floer homology theories.

**Theorem 1.11: Maslov Index for Symplectic Matrices**

There exists a unique function $\mu$, called the *Maslov index*, which assigns to each loop $\gamma : S^1 \to Sp(2n)$ of symplectic matrices an integer subject to the following axioms:

**Homotopy:** Loops are homotopic if and only if they have the same Maslov index. So, $\mu$ defines a bijective map from the free homotopy classes of loops in $Sp(2n)$ to $\mathbb{Z}$.

**Multiplicativity:** If we multiply two loops $\gamma_1 \cdot \gamma_2$ with respect to matrix multiplication, then the Maslov indices add:

$$\mu(\gamma_1 \cdot \gamma_2) = \mu(\gamma_1) + \mu(\gamma_2).$$

In particular, the identity loop has index zero.

**Direct Sum:** If we have loops $\gamma_1, \gamma_2$ in $Sp(2n)$ and $Sp(2m)$ respectively, then $\gamma_1 \oplus \gamma_2$ thought of as a loop in $Sp(2(m+n))$ in the natural way satisfies the index equality,

$$\mu(\gamma_1 \oplus \gamma_2) = \mu(\gamma_1) + \mu(\gamma_2).$$

**Normalization:** The loop $\gamma : S^1 \to Sp(2)$ given by,

$$\gamma(t) = \begin{pmatrix} \cos(2\pi t) & -\sin(2\pi t) \\ \sin(2\pi t) & \cos(2\pi t) \end{pmatrix},$$

has Maslov index one.

PROOF. Consider the map $f : Sp(2n) \to S^1$ given by,

$$f(X) = \det_{\mathbb{C}}(A+iB) \quad \text{where} \quad \begin{pmatrix} A & -B \\ B & A \end{pmatrix} = X(X^\intercal X)^{-1/2} \in U(n) \subset Sp(2n).$$

That is, we look at the unitary part of $X$ in its polar decomposition and take the complex determinant of that matrix.

Now given a path $\gamma(t)$ of symplectic matrices, its Maslov index is,

$$\mu(\gamma) = \deg(f \circ \gamma).$$

Really all this is, is we take our loop, use the deformation retraction above to get a loop in $U(n)$, and then look at its homotopy class in $\pi_1(U(n))$ under the canonical identification given by the determinant map with $\pi_1(S^1) = \mathbb{Z}$.

This is homotopy invariant. The multiplicativity and direct sum axioms are clear for unitary matrices from the definition and since every path is homotopic to a unitary one, the homotopy axiom implies this for all paths. The normalization is clear by computation.

Note the multiplicativity axiom implies the Maslov index is unchanged under a symplectic change of basis. By the spectral theorem, we can unitarily change basis so any unitary matrix is block diagonal with block consisting of $2 \times 2$ rotation matrices. Hence the normalization, homotopy, and direct sum axioms then determine the index of any path of unitary matrices. We can homotope a symplectic path to a unitary one, whence we prove uniqueness of the index. $\square$

Here is another interpretation of the Maslov index due to Arnold, Robbin, and Salamon. We define *Maslov cycle* $\Sigma \subset Sp(2n)$ to be the subset of matrices which have one as an eigenvalue. $\Sigma$ is the vanishing locus of $\det(I - X)$ for $X \in Sp(2n)$ and hence is an algebraic variety and a codimension one hypersurface; it has a natural co-orientation. The complement of $M$ is denoted $Sp^*(2n)$; it has two connected components determined by the sign of $\det(I - X)$.

Now given a loop $\gamma(t)$ of symplectic matrices, we can perturb it to intersect the Maslov cycle transversely away from the singular points of $\Sigma$. Then we can make an oriented count of the intersections of $\gamma$ with $\Sigma$; see Figure 1. This is homotopy invariant. Since $\Sigma$ splits $Sp(2n)$ into two connected components, the number of crossings with the loop $\gamma$ and the intersection count must be even. We define,

$$\mu(\gamma) = \frac{1}{2}\#(\gamma \cap \Sigma).$$

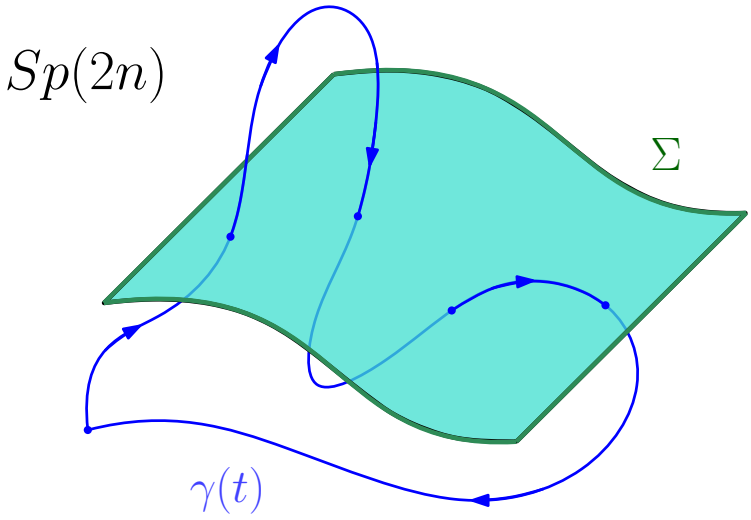



FIGURE 1. The Maslov index as an intersection number with the Maslov cycle $\Sigma$.

We will omit the verification that this satisfies the other axioms and so agrees with the Maslov index as previously constructed.

There is a very important twist on this construction called the *Conley–Zehnder index* for a particular collection of paths of symplectic matrices (it is also generalized to an index on all paths by Robbin–Salamon).

Let $\gamma(t) : [0,1] \to Sp(2n)$ be a path of symplectic matrices so that $\gamma(0) = I$ and so that $\gamma(1) \in Sp^*(2n)$. We define its Conley–Zehnder index $CZ(\gamma)$ as follows. Consider the matrices,

$$W_1 = -I \quad \text{and} \quad W_2 = \operatorname{diag}\left(-2, -1, \ldots, -1, \frac{1}{2}, -1, \ldots, -1\right).$$

These will live in the two distinct connected components of $Sp^*(2n)$. Whichever of the two connected components $\gamma(1)$ lives in, we can extend $\gamma$ to a path $\widetilde{\gamma} : [0,2] \to Sp(2n)$ so that $\widetilde{\gamma}$ remains within $Sp^*(2n)$ on $[1,2]$ and so that $\widetilde{\gamma}(2) = W_1$ or $W_2$. Note with $f : Sp(2n) \to S^1$ defined as within the preceding proof, we have $f(W_1), f(W_2) = \pm 1$. Thus, $f^2(W_1) = f^2(W_2) = 1$ using the degree two squaring map of $S^1$. Hence $f^2 \circ \widetilde{\gamma}$ defines a loop in $S^1$.

**Definition 1.12: Conley–Zehnder Index**

The *Conley–Zehnder index* of a path $\gamma : [0,1] \to Sp(2n)$ as above is defined as,

$$CZ(\gamma) := \deg(f^2 \circ \widetilde{\gamma}).$$

Since $\widetilde{\gamma}$ does not hit the Maslov cycle on $[1,2]$, this definition does not depend on the choice of extension of $\gamma$.

This also has a similar interpretation in terms of the Maslov cycle. We can extend $\gamma$ to $\widehat{\gamma} : [-\varepsilon, 1]$ so that on $[-\varepsilon, 0]$, one has $\widehat{\gamma}(t) = \operatorname{diag}(e^t, \ldots, e^t, e^{-t}, \ldots, e^{-t})$. $\widehat{\gamma}$ can then be perturbed to a path which intersects the Maslov cycle transversally away from its singular points. One then has,

$$CZ(\gamma) = \#(\widehat{\gamma} \cap \Sigma).$$

The Conley–Zehnder index satisfies similar properties to the Maslov index, which we summarize here.

**Theorem 1.13**

The Conley–Zehnder index has the following properties. Let $\gamma : [0,1] \to Sp(2n)$ be a path with $\gamma(0) = I$ and $\gamma(1) \in Sp^*(2n)$.
**Naturality:** For a path $\psi : [0,1] \to Sp(2n)$,

$$CZ(\gamma) = CZ(\psi\gamma\psi^{-1}).$$

**Homotopy:** If $\gamma$ and $\gamma'$ are homotopic via a family of paths satisfying our endpoint conditions, then $CZ(\gamma) = CZ(\gamma')$.

**Multiplicativity:** If $\psi$ is a loop in $Sp(2n)$ based at the identity,

$$CZ(\gamma \cdot \psi) = CZ(\gamma) + 2\mu(\psi).$$

**Sum:** If we have paths $\gamma_1, \gamma_2$ in $Sp(2n)$ and $Sp(2m)$ respectively satisfying our endpoint conditions, then $\gamma_1 \oplus \gamma_2$ thought of as a path in $Sp(2(m+n))$ in the natural way satisfies the index equality,

$$CZ(\gamma_1 \oplus \gamma_2) = CZ(\gamma_1) + CZ(\gamma_2).$$

**Parity:**

$$(-1)^{n-CZ(\gamma)} = \operatorname{sign} \det(I - \gamma(1)).$$

**Zero:** If the eigenvalues of $\gamma(t)$ do not lie on the unit circle for $t > 0$, $CZ(\gamma) = 0$.

**Rotation:** For $\theta \notin \mathbb{Z}$ the path of rotations,

$$\gamma(t) = \begin{pmatrix} \cos 2\pi\theta t & -\sin 2\pi\theta t \\ \sin 2\pi\theta t & \cos 2\pi\theta t \end{pmatrix}$$

has $CZ(\gamma) = 2\lfloor\theta\rfloor + 1$.

This can all be made nicely explicit in the case $n = 1$. Then $Sp(2)$ can be identified with a solid (open) torus $D^2 \times S^1$. A picture is provided in Figure 2, which we explain now. The central circle of the torus corresponds to the rotation matrices forming $U(1)$. The Maslov cycle $\Sigma$ consists of two cones with a common vertex at the identity. Mirror two $\Sigma$ are another two cones consisting of matrices with $-1$ as an eigenvalue which meet at minus the identity. Together these cones form two "banana" shapes which split $Sp(2)$ into four components. Inside the bananas are the *elliptic matrices*, i.e. those with complex eigenvalues on the unit circle. In the upper banana are those elliptic matrices $A$ which rotate positively ($\omega(v, Av) > 0$), while those in the lower banana rotate negatively. Outside the bananas are two components; the component touching the identity consists of the *positive hyperbolic* matrices with positive real eigenvalues, while the other consists of *negative hyperbolic* matrices with negative real eigenvalues.

We can use this picture to find the Conley–Zehnder index of a path in the two-dimensional case. Consider the red path $\gamma$ shown in Figure 3 (a). We can find $CZ(\gamma)$ using our two schema. In Figure 3 (b), we extend the path to $\widetilde{\gamma}$ which ends at $W_1$ without crossing the Maslov cycle in the extension. Now we can apply the retraction $f$, which will send the solid torus to its core circle. We see that $f \circ \widetilde{\gamma}$ winds one and a half times around the circle. Thus, $f^2 \circ \widetilde{\gamma}$ winds three times around the circle and so $CZ(\gamma) = 3$. In Figure 3 (b), we

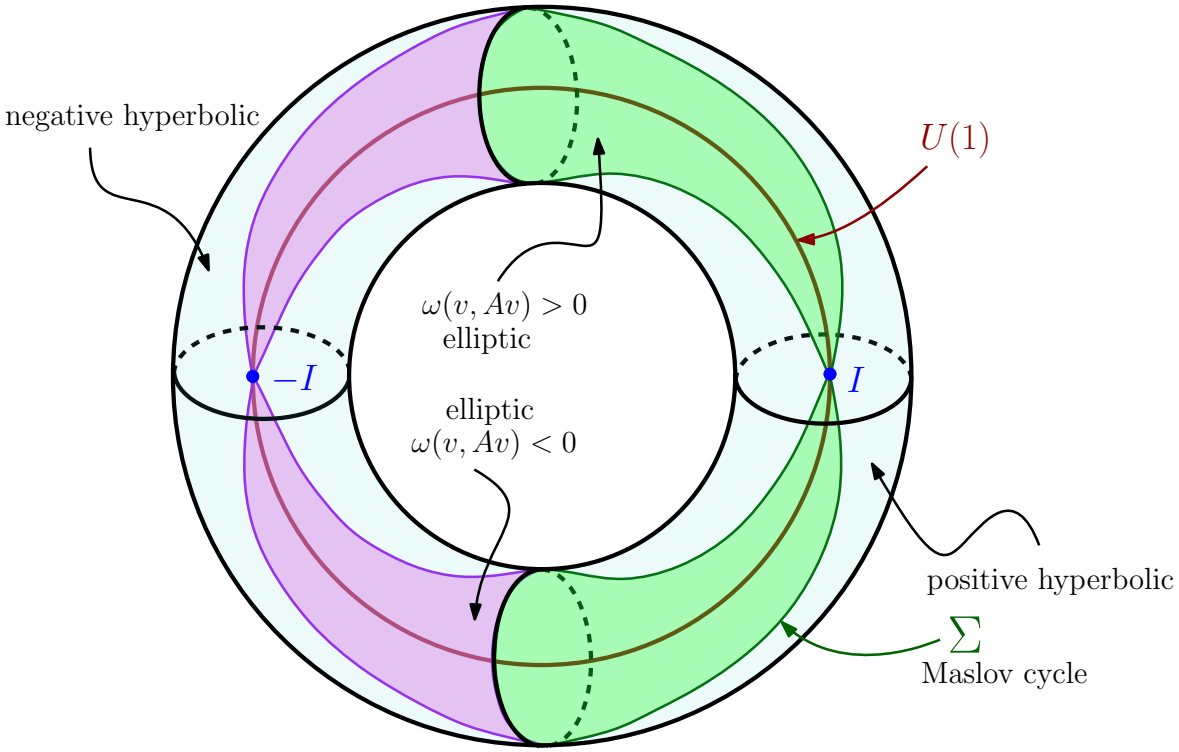



FIGURE 2. The space $Sp(2)$ as a solid torus.

perturb the start of the path to form $\widehat{\gamma}$. Now we count the intersections of $\widehat{\gamma}$ with the Maslov cycle $\Sigma$; we should orient $\Sigma$ as show with the green arrows, so the positive direction points into the upper cone and out of the lower one. We see $\widehat{\gamma}$ has three intersections with $\Sigma$, all with positive orientation, so again we find $CZ(\gamma) = 3$.

REMARK 1.14. For this remark, we assume the reader is familiar with some of the topics discussed in the succeeding section and chapters. In Floer theory, one is usually interested in the *Conley–Zehnder index of a Hamiltonian orbit* $\gamma(t) : I \to M$ (or sometimes a Reeb orbit, which is defined similarly). To define this, one first picks a trivialization $\tau$ of the symplectic vector bundle $\gamma^*TM$ (in favourable settings the resulting trivialization can be chosen canonically, or the resulting index depends nicely on our choice of $\tau$). Such a trivialization always exists since any complex, and thus symplectic, bundle on a circle is trivial. Under the trivialization $\tau$, the differential $\mathrm{d}_{\gamma(t)}\varphi_t$ of the Hamiltonian flow $\varphi_t$ along $\gamma$ determines a path of symplectic matrices $A(t)$. We have $\varphi_0 = \mathrm{id}$, and so if we assume $\det(I - \mathrm{d}\varphi_1) \neq 0$, then this path of matrices satisfies the conditions to define a Conley–Zehnder index. If this condition on $\mathrm{d}\varphi_1$ is met,

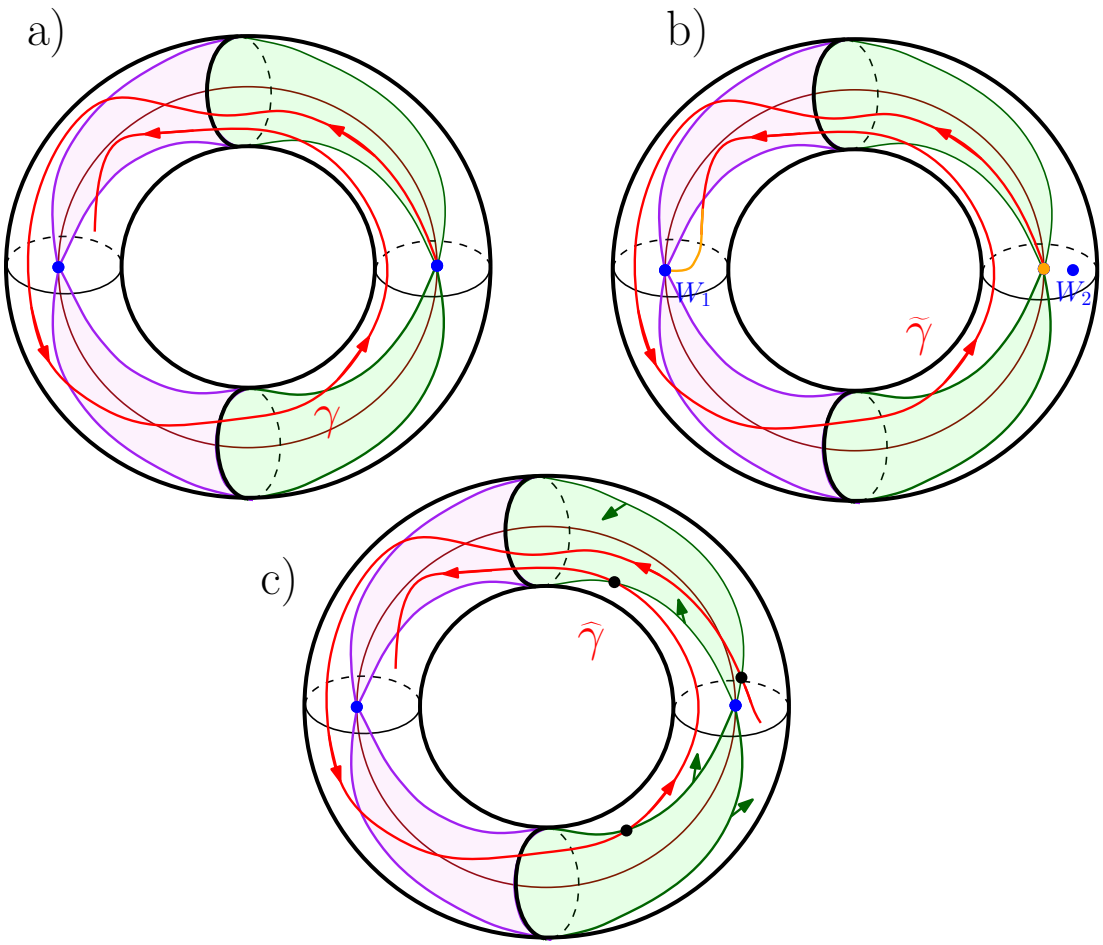



FIGURE 3. (a) A path $\gamma$ in $Sp(2)$ and (b), (c) two ways to compute its Conley–Zehnder index.

we say the orbit is *non-degenerate* and we define its Conley–Zehnder index to be,

$$CZ_\tau(\gamma) := CZ(A(t)).$$

This Conley–Zehnder index helps define the grading in many version of Floer homology. △

*Lagrangian Grassmannian.* There is an important version of this Maslov construction for paths of Lagrangian subspaces. To understand this, first we should say a little more about Lagrangian subspaces themselves. We will denote, for $(V, \omega)$ a symplectic vector space,

$$\Lambda(V, \omega) = \{\text{Lagrangian subspaces of } (V, \omega)\}.$$

We abbreviate $\Lambda(n) = \Lambda(\mathbb{R}^{2n}, \omega_{\text{std}})$. These spaces are called *Lagrangian Grassmannians*.

Any subspace of $\mathbb{R}^{2n}$ of dimension at most $n$ can be obtained as the image of a $2n \times n$ matrix $Z = \begin{pmatrix} X \\ Y \end{pmatrix}$; here $X, Y$ are real $n \times n$ matrices.

LEMMA 1.15. *Suppose we are subspace $L \subset \mathbb{R}^{2n}$ given as $L = \text{im}(Z)$ with $Z$ as above. Then $L \in \Lambda(n)$ if and only if $Z$ has rank $n$ and,*

$$X^\intercal Y = Y^\intercal X.$$

PROOF. Consider vectors $z = (Xu, Yu)$ and $w = (Xv, Yv)$ in $L$. Then,

$$\omega_{\text{std}}(z, w) = \omega(Xu, Yv) - \omega(Yu, Xv) = u^\intercal (X^\intercal Y - Y^\intercal X)v.$$

This gives the result. □

We will call such a matrix $Z$ satisfying the conditions of this lemma a *Lagrangian frame*. It is easy to show that the columns of $Z$ are orthonormal if and only if $X + iY$ is unitary. In this case, we say $Z$ is a *unitary Lagrangian frame*. Note every Lagrangian subspace has a unitary frame by the Gram–Schmidt procedure.

PROPOSITION 1.16. *$\Lambda(n)$ is diffeomorphic to the homogeneous space $U(n)/O(n)$.*

PROOF. To prove our proposition, note we can pick a unitary frame for every Lagrangian $L$. This is not unique, but it is unique up to right multiplication by $O(n)$, which corresponds to a change of orthonormal basis for $L$. Thus we get a smooth map $\Lambda(n) \to U(n)/O(n)$ compatible with left multiplication by $U(n)$ (i.e a map of homogeneous spaces). Injectivity is clear and surjectivity follows from the fact every unitary Lagrangian frame gives a Lagrangian. □

*Maslov Index for Lagrangian Subspaces.* The Lagrangian Grassmannian, being by above a homogeneous space, fits into a fibration. We get a homotopy long exact sequence for $n \geq 3$,

$$\cdots \to \pi_1(O(n)) = \mathbb{Z}_2 \xrightarrow{0} \mathbb{Z} = \pi_1(U(n)) \to \pi_1(\Lambda(n)) \to \pi_0(O(n)) = \mathbb{Z}_2 \to 0.$$

Hence the quotient map $U(n) \to \Lambda(n)$ is multiplication by two on $\pi_1$. In particular $\pi_1(\Lambda(n)) \cong \mathbb{Z}$. We now give a Maslov index to assign a loop in $\Lambda(n)$ an integer under this identification.

**Theorem 1.17: Maslov Index for Lagrangian Subspaces**

There exists a unique function $\mu$, called the *Maslov index*, which assigns to each loop $\gamma : S^1 \to \Lambda(n)$ of Lagrangian subspaces an integer subject to the following axioms:

**Homotopy:** Loops are homotopic if and only if they have the same Maslov index.

**Multiplicativity:** Given a loop $\gamma : S^1 \to \Lambda(n)$ and a loop of matrices $\rho : S^1 \to Sp(2n)$, we have,

$$\mu(\rho \cdot \gamma) = \mu(\gamma) + 2\mu(\rho).$$

**Identity:** Any constant loop has index zero.

**Normalization:** The loop $\gamma : S^1 \to \Lambda(1)$ given by,

$$\gamma(t) = e^{i\pi t}\mathbb{R} \subset \mathbb{C} = \mathbb{R}^2,$$

has Maslov index one.

PROOF. Given a Lagrangian subspace $L$ associated with a unitary Lagrangian frame $L = \operatorname{im}\begin{pmatrix} X \\ Y \end{pmatrix}$, let $U = X + iY$. Define $f : \Lambda(n) \to S^1$ by $f(L) = \det_{\mathbb{C}}(U^2)$. This is clearly independent of the choice of unitary Lagrangian frame. Note we need to include the square in this formula for this to work (this corresponds to the fact the map on $\pi_1$ was multiplication by two); we could remove the factor if our Lagrangians are unoriented.

Now we define the Maslov index of $\gamma(t)$ to be,

$$\mu(\gamma) = \deg(f \circ \gamma).$$

We will omit the routine verification of the axioms. This can be found in [McS1, pp. 51–52] □

There is again an intersection theory interpretation of this Maslov index as the intersection number with the *Maslov cycle* of Lagrangian subspaces intersecting $\{0\} \times \mathbb{R}^n$ non-transversally.

**1.4. Compatible Triples.** For many technical constructions in symplectic geometry (in particular all of Floer theory) it is important to make a choice of some extra geometric structure. The relevant choices are usually a metric and an almost-complex structure. We will want these to interact with the symplectic form in a compatible way, as we describe below.

**Definition 1.18: Linear Complex Structures**

A *complex structure* on a (even-dimensional) vector space $V$ is an automorphism $J \in \mathrm{Aut}(V)$ so that $J^2 = -I$. One can think of $J$ as endowing the vector space with a multiplication by the imaginary unit $i$. The set of complex structures on $V$ is denoted by $\mathcal{J}(V)$.

Suppose $(V, \omega)$ is a symplectic vector space. A complex structure $J$ is *$\omega$-compatible* if $g(v, w) = \omega(v, Jw)$ defines an inner product on $V$. I.e.,

$$\omega(v, Jv) > 0 \quad \text{for } v \neq 0 \text{ and } \quad \omega(Jv, Jw) = \omega(v, w).$$

If $g(v, w)$ is positive definite but not necessarily symmetric, we say $J$ is an *$\omega$-tame* complex structure. We may still obtain an inner product from an $\omega$-tame complex structure via,

$$\widetilde{g}(v, w) = \frac{1}{2}\big(\omega(v, Jw) - \omega(Jv, w)\big).$$

The set of $\omega$-compatible complex structures is denoted $\mathcal{J}(V, \omega)$ and the set of $\omega$-tame complex structures is denoted $\mathcal{J}_\tau(V, \omega)$.

Given a symplectic vector space $(V, \omega)$ with a compatible complex structure $J$ and an induced metric $g$, we call the tuple $(\omega, g, J)$ a *compatible triple* on $V$.

PROPOSITION 1.19. *The space $\mathcal{J}(\mathbb{R}^{2n})$ of complex structures is diffeomorphic to the homogeneous space $GL(\mathbb{R}^{2n})/GL(\mathbb{C}^n)$ and homotopy equivalent to $O(2n)/U(n)$. It has two connected components given by which orientation a complex structure induces.*

PROOF. First we note that given a complex structure $J$ on $\mathbb{R}^{2n}$, there is $A \in GL(\mathbb{R}^{2n})$ so that $JA = AJ_0$, where $J_0$ is the standard complex structure. To see this, observe we can always find a basis of $\mathbb{R}^{2n}$ of the form $v_1, Jv_1, \ldots, v_n, Jv_n$. If $v = (e_1, \ldots, e_n, f_1, \ldots, f_n)$ in the usual basis of $\mathbb{R}^{2n}$, then simply take,

$$Av = \sum_{i=1}^{n} (e_i v_i + f_i J v_i).$$

From this fact, we conclude that the map,

$$GL(\mathbb{R}^{2n}) \to \mathcal{J}(\mathbb{R}^{2n}) \quad \text{given by} \quad A \mapsto AJ_0A^{-1}$$

is surjective. It clearly descends to a bijection $GL(\mathbb{R}^{2n})/GL(\mathbb{C}^n) \to \mathcal{J}(\mathbb{R}^{2n})$ which will be a diffeomorphism. The remaining statements follow easily. □

We know given the compatible pair $(\omega, J)$ we obtain $g(v, w) = \omega(v, Jw)$. Not every metric arises in this way, but the only barrier is essentially a matter of rescaling, as the following lemma illustrates.

LEMMA 1.20. *Let $(V, \omega)$ be a symplectic vector space. There is a canonical continuous surjection,*

$$\mathrm{Riem}(V) \to \mathcal{J}(V, \omega),$$

*where* $\mathrm{Riem}(V)$ *denotes the set of inner products on $V$. This is a left inverse to the map sending a compatible complex structure to the metric it induces.*

PROOF. Let $g$ be an inner product on $V$. Let $T$ be the invertible linear map such that,

$$g(v, w) = \omega(v, Tw).$$

By skew-symmetry of $\omega$, $T^\intercal = -T$. Let $T = J|T|$ be the polar decomposition of $T$. Note $|T| = \sqrt{T^\intercal T} = \sqrt{-T^2}$. Since $T$ is normal, $J$ and $|T|$ commute, hence $T^2 = J^2(-T^2)$. Thus, $J^2 = -I$ defines an almost-complex structure. One has,

$$\omega(v, Jw) = \omega(v, J|J|^{-1}w) = g(v, |J|^{-1}w) = g(|J|^{-1/2}v, |J|^{-1/2}w).$$

Hence $\omega(v, Jw)$ is an inner product and so $J$ is a compatible complex structure. By construction this is left inverse to the map sending $J$ to $\omega(\cdot, J\cdot)$. In particular it is surjective. □

PROPOSITION 1.21. *The space of compatible complex structures $\mathcal{J}(V, \omega)$ is contractible.*

PROOF. The space $\mathrm{Riem}(V)$ is contractible, since it is a convex subspace of the vector space of symmetric bilinear forms $S^2V^*$. Let $r : I \times \mathrm{Riem}(V) \to \mathrm{Riem}(V)$ be a contraction. Then we can embed $\mathcal{J}(V, \omega)$ into $\mathrm{Riem}(V)$, apply $r$ and then apply the map from the above lemma to obtain a contraction of $\mathcal{J}(V, \omega)$.

Essentially we have just restated the proof that a retract of a contractible space is contractible. □

A number of other relevant spaces of are all naturally homotopy equivalent. Let $V$ be a vector space. Let $\Omega(V)$ denote the space of symplectic forms on $V$. For a fixed metric $g$ on $V$, let $\mathcal{J}(V, g)$ denote the set of complex structures on $V$ so that $J^*g = g$. Let $\Omega(V, g)$ denote all symplectic forms given as $g(J\cdot, \cdot)$ for some $J \in \mathcal{J}(V, g)$. Finally let $\mathcal{C}(V)$ denote the set of compatible triples $(\omega, J, \omega(\cdot, J\cdot))$.

PROPOSITION 1.22. *For any vector space $V$ of dimension $2n$ one has the following commutative diagram.*

$$\begin{array}{ccccccccc}
\frac{O(2n)}{U(n)} & \longrightarrow & \frac{GL(\mathbb{R}^{2n})}{GL(\mathbb{C}^n)} & \longleftarrow & \frac{GL(\mathbb{R}^{2n})}{U(n)} & \longrightarrow & \frac{GL(\mathbb{R}^{2n})}{Sp(2n)} & \longleftarrow & \frac{O(2n)}{U(n)} \\
\downarrow\cong & & \downarrow\cong & & \downarrow\cong & & \downarrow\cong & & \downarrow\cong \\
\mathcal{J}(V,g) & \longrightarrow & \mathcal{J}(V) & \longleftarrow & \mathcal{C}(V) & \longrightarrow & \Omega(V) & \longleftarrow & \Omega(V,g).
\end{array}$$

*The vertical maps are all diffeomorphisms and the horizontal maps are all homotopy equivalences. Here the outer two maps on the bottom are given by forgetting the metric and the outer two maps are given by forgetting the complex structure or symplectic form.*

PROOF. The vertical diffeomorphisms are clear by remembering what structures correspond to a symplectic form, metric, and complex structure. The horizontal maps are all given by inclusions of the top or bottom of the quotient as a maximal compact subgroup, and hence give homotopy equivalences. □

The contractibility of $\mathcal{J}(V,\omega)$ is important because it means that homotopy invariant constructions requiring a compatible almost-complex structure are independent of the choice. Sometime we will want to work with tame instead of compatible almost-complex structures, and so we want this to be contractible too.

PROPOSITION 1.23. *The space $\mathcal{J}_\tau(V,\omega)$ is contractible.*

PROOF. We have the space of compatible triples $\mathcal{C}(V)$ and similarly we can consider the space $\mathcal{C}_\tau(V)$ of symplectic forms $\omega$ and $\omega$-tame complex structures $J$. Define,

$$\rho_1 : \mathcal{C}_\tau(V) \to \Omega(V) \quad \text{given by} \quad \rho_1(\omega, J) = \frac{1}{2}(\omega + J^*\omega).$$

Note that this is well defined since,

$$\frac{1}{2}(\omega + J^*\omega)(\cdot, J\cdot) = \frac{1}{2}\big(\omega(\cdot, J\cdot) - \omega(J\cdot, \cdot)\big),$$

is non-degenerate by $\omega$-tameness of $J$ and thus is symplectic. $\rho_1(\omega, J)$ is compatible with $J$ and determines the same metric as $\omega$, which we call $g$.

Consider the fibration,

$$p : \mathcal{C}_\tau(V) \to \mathcal{C}(V) \quad \text{given by} \quad (\omega, J) \mapsto (\frac{1}{2}(\omega + J^*\omega), J).$$

Its fibre over $(\omega, J)$ is the space of pairs $(\omega + \sigma, J)$ where $\sigma \in \Lambda^2 V^*$ satisfies $\sigma + J^*\sigma = 0$. This conditions precisely means $\sigma$ is the real part of a $(2,0)$-form. Thus the fibre of this map is a vector space and $\mathcal{C}_\tau(V)$ is homotopy equivalent to $\mathcal{C}(V)$. Composing with the homotopy equivalent projection $\mathcal{C}(V) \to \Omega(V)$, $\rho_1$ is a homotopy equivalence. Let $\rho_0(\omega, J) = \omega$ be the projection.

Then $\rho_1$ and $\rho_2$ are homotopy equivalent via linear interpolation: $\rho_t = t\rho_1 + (1-t)\rho_0$, which one can check has image in symplectic forms. Hence $\rho_0$ is a homotopy equivalence.

Note $\rho_0$ is a fibration with fibre $\mathcal{J}_\tau(V, \omega)$ and so by the homotopy long exact sequence of a fibration, $\mathcal{J}_\tau(V, \omega)$ is weakly contractible. It is a manifold and hence contractible by Whitehead. □

Finally, let us clarify some notions around compatible triples. Recall, compatible $(\omega, J)$ are compatible if $g(\cdot, \cdot) = \omega(\cdot, J\cdot)$ defines a metric. Similarly, we call the pair $(g, J)$ compatible if $\omega(\cdot, \cdot) = g(J\cdot, \cdot)$ defines a symplectic form, and we call the pair $(\omega, g)$ compatible if $g(\cdot, \cdot) = \omega(\cdot, J\cdot)$ for a complex structure $J$.

In this way, any two elements of the compatible triple $(\omega, g, J)$ determine the third, and the elements of the triple will be pairwise compatible. The choice of compatible triple on $V$ is equivalent to making $V$ a *Hermitian vector space*. That is, a complex vector space with a Hermitian inner product $h$. The complex multiplication by $i$ on $V$ is equivalent to acting by the complex structure $J$ and the Hermitian metric $h$ is equivalent to $(g, \omega)$ via,

$$h(u, v) = g(u, v) + i\omega(u, v).$$

All of this is really just a consquence of the 2-out-of-3 property: Proposition 1.8. A symplectic form is a $Sp(n)$-structure, a metric is an $O(n)$-structure, a complex structure is a $GL(\mathbb{C}^n)$-structure, and the intersection of any two of these is a Hermitian metric, i.e. a $U(n)$-structure.

**1.5. Symplectic Vector Bundles.** We now consider the theory of symplectic vector bundles, i.e. bundles with a $Sp(n)$-structure. Because $U(n)$ is the maximal compact subgroup of $Sp(n)$, there is really no difference from the study of complex vector bundles.

**Definition 1.24: $G$-Bundles**

A *symplectic vector bundle* $(E, \omega)$ is a rank $2n$-vector bundle $E \to M$ over a manifold with a smoothly varying linear symplectic form $\omega$ on the fibres; that is, $\omega$ is any a non-degenerate two-form on $E$ which we call

an *almost-symplectic form*. This is the same as asking for a reduction of the structure group of $E$ from $GL(\mathbb{R}^{2n})$ to $Sp(2n)$.

An *isomorphism of symplectic vector bundles* is an isomorphism of real bundles so that the symplectic form on the codomain pulls back to the symplectic form on the domain. We call $M$ an *almost-symplectic manifold* if it is equipped with a symplectic vector bundle structure on its tangent bundle.

A *Riemannian vector bundle* $(E, g)$ is a vector bundle with a smoothly varying inner product $g$ called a *Riemannian metric*. This is the same as a vector bundle with a reduction to an $O(n)$-structure.

An *complex vector bundle* $(E, J)$ is a vector bundle with a smoothly varying linear complex structure $J$, called an *almost-complex structure*. This is the same as a vector bundle with a reduction to a $GL(\mathbb{C}^n)$-structure.

An *almost-Hermitian vector bundle* $(E, h)$ is a complex vector bundle with a smoothly varying Hermitian metric $h$. This is the same as a vector bundle with a reduction to a $U(n)$-structure.

Recall we showed above that $Sp(2n)$ deformation retracts to $U(n)$; this is also a deformation retract of $GL(\mathbb{C}^n)$. This means the classifying spaces $BU(n)$, $BGL(\mathbb{C}^n)$ and $BSp(n)$ are all homotopy equivalent. Hence isomorphism classes of symplectic and complex vectors agree.

**Theorem 1.25**

Every symplectic vector bundle $(E, \omega)$ has a unique complex structure. Two symplectic vector bundles are isomorphic if and only if the complex vector bundles they determine are.

Proof. As explained in the comment above, this is simple classifying space theory. □

**Definition 1.26**

By taking smoothly varying versions of the corresponding linear structure, we can define a compatible almost-complex structure on $(E, \omega)$ and a compatible triple $(\omega, g, J)$ of an almost-symplectic form, a Riemannian metric, and an almost-complex structure. As in the linear case, such a

compatible triple is the same as a Hermitian metric $h$ on $E$. Similarly we can define a tame almost-complex structure on $(E, \omega)$.

The space of compatible almost-complex structures on $(E, \omega)$ is denoted $\mathcal{J}(E, \omega)$. The space of tame almost-complex structures on $(E, \omega)$ is denoted $\mathcal{J}_\tau(E, \omega)$.

**Theorem 1.27**

Every symplectic vector bundle has a unique almost-Hermitian structure. In particular, every symplectic vector bundle admits a compatible almost-complex structure. Moreover, the space of compatible complex structures is contractible.

PROOF. Since an almost-Hermitian bundle is one equipped with a $U(n)$-structure, the same argument as above, shows isomorphism classes of symplectic and Hermitian bundles agree.

Since a symplectic bundle $(E, \omega)$ always admits an almost-Hermitian structure, we obtain from it a compatible triple $(\omega, g, J)$ which in particular yields a compatible almost-complex structure $J$.

Applying the same argument as in the linear case, $J \mapsto g = \omega(\cdot, J\cdot)$ will define a homotopy equivalence from the space of compatible almost-complex structures to Riemannian metrics. The space of metrics is a convex subset of a vector space and hence is contractible, thus so is the space of compatible almost-complex structures. □

Note that a choice of compatible almost-complex structure $J$ on $(E, \omega)$, endows $E$ with its unique up to isomorphism complex vector bundle structure $(E, J)$. The fact that the space of compatible almost-complex structures is contractible gives another reason that any choice of complex structure $(E, J)$ on $E$ is isomorphic to any other.

We can also more generally define a tame almost-complex structures on $(E, \omega)$ in analogy with the linear case.

**Theorem 1.28**

Every symplectic vector bundle $(E, \omega)$ admits a tame almost-complex structure. The space of tame almost-complex structures is contractible.

PROOF. Any compatible almost-complex structure is tame, and so the first statement follows from the previous theorem. For the second statement, choose a compatible almost-complex structure $J_0$ on $(E, \omega)$ inducing a metric $g$. With respect to this metric, a vector $v$ in a fibre of $E$ has a norm $\|v\|$. This induces an operator norm on endomorphisms $A : E_p \to E_p$ of a fibre of $E$:

$$\|A\|_{op} = \sup_{v \in E_p, \|v\|=1} \|Av\|.$$

We can set,

$$\mathcal{A} = \Big\{ A \in \Gamma(\mathrm{End}(E)) : AJ_0 = -J_0 A, \|A_p\| < 1 \, \forall p \in M \Big\}.$$

This is a convex set. We can define a homeomorphism $\mathcal{F} : \mathcal{J}_\tau(E, \omega) \to \mathcal{A}$ by,

$$\mathcal{F}(J) = (I + J_0 J)(I - J_0 J)^{-1} \quad \text{with inverse} \quad \mathcal{F}^{-1}(A) = J_0(I + A)^{-1}(I - A).$$

Since $\mathcal{A}$ is contractible, so is $\mathcal{J}_\tau(E, \omega)$. □

Vector bundles can be classified by their characteristic classes, which correspond to elements of the cohomology of their classifying space. So, symplectic vector bundles have characteristic classes given by $H^*(BSp(2n)) = H^*(BU(n))$. This means that the characteristic classes of symplectic vector bundles are just the Chern classes. To compute the Chern classes of a symplectic vector bundle $(E, \omega)$, one can pick any compatible almost-complex structure $J$ and compute $c_i(E, J)$ by any methods available for computing Chern classes of a complex bundle. The classes obtained in this way will be independent of the choice of $J$ as long as it is compatible with $\omega$.

Given a symplectic bundle $(E, \omega)$, we can ask about a *symplectic trivialization*, i.e. an isomorphism with a trivial symplectic bundle. Similarly, to a complex bundle $(E, J)$ or an almost-Hermitian bundle $(E, h)$ we can ask about a *complex or unitary trivialization*, which is an isomorphism with a trivial complex/Hermitian bundle. By our earlier result, a unitary and a symplectic trivialization are equivalent.

PROPOSITION 1.29. *Given a Hermitian bundle $(E, h) \to M$, and a smooth map $\gamma : D^n \to M$ for some $n$-dimensional disk $D^n$, $\gamma^* E$ has a unitary trivialization.*

PROOF. This follows from basic algebraic topology. $D^n$ is contractible so every unitary bundle on $D^n$ is trivial. □

## 2. Symplectic Manifolds

**2.1. Basic Definitions.** We begin with the foundational concepts of symplectic geometry which extend the ideas now familiar to us from symplectic linear algebra to the smooth category.

**Definition 2.1: Symplectic Manifold**

A *symplectic manifold* $(M, \omega)$ is a manifold $M$ equipped with a closed non-degenerate two form $\omega$. This is the same as saying $(TM, \omega)$ is a symplectic vector bundle with the additional integrability constraint $\mathrm{d}\omega = 0$.

A smooth map $\varphi : (M_1, \omega_1) \to (M_2, \omega_2)$ is called *symplectic* if $\varphi^*\omega_2 = \omega_1$. A symplectic map which is also a diffeomorphism is called a *symplectomorphism*.

A submanifold $W \subset (M, \omega)$ is called *Lagrangian* (resp. *isotropic*, *coisotropic*, *symplectic*) if for each $q \in W$, $T_qW$ is a Lagrangian (resp. isotropic, coisotropic, symplectic) subspace of $(T_qM, \omega_q)$.

Note there are several immediate topological restrictions to a manifold being symplectic. Clearly it must be even dimensional. Also, non-degeneracy is equivalent to the non-vanishing of the top form $\omega^n$ for $n$-dimensional $M$ . This implies that $(M, \omega)$ is orientable (and when compact is naturally an oriented manifold with volume form). Further, when $M$ is closed, $\omega^n$ must be non-trivial in de Rham cohomology (otherwise Stokes' formula implies $\mathrm{vol}(M) = 0$). This implies that $M$ has a non-trivial element $[\omega] \in H^2(X; \mathbb{R})$ so that $[\omega] \smile \cdots \smile [\omega] \in H^{2k}(X; \mathbb{R})$ is non-zero for $k \leq n$. In particular this rules out all the spheres except $S^2$ from being symplectic.

In the non-closed case, $\omega$ may be exact in which case we call $M$ an *exact symplectic manifold*.

Because $\omega$ is non-degenerate it defines, as with a metric, a musical bundle isomorphism,

$$\omega^\flat : TM \to T^*M \quad \text{given by} \quad \omega^\flat v = \iota_v \omega.$$

**Definition 2.2**

Given a function $H \in C^\infty(M)$, we define its *Hamiltonian vector field* $X_H \in \Gamma(TM)$ by the condition,

$$X_H = -(\omega^\flat)^{-1}\,\mathrm{d}H\,, \quad \text{i.e.} \quad \omega(\cdot, X_H) = \mathrm{d}H\,.$$

Note that many sources including [McS1] and [daS] take the opposite sign convention. More generally, a vector field $X$ is *symplectic* if $\omega^\flat X$ is a closed one form. Or equivalently, if $\mathcal{L}_X\omega = 0$. We will say a vector field $X$ is *Liouville* if $\mathcal{L}_X\omega = \omega$.

EXERCISE 2.3. These simple exercises are very important.

(1) Show a time-dependent complete vector field $X_t$ is symplectic for each time $t$ iff its flow $\phi_t$ is a symplectomorphism for each time $t$.

(2) Show if $Y, Z$ are symplectic then $[Y, Z]$ is the Hamiltonian vector field of $\omega(Y, Z)$. In particular, if $X_f, X_g$ are Hamiltonian, then $[X_f, X_g]$ is Hamiltonian for $\{f, g\}$, where
$$\{f, g\} := \omega(X_f, X_g),$$
is the *Poisson bracket* on $(M, \omega)$.

(3) Work out the Poisson bracket for $(\mathbb{R}^{2n}, \omega_{\text{std}})$ and show it is a Lie bracket. Deduce by Darboux's theorem (proved below) that the Poisson bracket on any symplectic manifold is a Lie bracket on $C^\infty(M)$. Deduce the map from smooth functions to Hamiltonian vector fields given by $f \mapsto X_f$ is a Lie algebra homomorphism.

(4) The flow of a Hamiltonian vector field $X_H$ is called the *Hamiltonian flow* of $H$. Show that the Hamiltonian vector field of $H$ is orthogonal to the gradient of $H$ with respect to any metric, so the Hamiltonian flow preserves the level sets of $H$. ▲

Consider a codimension one hypersurface $\Sigma \subset M$ in a symplectic manifold. It is necessarily coisotropic and $(T\Sigma)^\omega$ defines a sub-line bundle of $T\Sigma$. By the Frobenius theorem, this is necessarily integrable and so $\Sigma$ is foliated by curves; this is called the *characteristic foliation* and leaves of the foliation are called *characteristics*. If $\Sigma$ is a regular level set of a Hamiltonian $H$, then integral curves of $X_H$ form characteristics. One of the central areas of study in modern symplectic geometry is devoted to proving the existence of closed characteristics, i.e. periodic trajectories of some $X_H$, on such a hypersurface.

REMARK 2.4. We will later discuss further these concepts in the context of physics where Hamiltonian flows are intricately linked with the motion of particles in classical mechanics. △

A smooth one parameter family $\psi_t$ of symplectomorphisms of $M$ with $\psi_0 = \text{id}$ is called a *symplectic isotopy*. This is necessarily the flow of a time-dependent

symplectic vector field. If $\psi_t$ is in fact the flow of a time-dependent Hamiltonian vector field, then we say it is a *Hamiltonian isotopy*. The time one map of a Hamiltonian isotopy is called a *Hamiltonian diffeomorphism*. The group of such symplectomorphisms is denoted $\mathrm{Ham}(M, \omega)$.

We say a diffeomorphism $\varphi$ of $M$ is *non-degenerate* if for every fixed point $x$ of $\varphi$, $\mathrm{d}\varphi_x$ does not have one as an eigenvalue. A motivating question in the beginning of modern symplectic geometry was Arnold's conjecture on the fixed points of a non-degenerate Hamiltonian diffeomorphism.

CONJECTURE 2.5 (Arnold). *Let $(M^{2n}, \omega)$ be a closed symplectic manifold and $\varphi$ a non-degenerate Hamiltonian diffeomorphism. Then,*

$$\#\mathrm{Fix}(\varphi) \geq \sum_{i=0}^{2n} H_i(M; \mathbb{Z}).$$

Note if $\varphi$ is the Hamiltonian diffeomorphism associated to an *autonomous* (time-independent) Hamiltonian $H$, then there are at least as many fixed points of $\varphi$ as critical points of $H$. If the map $\varphi$ is non-degenerate then all critical points of $H$ are non-degenerate. Hence the conjecture follows from Morse homology.

The Arnold conjecture was proved by Conley–Zehnder for the $2n$ torus, making use of the fact Fourier series work on the torus. In his revolutionary work in the late 1980s, Floer proved the conjecture for all *monotone* symplectic manifolds:

$$\langle c_1(TM), h(\pi_2(M))\rangle = \lambda\langle \omega, h(\pi_2(M))\rangle \quad \text{for fixed} \quad \lambda > 0,$$

where $h$ is the Hurewicz homomorphism. Floer's proof hinged on the introduction of *Floer homology*, which is an infinite dimensional version of Morse homology. If we replace $\mathbb{Z}$ by $\mathbb{Q}$ in the statement of the Arnold conjecture, the conjecture is fully resolved due to a long series of works by Hofer–Salamon, Ono, Fukaya, Liu–Tian, and others. The conjecture is also resolved over a field of characteristic $p$ due to Abouzaid–Blumberg. As far as I am aware, the conjecture is still partially open over $\mathbb{Z}$. We hope to discuss Floer homology a little further towards the end of these notes.

**2.2. Examples.** We review some fundamental examples of symplectic manifolds.

(1) Symplectic vector spaces are symplectic manifolds. Up to symplectomorphism, there is just one such linear symplectic manifold in each dimension $(\mathbb{R}^{2n}, \omega_{\mathrm{std}})$.

(2) Given any manifold $M$, its cotangent bundle $T^*M$ is always an exact symplectic manifold. Locally, if $q_1, \ldots, q_n$ are coordinates on $M$ and $p_1, \ldots, p_n$ are the coordinates on $T^*M$ with $p_i$ corresponding to $\mathrm{d}x_i$, then we have,

$$\omega_{\mathrm{can}} = \sum_{i=1}^{n} \mathrm{d}p_i \wedge \mathrm{d}q_i \,.$$

This has a primitive,

$$\lambda_{\mathrm{can}} = \sum_{i=1}^{n} p_i \,\mathrm{d}q_i \,.$$

Sometimes $\lambda_{\mathrm{can}}$ is called the *tautological or canonical one-form* on $T^*M$. It admits the following coordinate-free description. Given $(v, w) \in T_qM \oplus T_q^*M = T_{(q,u)}T^*M$, we have,

$$\lambda_{\mathrm{can},(q,u)}(v, w) = u(v).$$

Equivalently, if $\pi : T^*M \to M$ is the projection then,

$$\lambda_{\mathrm{can},(q,u)} = \pi^* u.$$

Why is this tautological? One reason is for $\eta \in \Omega^1(M)$, $\eta^*\lambda_{\mathrm{can}} = \eta$ and this uniquely characterizes $\lambda_{\mathrm{can}}$. This also implies $\eta^*\omega_{\mathrm{can}} = \mathrm{d}\eta$.

Note that the case of $(\mathbb{R}^{2n}, \omega_{\mathrm{std}})$ is a special case of this (up to maybe a sign depending on how one orders the coordinates).

REMARK 2.6. We have taken the opposite sign convention from [**McS1**] and [**daS**] for $\omega_{\mathrm{std}}$. However we have also differed from them by a sign in defining the Hamiltonian vector field. Together these signs ensure that both our convention and theirs recovers the usual formulation of Hamilton's equations which we will discuss below. Essential reading in keeping track of signs in symplecitc geometry is Wendl's blog post, linked here. △

(3) Two dimensional symplectic manifolds are precisely orientable surfaces equipped with an area form. By Moser's trick below, two closed surfaces are symplectomorphic if and only if they have the same genus and area. The non-closed case is a little more complicated but solved due to Greene and Shiohama; see here. Symplectomorphisms of surfaces are just area preserving diffeomorphisms.

(4) A *Kähler manifold* $X$ is a complex manifold equipped with a symplectic form so that $\omega$ is compatible with the natural complex structure

on $TX$ (we call $\omega$ the *Kähler form* in this case). Examples of this include Riemann surfaces (as in the last example), K3 surfaces, smooth projective algebraic varieties, and *Stein manifolds*, which are analytic submanifolds of $\mathbb{C}^n$ with the induced complex structure and metric. Stein manifolds are necessarily non-compact by the maximum principle and include all non-compact Riemann surfaces.

(5) A special Kähler manifold is $\mathbb{C}P^n$. Consider on $\mathbb{C}^n$ the Kähler form,

$$\omega = \frac{i}{2}\partial\overline{\partial}\log\big(|z|^2+1\big).$$

Projective space $\mathbb{C}P^n$ is covered by open affine holomorphic charts,

$$\varphi_i : \mathbb{C}^n \to \mathbb{C}P^n, \quad (z_1,\dots,z_n) \mapsto [z_1 : \cdots : z_{i-1} : 1 : z_i : \cdots : z_n],$$

where we use homogeneous coordinates. It is a simple exercise to see that the forms $\varphi_i^*\omega$ agree on overlaps and hence patch together to define a Kähler form $\omega_{FS}$ on $\mathbb{C}P^n$ called the *Fubini-Study form*. Another way to obtain this symplectic structure is by the symplectic reduction (which we discuss later) of $(\mathbb{C}^{n+1}, \omega_{\text{std}})$ with respect to the $S^1$-action by complex multiplication and the moment map,

$$\mu(z) = \frac{1}{2}(1 - \|z\|^2).$$

▲

It is also good to see some examples of Lagrangian submanifolds. In $(\mathbb{R}^2, \omega_{\text{std}})$, any one dimensional submanifold is Lagrangian and in particular any embedded circle is a compact Lagrangian. Taking products of circles in each copy of $\mathbb{R}^2$, we obtain Lagrangian tori $T^n$ in $(\mathbb{R}^{2n}, \omega_{\text{std}})$. Also in $\mathbb{R}^{2n}$ with standard symplectic coordinates $x_1,\dots,x_n,y_1,\dots,y_n$, the span of the $x_i$'s and the span of the $y_i$'s are both non-compact Lagrangian.

In a cotangent bundle $(T^*M, \omega_{\text{can}})$, both the zero section and any fibre $T^*_xM$ are Lagrangian. More generally, the graph of any closed one form is Lagrangian.

Exercise 2.7. Prove the graph of a one form in $T^*M$ is Lagrangian if and only if the form is closed. ▲

Consider $(M,\omega)$ symplectic. We can form the product $M \times M$ with symplectic form $\omega \oplus (-\omega) = \pi_1^*\omega - \pi_2^*\omega$. Then the graph of any symplectomorphism $\psi : M \to M$ is a Lagrangian submanifold.

Exercise 2.8. Prove the graph of a smooth map $M \to M$ in $M^2$ is Lagrangian if and only if the map is a symplectomorphism. ▲

Given a complex projective variety, its real locus, i.e. the subset fixed under the involution of complex conjugation, is a Lagrangian submanifold, since this involution reverses the sign of the symplectic form but preserves tangent vectors to the real locus. In particular, $\mathbb{R}P^n$ is a Lagrangian submanifold of $\mathbb{C}P^n$. Another example of a Lagrangian submanifold of $\mathbb{C}P^n$ is the Clifford torus,

$$T = \{[z_0 : \cdots : z_n] \in \mathbb{C}P^n : |z_0| = |z_1| = \cdots = |z_n|\}.$$

Since $\omega_{FS}$ only depends on the norms of the coordinates, it pairs trivially with vectors tangent to $T$.

**2.3. Moser's Trick.** Our integrability condition in the symplectic form ensures that symplectic manifolds are all locally symplectomorphic to $(\mathbb{R}^{2n}, \omega_{\text{std}})$. This result, called Darboux's theorem, expresses the local flexibility of symplectic manifolds. The standard proof of Darboux's theorem and several related results uses *Moser's trick*, which constructs a desired diffeomorphism as the time one map of an isotopy built from a time-dependent vector field.

**Theorem 2.9: Moser's Trick**

Suppose $M$ is a compact manifold and $\omega_t$ is a one-parameter family of differential forms so that the equation,

$$\frac{\mathrm{d}}{\mathrm{d}t}\omega_t + \mathcal{L}_{X_t}\omega_t = 0,$$

admits a solution $X_t$. Then the flow $\psi_t$ of $X_t$ satisfies $\psi_t^*\omega_t = \omega_0$.

Proof. The proof of this is basically immediate from calculus. Because $\psi_t$ is the flow of $X_t$, we have $\frac{\mathrm{d}}{\mathrm{d}t}\psi_t = X_t \circ \psi_t$. Thus we have,

$$\frac{\mathrm{d}}{\mathrm{d}t}\psi_t^*\omega_t = \psi_t^*\left(\frac{\mathrm{d}}{\mathrm{d}t}\omega_t + \mathcal{L}_{X_t}\omega_t\right) = 0.$$

So $\psi_t^*\omega_t = \psi_0^*\omega_0 = \omega_0$. □

First we apply this to an interesting non-symplectic example.

Proposition 2.10. *Consider a compact manifold $M$ equipped with two volume forms $\alpha_1$ and $\alpha_2$. Then there exists a volume preserving diffeomorphism $\psi : (M, \alpha_1) \to (M, \alpha_2)$ if and only if the volumes $\int_M \alpha_1$ and $\int_M \alpha_2$ agree.*

Proof. The "only if" is clear by the change of variables theorem. Consider $\alpha_t = (1-t)\alpha_0 + t\alpha_1$. By Moser's trick, it suffices to show,

$$\alpha_0 - \alpha_1 = \mathcal{L}_{X_t}\alpha_t,$$

has a solution for $X_t$. By Cartan's formula,

$$\mathcal{L}_{X_t}\alpha_t = \iota_{X_t}\mathrm{d}\alpha_t + \mathrm{d}\iota_{X_t}\alpha_t = \mathrm{d}\iota_{X_t}\alpha_t,$$

where $\mathrm{d}\alpha_t = 0$, since $\alpha_t$ is a top-form. On the other hand, note since $\alpha_1, \alpha_2$ have the same volume, that $\alpha_1 - \alpha_2$ is trivial in cohomology and so equals $\mathrm{d}\beta$ for some form $\beta$. Thus we need a solution $X_t$ of,

$$\mathrm{d}(\beta - \iota_{X_t}\alpha_t) = 0.$$

But we can indeed always solve $\beta = \iota_{X_t}\alpha_t$ for $X_t$ since each $\alpha_t$ is a volume form and so fibrewise defines a non-degenerate pairing between codimension-one-forms and tangent vectors. □

PROPOSITION 2.11 (Moser Stability). *Let $M$ be a closed manifold and $\omega_t$ a one-parameter family of cohomologous symplectic forms. Then there is an isotopy $\psi : M \times [0,1] \to M$ so that $\psi_t^*\omega_t = \omega_0$ and $\psi_0 = \mathrm{id}$.*

PROOF. By Moser's trick, since $\omega_t$ is always closed, we need to find $X_t$ so that,

$$\frac{\mathrm{d}}{\mathrm{d}t}\omega_t + \mathrm{d}\iota_{X_t}\omega_t = 0.$$

Note,

$$\frac{\mathrm{d}}{\mathrm{d}t}\omega_t = \frac{\mathrm{d}}{\mathrm{d}t}(\omega_t - \omega_0).$$

Since $\omega_t$ and $\omega_0$ are cohomologous, this is null cohomologous. So, $\frac{\mathrm{d}}{\mathrm{d}t}\omega_t = \mathrm{d}\sigma_t$. Thus it suffices to solve,

$$\sigma_t + \iota_{X_t}\omega_t = 0$$

for $X_t$, which is always possible by non-degeneracy to obtain an isotopy $\psi$ from the flow of $X_t$. In order that $\psi$ is smooth, we need $\sigma_t$ to be smooth in $t$. Locally this is always possible on a coordinate chart and we can patch together by Mayer–Vietoris. □

**Theorem 2.12: Moser Isotopy Theorem**

Suppose $M$ is a smooth manifold and $Q$ a compact submanifold. Suppose $\omega_0, \omega_1$ are closed two forms which agree and are non-degenerate on $TM|_Q$. Then there are neighbourhoods $N_0, N_1$ of $Q$ and a diffeomorphism $\psi : N_0 \to N_1$ so that $\psi|_Q = \mathrm{id}$ and $\psi^*\omega_1 = \omega_0$.

PROOF. Suppose we may find a one form $\sigma$ on some $N_0$ so that $\mathrm{d}\sigma = \omega_1 - \omega_0$ and $\sigma|_Q = 0$. Then defining $\omega_t = \omega_0 + t\mathrm{d}\sigma$, we can shrink $N_0$ so $\omega_t$ is non-degenerate and obtain a vector field $X_t$ satisfying $\sigma + \iota_{X_t}\omega_t = 0$ and vanishing

on $Q$. By shrinking $N_0$ again, we can make this solution global on $t \in [0,1]$ and by Moser's trick obtain the result by taking the time one flow of $X_t$.

Thus it suffices to find $\sigma$. Up to diffeomorphism, we may suppose $N_0$ is neighbourhood of the zero section in the normal bundle $NQ$. Define an isotopy $\varphi_t : N_0 \to N_0$ by $\varphi_t(q,v) = (q,tv)$. Then $\varphi_0$ is the projection to the base and $\varphi_1 = \mathrm{id}$; this is a diffeomorphism for $t > 0$. So,

$$\varphi_0^*(\omega_1 - \omega_0) = 0 \quad \text{and} \quad \varphi_1^*(\omega_1 - \omega_0) = \omega_1 - \omega_0.$$

Define a vector field,

$$Y_t = \left(\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t\right) \circ \varphi_t^{-1} \quad \text{for} \quad t > 0.$$

Define,

$$\sigma_t = \varphi_t^*(\iota_{Y_t}(\omega_1 - \omega_0)) \quad \text{for} \quad t > 0.$$

Note this vanishes on $Q$ and extends smoothly to $\sigma_0 = 0$. We have,

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t^*(\omega_1 - \omega_0) = \varphi_t^*(\mathcal{L}_{Y_t}(\omega_1 - \omega_0)) = \mathrm{d}(\varphi_t^*\iota_{Y_t}(\omega_1 - \omega_0)) = \mathrm{d}\sigma_t.$$

Hence,

$$\omega_1 - \omega_0 = (\varphi_1^* - \varphi_0^*)(\omega_1 - \omega_0) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}(\varphi_t^*(\omega_1 - \omega_0))\,\mathrm{d}t = \mathrm{d}\sigma\,,$$

where $\sigma = \int_0^1 \sigma_t\,\mathrm{d}t$. □

Remark 2.13. Another way to state this argument is that we have obtained a chain homotopy, $K : \Omega^k(N_0) \to \Omega^{k-1}(N_0)$ defined by,

$$K(\alpha) = \int_0^1 \varphi_t^*(\iota_{Y_t}\alpha)\,\mathrm{d}t\,.$$

The above argument shows,

$$\mathrm{d} \circ K + K \circ \mathrm{d} = \mathrm{id} - \pi^* i^*,$$

where $\pi : N_0 \to Q$ is projection and $i : Q \to N_0$ is inclusion. Thus since $\pi^* i^*(\omega_1 - \omega_0)$ is exact, so is $\omega_1 - \omega_0$. △

**Theorem 2.14: Darboux's Theorem**

Every symplectic manifold $M$ is locally symplectomorphic to $\mathbb{R}^{2n}$ with its standard form.

PROOF. Consider the case of the relative Moser isotopy theorem when $Q$ is a single point $x \in M$. By symplectic linear algebra, we may find coordinates $x_1, \ldots, x_n, y_1, \ldots, y_n$ around $x$ so that $\omega_x$ is the standard form in these coordinates. Now $\omega_{\text{std}}$ and $\omega$ define two symplectic forms on an open ball around $x$ which agree at $x$. By the Moser isotopy theorem, there is a diffeomorphism $\psi$ of this neighbourhood taking $\omega$ to $\omega_{\text{std}}$. Hence in the coordinates $\psi^* x_1, \ldots, \psi^* x_n, \ldots \psi^* y_1, \ldots, \psi^* y_n$ on the neighbourhood of $x$, $\omega$ is the standard symplectic form. □

REMARK 2.15. Note this implies for example that all symplectic manifolds admit lots of (i.e. infinitely many) Lagrangian submanifolds. In any local neighbourhood, we get a symplectomorphism with $\mathbb{R}^{2n}$, which admits many Lagrangian tori. △

**2.4. Tubular Neighbourhood Theorems.** There are some more applications of Moser's trick to prove normal form results for tubular neighbourhoods of special submanifolds.

PROPOSITION 2.16 (Symplectic Neighbourhood Theorem). *Let $(M_1, \omega_1), (M_2, \omega_2)$ be symplectic manifolds with compact symplectic submanifolds $Q_1, Q_2$. Suppose there is an isomorphism of symplectic normal bundles $\Phi : \nu_{Q_0} \to \nu_{Q_1}$ restricting on the zero sections to a symplectomorphism $\varphi : (Q_0, \omega_0) \to (Q_1, \omega_1)$. Then there are neighbourhoods $N_0, N_1$ of $Q_0, Q_1$ so that $\varphi$ extends to a symplectomorphism $\psi : (N_0, \omega_0) \to (N_1, \omega_1)$ with $\mathrm{d}\psi$ inducing $\Phi$ on $\nu_{Q_0} = (TQ_0)^\omega$.*

PROOF. Let $\exp_j$ be the exponential map with respect to some metric taking a neighbourhood of the zero section in $\nu_{Q_j}$ to a tubular neighbourhood of $Q_j$. Then $\varphi$ extends to a diffeomorphism,

$$\varphi' = \exp_1 \circ \Phi \circ \exp_0^{-1} : N_0 \to N_1,$$

so that $\mathrm{d}\varphi'$ induces $\Phi$ on $\nu_{Q_0} = (TQ_0)^\omega$. Thus we reduce to the case where $M_1 = M_2, Q_1 = Q_2$ and $\omega_1|_Q = \omega_2|_Q$. Now the result follows from relative Moser isotopy. □

PROPOSITION 2.17 (Weinstein Tubular Neighbourhood Theorem). *Let $(M, \omega)$ be symplectic manifold and $L \subset M$ a compact Lagrangian. Then there is a tubular neighbourhood $N$ of the zero section in $T^*L$, a tubular neighbourhood $V$ of $L$ in $M$, and a symplectomorphism $\psi : (N, \omega_{\text{can}}) \to (V, \omega)$ so that $\psi|_L = \mathrm{id}$.*

PROOF. Pick a metric $g$ on $M$ and a compatible almost complex structure $J$ (see the end of the section). The action of $J$ then defines an isomorphism $TL \to TL^\perp \cong NL$. And $(T_qL)^\perp$ will be a Lagrangian subspace of $T_qM$ for

any $q \in L$. Consider the musical isomorphism, $g^\sharp : T^*L \to TL$. We can then define,

$$\psi : T^*L \to M \quad \text{by} \quad \psi(q,v) = \exp_q(Jg^\sharp(v)).$$

That is, we take a covector, dualize to get a vector, rotate by $J$ to get a normal vector, and then exponentiate to get an element of the tubular neighbourhood.

Given $v = (v_1, v_2) \in T_qL \oplus T_q^*L \cong T_{(q,0)}T^*L$, we have,

$$\mathrm{d}\psi_{(q,0)}(v) = v_1 + Jg^\sharp(v_2).$$

So, since $L$ is Lagrangian,

$$\begin{aligned}
\psi^*\omega_{(q,0)}(v,w) &= \omega_q(v_1 + Jg^\sharp v_2, w_1 + Jg^\sharp w_2) \\
&= \omega_q(v_1, Jg^\sharp w_2) - \omega_q(w_1, Jg^\sharp v_2) \\
&= g(v_1, g^\sharp w_1) - g(w_1, g^\sharp v_2) \\
&= w_1(v_1) - v_2(w_1) \\
&= \mathrm{d}\lambda_{can}(v,w),
\end{aligned}$$

where $\lambda$ is the tautological one form on the cotangent bundle. Now by the Moser isotopy theorem we get the desired map. □

Note this result puts lots of restrictions on what Lagrangians $L$ we can have by the fact that $TL$ needs to be isomorphic to the normal bundle $NL$ of the inclusion. For example, consider a Lagrangian surface $L$ in $\mathbb{R}^4$. Note that $L$ is always displaceable in $\mathbb{R}^4$ and hence the self intersection number $\chi(NL)$ is zero in the oriented case and is zero modulo two in the non-oriented case. Since $NL \cong TL$, we conclude if $L$ is orientable then $\chi(L) = 0$ and if $L$ is non-orientable $\chi(L)$ is even. Hence the only orientable Lagrangians are tori and the only non-orientable Lagrangians are connect sums of Klein bottles. In fact, a Klein bottle $K$ cannot be a Lagrangian in $\mathbb{R}^4$ but all higher connect sums $K^{\#n}$ for $n > 1$ do appear, as shown by Givental. In general $\mathbb{R}^{2n}$ can never have an *exact Lagrangian*. This is a Lagrangian submanifold $L$ of an exact symplectic manifold $(M, \mathrm{d}\lambda)$ so that $\lambda|_L$ is exact. A proof of this fact follows from Viterbo's work on symplectic homology.

An important open problem related to exact Lagrangians is the following.

Conjecture 2.18 (Arnold Nearby Lagrangian Conjecture). *Let $M$ be a closed manifold. Any closed exact Lagrangian in $T^*M$ is Hamiltonian isotopic to the zero section.*

This conjecture is basically wide open. The best progress is due to Abouzaid–Kragh who show that if $L \hookrightarrow T^*M$ is an exact closed Lagrangian embedding, then the projection $L \to M$ is a simple homotopy equivalence.

**2.5. Almost-Complex Structures.** Note that the tangent bundle $TM$ of a symplectic manifold $(M, \omega)$ is in particular a symplectic vector bundle. We can apply our theory of symplectic vector bundles to consider compatible and tame almost-complex structures on $TM$ (or usually we just say on $M$) and compatible triples.

We write $\mathcal{J}(M, \omega)$ and $\mathcal{J}_\tau(M, \omega)$ respectively for the space of compatible and tame almost-complex structures on $(M, \omega)$. By our earlier result, these spaces are non-empty and contractible.

Suppose we choose a compatible triple $(\omega, g, J)$ on a symplectic manifold, determining a Hermitian metric $h$. This gives $M$ the structure of an almost-Hermitian manifold meaning $TM$ is an almost-Hermitian bundle. Because $\omega$ additionally satisfies the integrability constraint $\mathrm{d}\omega = 0$, one sometimes calls this an *almost-Kähler manifold*. Hence any symplectic manifold can be made almost-Kähler, and this structure is unique up to isomorphism.

As with symplectic forms, there are some topological constraints on what manifolds admit almost-complex structures. Clearly they must be even dimensional. There are also constraints on the manifold's characteristic classes which for example implies $S^2$ and $S^6$ are the only spheres which can carry almost-complex structures. There is a natural almost-complex structure on $S^6$ coming from octonion multiplication on $\mathbb{R}^7 = \mathrm{Im}(\mathbb{O})$.

There is an integrability condition one can impose on an almost-complex manifold $(M, J)$, namely that $N_J = 0$. Here $N_J$ is the *Nijenhuis tensor*, a vector-valued two-form given by,

$$N_J(X, Y) = [X, Y] + J[JX, Y] + J[X, JY] - [JX, JY].$$

This is equivalent to asking that the space of holomorphic vector fields $T^{1,0}_{\mathbb{C}} M$ inside the complexified tangent bundle of $M$ is involutive.

We will say that an almost-complex structure $J$ on $M$ is *integrable* when $M$ can be made into a complex manifold, i.e. one with biholomorphic transition functions, with respect to the complex structure of $J$. A deep theorem asserts that the condition above suffices to give integrability.

**Theorem 2.19: Newlander–Nirenberg**

An almost complex manifold $(M, J)$ is integrable if and only if $N_J = 0$.

**Definition 2.20**

Suppose $(M, \omega)$ is a symplectic manifold equipped with a compatible triple $(\omega, g, J)$. If $J$ is integrable, we say $(M, J, \omega)$ is a *Kähler manifold*. We call $\omega$ the *Kähler form*.

Kähler manifolds are very special and well studied in complex and algebraic variety. Every smooth projective variety is Kähler, as is every Riemann surface. An important property of compact Kähler manifolds is that they satisfy the *Hodge decomposition theorem*:

$$H^k(M; \mathbb{C}) = \bigoplus_{p+q=k} H^{p,q}(M),$$

where $H^{p,q}$ are the *Dolbeault cohomology groups*, defined from the complex obtained from the anti-holomorphic differential on complex $(p, q)$-forms. Complex conjugation proves that $H^{p,q}$ and $H^{q,p}$ are isomorphic for Kähler manifolds and so one purely topological consequence of the above fact is that the odd Betti numbers of a compact Kähler manifold must be even. There is much more to say about Kähler geometry, some of which is summarized in [**McS1**, §4.4] and much more can be found in [**Huy**].

EXAMPLE 2.21. Not every symplectic manifold is Kähler. This is not at all an obvious fact. The first discovered example is called the *Kodaira-Thurston manifold*. One considers $\mathbb{R}^4$ with its standard symplectic form. Consider $\Gamma = \mathbb{Z}^2 \ltimes \mathbb{Z}^2$ equipped with the operation,

$$(j', k') \circ (j, k) = (j + j', A_{j'}k + k') \quad \text{for} \quad A_{j'} = \begin{pmatrix} 1 & j+1' \\ 0 & 1 \end{pmatrix}.$$

This discrete group acts by symplectomorphisms on $\mathbb{R}^4$ via,

$$(j, k) \cdot (x, y) = (x + j, A_j y + k).$$

The quotient $M = \mathbb{R}^4/\Gamma$ is a symplectic manifold and also complex. $M$ is a torus bundle over the torus. Its first homology is,

$$H_1(M) = \Gamma/[\Gamma, \Gamma] = \mathbb{Z}^3.$$

This is odd dimensional and hence can not be Kähler, since the odd Betti numbers should be even. ▲

In general, a complex manifold need not be symplectic, a symplectic manifold need not be complex and an almost complex manifold need not be complex or symplectic. See [**daS**, §17.3]. A chart illustrating the various examples is shown in figure 4. The example by Fernandez–Gotay–Gray of symplectic non-complex manifolds comes from circle bundles over circle bundles over $T^2$.

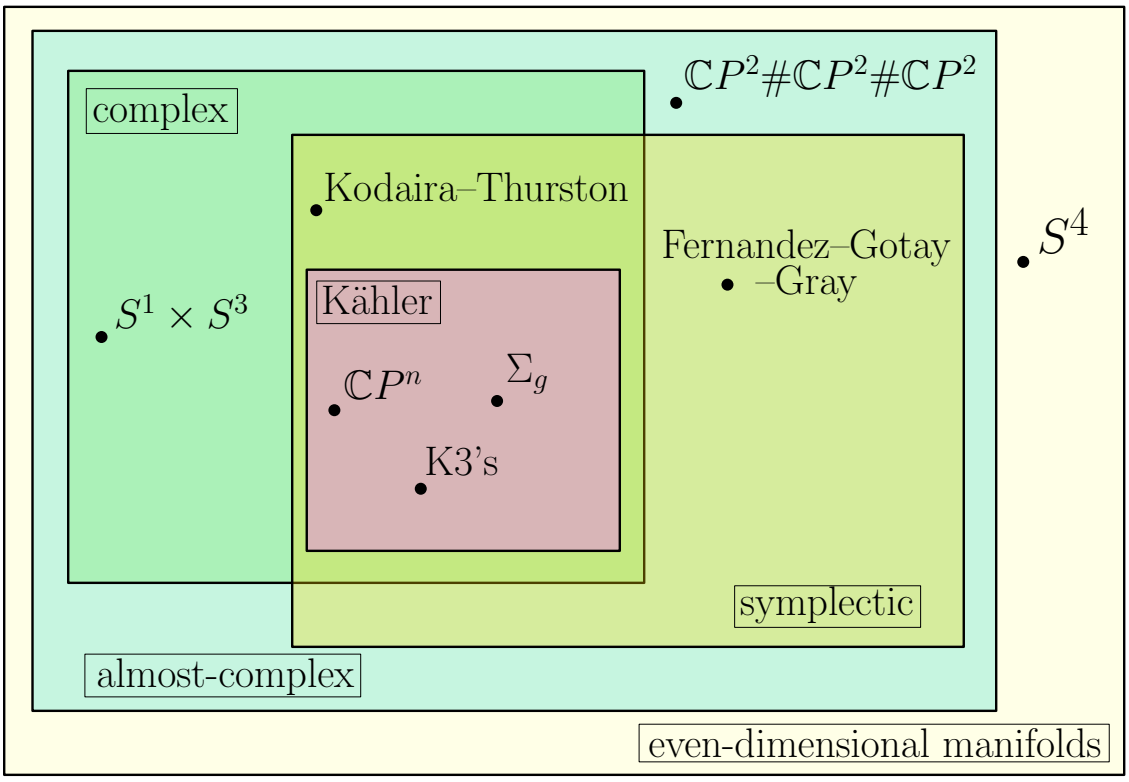



FIGURE 4. Chart of the relations between almost–complex, complex, symplectic, and Kähler manifolds.

Note that in general determining where a manifold lives on this chart may be very difficult. For example, the standard almost-complex structure on $S^6$ is not integrable as follows from the non-associativity of octonion multiplication. It is a well known open problem to determine if $S^6$ admits a complex structure. As an example of the difficulty of determining symplectic structures, the fact $\mathbb{C}P^2\#\mathbb{C}P^2\#\mathbb{C}P^2$ is not symplectic is due to Taubes and requires Seiberg–Witten theory.

*Pseudoholomorphic Curves.* Complex and almost-complex geometry has become an integral part of symplectic geometry even beyond the study of complex and Kähler symplectic manifolds. The key notion, directly related to an almost-complex structure, is the following.

**Definition 2.22: Pseudoholomorphic Curve**

Let $(M, J)$ be an almost-complex manifold. A *pseudoholomorphic curve* or $J$-holomorphic curve is a smooth map $u : \Sigma : J \to M$ from a punctured Riemann surface $(\Sigma, j)$ satisfying the non-linear Cauchy–Riemann equation,

$$\overline{\partial}_J u := \frac{1}{2}(\mathrm{d}u + J \circ \mathrm{d}u \circ j) = 0.$$

We will be interested in studying these curves in a symplectic manifold $(M, \omega)$ equipped with a compatible or tame almost-complex structure. This equation can be seen as describing some kind of symplectic minimal surface. Embedded compatible $J$-holomorphic curves are in particular two dimensional symplectic submanifolds (and the converse is always true for some choice of $J$).

Pseudoholomorphic curves share many properties with holomorphic curves in a complex manifold that are familiar from studying the Cauchy–Riemann equations in complex analysis. This includes unique continuation, elliptic regularity results, and intersection positivity phenomena. One is also interested in the moduli space of all pseudoholomorphic curves satisfying certain constraints modulo automorphisms. In some cases, this defines a finite dimensional smooth manifold with natural compactification whose properties carry important information about the symplectic manifold.

$J$-holomorphic curves were first seriously applied in symplectic geometry by Gromov, who used them to begin the study of symplectic embeddings.

**Theorem 2.23: Gromov Non-Squeezing**

Let $(B^{2n}(r), J)$ be the $2n$-dimensional open ball of radius $r$ with its standard complex structure and let $(Z^{2n}(R), J)$ be the open cylinder $Z^{2n}(R) = B^2(R) \times \mathbb{R}^{2n-2}$ also with its standard symplectic structure. Then there is a symplectic embedding,

$$\varphi : (B^{2n}(r), J) \overset{s}{\hookrightarrow} (Z^{2n}(R), J) \quad \text{if and only if } r \leq R.$$

Since Gromov's work, $J$-holomorphic curves have found numerous applications in symplectic topology. They also are the basis for defining *Gromov–Witten invariants* which are given by evaluating certain cohomology classes on the moduli space of $J$-holomorphic curves. The construction of various Floer theories relies on studying $J$-holomorphic curves or a zeroth order perturbation of them.

We will discuss pseudoholomorphic curves in serious detail later in these notes.

## 3. Classical Mechanics

The origins of symplectic geometry lie in the study of classical mechanics and its Hamiltonian formulation. In fact one could say that Hamiltonian mechanics is just the study of the dynamics of an arbitrary (quadratic at infinity) Hamiltonian flow on $(T^*M, \omega_{\text{can}})$.

**3.1. Hamilton's Equations.** To expand out the relation to mechanics, consider a particle constrained to move on some smooth manifold $M$ (if we want multiple particles, we should just consider the product space $M^n$). The manifold $M$ is called the *configuration space* of our mechanical system. Take coordinate $q_1, \ldots, q_n$ on the manifold. Associated to these coordinates are momenta $p_1, \ldots, p_n$ which we consider as the coordinates $\mathrm{d}q_1, \ldots, \mathrm{d}q_n$ on $T^*M$. The cotangent bundle $T^*M$ is called the *phase space* of our system.

Now pick a Hamiltonian $H(q_i, p_i)$ on $(T^*M, \omega_{\text{can}})$. We will consider this to represent the energy of our system written in terms of the position and momenta. The trio $(T^*M, \omega_{\text{can}}, H)$ is called a *Hamiltonian system*; it is a special type of continuous dynamical system. As we have defined before, the Hamiltonian vector field $X_H$ is determined by,

$$\omega(\alpha, X_H) = \mathrm{d}H(\alpha).$$

This should define the flow of our Hamiltonian system. Let $f : T^*M \times \mathbb{R} \to \mathbb{R}$ be a time-dependent smooth function (what physicists might call an "observable"). Suppose $\gamma(t)$ is a integral curve of $X_H$; we can determine how $f$ changes along the flow of $X_H$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\gamma^* f = \gamma^* \frac{\partial f}{\partial t} + \gamma^* \mathcal{L}_{X_H} f = \gamma^* \frac{\partial f}{\partial t} + \gamma^* X_H(f) = \gamma^* \frac{\partial f}{\partial t} + \gamma^* \{H, f\}.$$

One informally forgets the $\gamma^*$ and writes,

$$\frac{\mathrm{d}}{\mathrm{d}t} f = \frac{\partial f}{\partial t} + \{H, f\}.$$

This is the *general Hamilton equation* describing the time evolution of an observable. In our same system of local coordinates, $(q_1, \ldots, q_n, p_1, \ldots, p_n)$, since $\omega = \sum \mathrm{d}p_i \wedge \mathrm{d}q_i$, we have,

$$X_{p_i} = \frac{\partial}{\partial q_i} \quad \text{and} \quad X_{q_i} = -\frac{\partial}{\partial p_i}.$$

The Hamilton equation for the position and momenta along an integral curve becomes,

$$\frac{\mathrm{d}}{\mathrm{d}t}q_i = \{H, q_i\} = \frac{\partial H}{\partial p_i}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}p_i = \{H, p_i\} = -\frac{\partial H}{\partial q_i}.$$

These are *Hamilton's equations*. They are equivalent to Newton's laws (in an arbitrary coordinate system).

EXAMPLE 3.1. Let us see an explicit example of this where it reduces to a very ordinary description of classical mechanics. Consider $\mathbb{R}^n$ with coordinates $q_1, \ldots, q_n$. We can form $T^*\mathbb{R}^n = \mathbb{R}^{2n}$ with fibre coordinates $p_1, \ldots, p_n = \mathrm{d}q_1, \ldots, \mathrm{d}q_n$ and globally,

$$\omega_{\mathrm{can}} = \sum_{i=1}^{n} \mathrm{d}p_i \wedge \mathrm{d}q_i \,.$$

Suppose we have some Hamiltonian of the form,

$$H(p, q) = \frac{1}{2m}\|\vec{p}\|^2 + V(\vec{q}),$$

where $V(q)$ is smooth function on the base, corresponding to the potential energy and $m$ is a parameter representing mass. Hamilton's equations become,

$$\frac{\mathrm{d}}{\mathrm{d}t}\vec{q} = \frac{\vec{p}}{m} \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}t}\vec{p} = -\nabla V(\vec{q}).$$

The first equation says that $\vec{p} = m\vec{v}$, where $\vec{v} = \dot{\vec{q}}$ is the velocity of a particle moving on our integral curve; i.e. $\vec{p}$ corresponds to momentum. The second says that the change in momentum is minus the gradient of $V$. Recall that force $F$ in classical mechanics is defined to be $-\nabla V$. Hence Hamilton's equations combine to yield,

$$\vec{F} = m\vec{a},$$

where $\vec{a} = \ddot{\vec{q}}$ is acceleration. We have just recovered Newton's second law. ▲

EXAMPLE 3.2. Consider a smooth manifold $M$ equipped with a metric $g$. The inner product $g$ on the fibres of $TM$ induces one on $T^*M$ under the canonical isomorphism of $TM$ and $T^*M$ determined by $g$. We can define a Hamiltonian on $(T^*M, \omega_{\mathrm{can}})$ given by,

$$H(x, v) = \frac{1}{2}\|v\|_g^2.$$

This can be thought of as the Hamiltonian of a free particle on $M$, experiencing no force. Let $X_H$ be the Hamiltonian vector field. Now suppose $q_1, \ldots, q_n$ are

local coordinates on $M$ so that the metric $g$ on $TM$ is represented by the matrix $(g_{ij})$ acting on $\frac{\partial}{\partial q_i}$ and the inverse matrix $(g^{ij})$ acting on $p_i = \mathrm{d}p_i$. Hamilton's equations yield, in Einstein notation,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}q^i &= \frac{\partial}{\partial p_i}\left[\frac{1}{2}g^{jk}p_jp_k\right] \quad = g^{ij}p_j \\ \frac{\mathrm{d}}{\mathrm{d}t}p_i &= -\frac{\partial}{\partial q^i}\left[\frac{1}{2}g^{jk}p_jp_k\right] = -\frac{1}{2}p_jp_k\frac{\partial g^{jk}}{\partial q^i}.\end{aligned}$$

Together these give,

$$\begin{aligned}\ddot{q}_i &= -\frac{1}{2}\dot{q^k}\dot{q^j}\frac{\partial g_{jk}}{\partial q^i} + \frac{\partial g_{ij}}{\partial q^k}\dot{q^j}\dot{q^k} \\ &= \dot{q^j}\dot{q^k}\cdot\frac{1}{2}\left(\frac{\partial g_{ij}}{\partial q^k} + \frac{\partial g_{ik}}{\partial q^j} - \frac{\partial g_{jk}}{\partial q^i}\right) \\ &= g_{i\ell}\Gamma^{\ell}_{jk}\dot{q^j}\dot{q^k}\end{aligned}$$

Thus the flow satisifies,

$$\frac{\mathrm{d}^2q^i}{\mathrm{d}t^2} + \Gamma^i_{jk}\frac{\mathrm{d}q^j}{\mathrm{d}t}\frac{\mathrm{d}q^k}{\mathrm{d}t} = 0.$$

This is the *geodesic equation*. Hence the Hamiltonian flow of $H$ on $T^*M$ reproduces the (co)geodesic flow with respect to $g$.

More generally, one can consider the Hamiltonian,

$$H(x,v) = \frac{1}{2}\|v\|_g^2 + V(x),$$

for a function $V$ on the base. The flow of $X_H$ is now that of a particle subject to a conservative force $-\nabla_g V(x)$.

Another generalization is to twist $(T^*M, \omega_{\mathrm{std}})$. Let $\sigma$ be a closed two-form on $M$ and let $\pi : T^*M \to M$ be the projection to the base. Then, $\omega_\sigma = \omega_{\mathrm{std}} + \pi^*\sigma$ defines a new symplectic form on $T^*M$. In this case we call $\sigma$ a *magnetic form* and $(T^*M, \omega_\sigma)$ a *magnetic cotangent bundle*. The flow $X_H$ of the above Hamiltonian on this manifold describes the motion of a particle subject to the force $-\nabla_g V(x)$ and a magnetic field $\sigma$. ▲

**3.2. Poisson Manifolds.** There is an interesting generalization of symplectic geometry to manifolds admitting some kind of bracket structure.

A *Poisson algebra* is an associative algebra $A$ over a field equipped with a *Poisson bracket* $\{\cdot,\cdot\}$ satisfying the following conditions:

(1) The bracket is a Lie bracket $\{\cdot,\cdot\} : A \times A \to A$, i.e. skew-symmetric, bilinear, and satisfying the Jacobi identity.

(2) The bracket is a derivation in its entries:

$$\{f, gh\} = \{f, g\}h + g\{f, h\}.$$

**Definition 3.3**

A *Poisson manifold* is a smooth manifold $M$ equipped with a Poisson bracket $\{\cdot, \cdot\}$ on its algebra of smooth functions $C^\infty(M)$. I.e. the bracket satisfies the following properties:
**Skew-Symmetry:** $\{f, g\} = -\{g, f\}$,
**Jacobi Identity:** $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$.
**Leibniz Rule:** $\{f, gh\} = \{f, g\}h + \{f, h\}g$.

Equivalently, a Poisson manifold is a smooth manifold $M$ equipped with a *Poisson bivector field* $\pi \in \Gamma(\bigwedge^2 TM)$ satisfying $[\pi, \pi] = 0$. Here, $[\cdot, \cdot]$ is the *Schouten bracket*, which is the unique bracket on multivectors making $\Gamma(\bigwedge^* TM)$ a super Lie algebra so that $[u, \cdot]$ is a degree $k$ graded superderivation for any $u \in \Gamma(\bigwedge^k TM)$. The bracket and bivector field are related by,

$$\{f, g\} = \pi(\mathrm{d}f\,, \mathrm{d}g).$$

A smooth map $f : (M, \pi) \to (N, \varpi)$ is called a *Poisson map* if the Poisson structures are $f$-related, i.e. $f_*\pi = \varpi$. If a Poisson map $f : M \to N$ is a diffeomorphism, then we call $f$ a *Poisson diffeomorphism* or[a] an *icthyomorphism*.

[a] With tongue in cheek.

Exercise 3.4. Show that if three points on an affine line in $\Gamma(\bigwedge^2 TM)$ are Poisson, then the whole line is. This is called a *Poisson pencil*. ▲

Example 3.5. Let us consider some important examples of Poisson manifolds.

(1) Every symplectic manifold $(M, \omega)$ is Poisson via the usual Poisson bracket,

$$\pi(\mathrm{d}f\,, \mathrm{d}g) = \{f, g\} = \omega(X_f, X_g).$$

The Poisson bivector field and $\omega$ are related by the musical isomorphisms they induce $\pi^\sharp = (\omega^\flat)^{-1}$. In fact symplectic manifolds are the same as *non-degenerate* Poisson structures, i.e. those for which $\pi^\sharp$ is an isomorphism.

(2) Unlike symplectic manifolds, there are no topological restrictions on a manifold being Poisson. In particular, as a rather dull example, every manifold $M$ is Poisson with respect to the trivial bracket $\pi = 0$.

(3) If $\mathfrak{g}$ is a finite dimensional Lie algebra with dual $\mathfrak{g}^*$, then $\mathfrak{g}^*$ has a linear Poisson structure due to Kirillov–Kostant–Souriau. Say $x_i$ is a basis for $\mathfrak{g}$ with dual basis $\mu_i$, so that $\mathfrak{g}$ has structure constants $c^i_{jk}$ defined by $[x_j, x_k] = \sum c^i_{jk} x_i$. Then, the Poisson structure is,

$$\pi = \frac{1}{2} \sum_{ijk} c^i_{jk}\, \mu_i \, \frac{\partial}{\partial \mu_j} \wedge \frac{\partial}{\partial \mu_k}.$$

In fact every linear (in the sense of first order coefficients) Poisson structure on a finite vector space is of this form.

(4) The bracket $[\pi, \pi]$ of bivector fields is degree three. Hence any bivector field on a surface is Poisson. ▲

We may define Hamiltonian vector fields for Poisson structures as for symplectic structures. The vector field $X_H$ associated to $H$ is simply given by $\pi^\sharp \, \mathrm{d}H = X_H$. The collection of all Hamiltonian vector fields will span a distribution $\pi^\sharp(T^*M) \subset TM$. This distribution is the entire tangent bundle precisely when $\pi$ is non-degenerate and hence defines a symplectic structure. Note as in the symplectic case, the Lie bracket of Hamiltonian vector fields is Hamiltonian and so this distribution is involutive.

If near a point $x$, $\pi$ is *regular*, so that $\pi^\sharp$ has fixed rank, then the distribution will be locally regular and involutive and so by the Frobenius theorem, a neighbourhood of $x$ will be foliated by leaves on which $\pi$ restricts to a non-degenerate Poisson structure, i.e. a symplectic structure. Even if $\pi$ is not regular near $x$, Weinstein proved using an extension of Darboux's theorem that in a neighbourhood of $x$ one can find coordinates $(q_1, \ldots, q_k, p_1, \ldots, p_k, x_1, \ldots, x_{n-2k})$ so that $\pi$ is the sum of the standard symplectic form in $(q_1, \ldots, q_k, p_1, \ldots, p_k)$ and a Poisson structure $\pi_0$ in terms of $(x_1, \ldots, x_{n-2k})$ so that $(\pi_0)_x = 0$. I.e. $(M, \pi)$ splits at $x$ as the direct sum of a symplectic manifold and a Poisson manifold with vanishing Poisson structure at $x$. We can inductively apply this result to the sub-Poisson structures we obtain and we conclude our distribution defines a *singular foliation*, that is a foliation where the leaves may have differing dimensions.

**Theorem 3.6: Weinstein Splitting Theorem**

Given a Poisson manifold $(M, \pi)$, the distribution $\pi^\sharp(T^*M)$ defines a singular foliation of $M$ so that on each leaf $L \subset M$, the condition,

$$\omega_L^\flat = (\pi|_L^\sharp)^{-1}$$

canonically defines a symplectic form $\omega_L$ on $L$. Hence $M$ admits a singular foliation by *symplectic leaves*.

EXAMPLE 3.7. Let us describe what the symplectic leaves are in the examples of Poisson structures we listed above.

(1) If $(M, \pi)$ is non-degenerate, then it is already symplectic and all of $M$ is a single symplectic leaf.

(2) If $\pi = 0$, then each point of $M$ is a zero dimensional symplectic leaf.

(3) For a linear Poisson structure $\pi$ on $\mathfrak{g}^*$, the symplectic leaves of $\pi$ are precisely the orbits of the coadjoint action of $G$ on $\mathfrak{g}^*$, where $\mathrm{Lie}(G) = \mathfrak{g}$. These leaves may clearly jump rank.

EXERCISE 3.8. Prove this. ▲

(4) Given any bivector $\pi$ on a two-manifold $\Sigma$, it will have some degeneration locus $Z$ where $\pi = 0$. On each connected component of the open complement of $Z$, $\pi$ defines two-dimensional symplectic leaves. On $Z$, $\pi$ will vanish and the symplectic leaves are the points of $Z$.

Here is a basic example of this. Consider $\Sigma = \mathbb{R}P^2$ equipped with the Poisson structure,

$$\pi = (g(x) - y^2)\frac{\partial}{\partial x} \wedge \frac{\partial}{\partial y},$$

defined on $\mathbb{R}^2$ and extended smoothly to the line at infinity. Here $g(x)$ is a cubic polynomial $g(x) = x(x-1)(x-\lambda)$ for some parameter $\lambda \in (0, 1)$. The Poisson manifold will look like that depicted in Figure 5. The degeneration locus $Z$ is the zero locus of the real cubic $y^2 = g(x)$. On $Z$, the symplectic leaves of $\pi$ are points. Away from $Z$, $\mathbb{R}P^2 \backslash Z$ has three connected components which will each be a symplectic manifold with respect to $\pi$.

(5) Generalizing the last example, suppose $(M^{2n}, \pi)$ is a Poisson submanifold so that $\pi^n$ intersects the zero section of $\det(TM)$ transversally. Then the intersection $Z = \pi^n \pitchfork 0$ is a hypersurface in $M$. Away from the hypersurface $Z$, $(M, \pi)$ is symplectic. One can show that all the symplectic leaves of $Z$ are of the maximal dimension $2n-2$. If in addition we assume the leaves of $M$ are compact, then each connected component of $Z$ is a mapping torus $S^1 \ltimes_\varphi F$, where $F$ is a leaf in $Z$

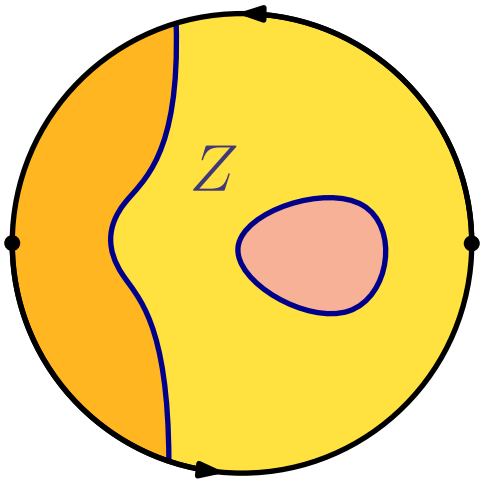



FIGURE 5. A Poisson structure on $\mathbb{R}P^2$ with a cubic degeneration locus $Z$.

and $\varphi \in \mathrm{Symp}(F)$. A Poisson manifold of this form is called a *log symplectic manifold*. ▲

Poisson geometry is an expansive field in of itself. The correspondence between linear Poisson structures and Lie algebras extends to a correspondence between fibrewise linear Poisson structures and *Lie algebroids*, which are like smooth many-object generalizations of Lie algebras. Moreover, any Poisson manifold $M$ induces a Lie algebroid structure on its cotangent bundle. Much of the study of Poisson manifolds can be converted to questions about Lie algebroids, whereupon one can incorporate information from the categorical study of algebroids and groupoids.

Recall (or accept) that the Hamiltonian formulation of quantum mechanics is obtained by converting the Poisson bracket relations of observables (smooth functions) in a classical system to commutator relations of observables (linear operators) in a quantum system. Hence Poisson geometry is a natural context in which to study the mathematics of quantization. This is the theory of *deformation quantization*, explored by Kontsevich.

**3.3. Integrable Systems and the Arnold–Liouville Theorem.** We have seen that Hamiltonian flow necessarily preserves the level sets of the Hamiltonian, i.e. orbits of the Hamiltonian system have constant energy. In many favourable settings and classical examples, one can find other observables (functions) which are conserved by the Hamiltonian flow. These are called *integrals of motion*. The existence of an integral of motion puts further constraints on the dynamics of the system.

One common instance of this is in Noether's theorem, which we will discuss in the section on Hamiltonian actions. Suppose a Lie group $G$ acts on a symplectic manifold $M$ by symplectomorphisms. Under certain conditions described later, we call this action a *Hamiltonian action* and we can associate to it a momentum map $\mu : M \to \mathfrak{g}^*$. Noether's theorem then says that $\mu$ is preserved by the Hamiltonian flow of some function $H$ if and only if $H$ is invariant under the $G$-action. Or in physics speak, $G$ is a symmetry of the Hamiltonian system $(M, \omega, H)$ if and only if $\mu$ is a conserved quantity. From a conserved quantity $\mu$, we can go backwards to obtain a symmetry by taking the group generating the flows of the components of $\mu$ as our $G$-action. For example, often our system may be independent of translations in some direction or rotations about some example, in which case linear momentum in that direction or angular momentum about that axis will be conserved.

Regardless of where it comes from, if $F$ is an observable and $\varphi_H^t$ is the flow of Hamiltonian $H$, we have from the general Hamilton equation,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\varphi_H^t\right)^* F = (\varphi_H^t)^* \{H, F\}.$$

Hence, $F$ is an integral of motion if and only if $F$ and $H$ Poisson commute.

We say a set of functions $f_1, \ldots, f_k$ on $M$ is *functionally independent* if the covectors $(\mathrm{d}f_1)_x, \ldots, (\mathrm{d}f_k)_x$ are linearly independent on an open dense set of $x \in M$, or equivalently $\mathrm{d}f_1 \wedge \cdots \wedge \mathrm{d}f_k$ is non-vanishing almost everywhere. This means that $(f_1, \ldots, f_k) : M \to \mathbb{R}^k$ is a submersion almost everywhere.

**Definition 3.9: Integrable Systems**

A Hamiltonian system $(M^{2n}, \omega, H)$ is *partially integrable* if there is at least one function $F$ functionally independent from $H$ which Poisson commutes with $H$.

The system is *(Liouville) integrable* if there are functionally independent functions $F_1 = H, F_2, \ldots, F_n$ which pairwise Poisson commute:

$$\{F_i, F_j\} = 0 \quad \text{for all} \quad i, j.$$

Note that there can be no more than $n$ functionally independent and simultaneously Poisson-commuting functions. We say the system is *superintegrable* if it is integrable with respect to $F_1 = H, F_2, \ldots, F_n$ and if we may find additional functions $G_1, \ldots, G_k$ so that $F_1, \ldots, F_n, \ldots, G_1, \ldots, G_k$ are functionally independent and,

$$\{G_i, H\} = 0 \quad \text{for all} \quad i.$$

If we may take $k = n$, then we say the system is *maximally superintegrable*.

Note that if our system $(M^{2n}, \omega, H)$ is partially integrable for functions $F_1 = H, F_2, \ldots, F_k$, then the flow must lie in the intersection of the level sets of each conserved function $F_i$. Hence (assuming we are at a regular value) we can restrict the orbits of $X_H$ to $2n - k$ dimensional submanifolds. In the case of Liouville integrability, the orbits are restricted to $n$ dimensional submanifolds. If the system is maximally superintegrable, then the orbits are restricted to one dimensional manifolds. Hence the dynamics of the system are fully determined (up to maybe solving some integrals to determine periods of orbits) by these conserved quantities.

Moreover, say the system is partially integrable for functions $F_1 = H, F_2, \ldots, F_k$, and $N$ is an intersection of regular level sets. Then with respect to some metric, the normal bundle $TN^\perp \subset TM|_N$ is spanned by $\nabla F_1, \ldots, \nabla F_k$. Because the functions Poisson commute, after picking a compatible almost complex structure $J$,

$$\omega(\nabla F_i, \nabla F_j) = \omega(JX_{F_1}, JX_{F_2}) = \omega(X_{F_1}, X_{F_2}) = \{F_1, F_2\} = 0.$$

Hence the fibres of $TN^\perp$ are isotropic linear spaces and so $N$ is a coisotropic submanifold. In the special case of Liouville integrability, the intersection of level sets $N$ will be a Lagrangian submanifold.

This implies that the map $F = (F_1, \ldots, F_n) : M^{2n} \to \mathbb{R}^n$ is a *Lagrangian fibration* when restricted to the preimage of the regular values.

PROPOSITION 3.10. *If the Hamiltonian vector fields $X_{F_i}$ of an integrable system are all complete on a regular level set $F^{-1}(c)$, then each connected component of $F^{-1}(c)$ is canonically diffeomorphic to an affine copy of $\mathbb{R}^{n-k} \times T^k$ for some $k$. In particular, if $F$ has compact connected fibres, then $F$ defines an affine torus bundle over the regular set.*

PROOF. Let $N$ denote a connected component of $F^{-1}(c)$ a regular Lagrangian level set. We know the functions $F_i$ Poisson commute. But then,

$$\mathcal{L}_{X_{F_i}} X_{F_j} = [X_{F_i}, X_{F_j}] = X_{\{F_i, F_j\}} = 0.$$

This implies the flows of the Hamiltonian vector fields commute and define an $n$ parameter abelian group acting by diffeomorphisms. Thus, picking a basepoint, there is a smooth map, $g : \mathbb{R}^n \to N$ which is a locally transitive group action on a connected manifold, hence transitive. This gives $N$ the structure of a principal homogeneous space $\mathbb{R}^n/\Gamma$, where $\Gamma$ is the stabilizer of

an element of $N$ (since the action is transitive, the stabilizer are all conjugate). $\Gamma$ must be a discrete subgroup of $\mathbb{R}^n$, i.e. a $k$-dimensional lattice for $k \leq n$. Thus $N$ is diffeomorphic to $\mathbb{R}^{n-k} \times T^k$. Changing the basepoint amounts to a translation of the action by $g$, so the diffeomorphism is canonical considering $\mathbb{R}^{n-k} \times T^k$ as an affine space (i.e. a group torsor). □

REMARK 3.11. This result also holds for general Lagrangian fibrations (see [**Mein**, Ch. 5]). △

Now suppose that $F$ has compact connected fibres, so that we have an affine torus bundle $F : M^{2n} \to B$. The fibrewise action $\mathbb{R}^n \to M$ can be recast as a vector bundle action $T^*B \times M \to M$. Here, we send a one-form $\alpha \in \Omega^1(B)$ to the corresponding vector field $X_\alpha$ under $\omega^\sharp \circ F^*$. Because the Hamiltonian vector fields $X_{F_i}$ are dual to $\mathrm{d}F_i$ which are constant on the fibres of $F$, $X_\alpha$ will be a vertical vector field. The time one flow of $X_\alpha$ then defines a map $M \to M$ preserving the fibres of $F$. All the constructions were linear so this gives a linear vector bundle action as claimed.

Let $\Lambda \subset T^*B$ denote the bundle of stabilizers of this bundle action. Over each point of $B$, we know $\Lambda$ is a copy of $\mathbb{Z}^n$.

LEMMA 3.12. *The stabilizer $\Lambda$ is a Lagrangian submanifold of $T^*B$.*

PROOF. On a small neighbourhood $U \subset B$, $\Lambda|_U$ trivializes as $U \times \mathbb{Z}^n$ whose sheets are graphs of 1-forms $\alpha$ on $B$. Suppose the flow associated to $\alpha$ is $\Phi^t_\alpha$. Using the Lie derivative,

$$(\Phi^t_\alpha)^*\omega = \omega + t\pi^* \,\mathrm{d}\alpha\,.$$

Because $\alpha$ stabilizes, $F^1_\alpha = 0$ and so $\mathrm{d}\alpha = 0$. The graph of a closed one-form is always Lagrangian. Thus $\Lambda$ is Lagrangian over $U$, and hence globally. □

LEMMA 3.13. *If $U \subset B$ is contractible, then it possesses a section $s : U \to M$ of the Lagrangian fibration $F : M \to B$ so that $s^*\omega = 0$. Furthermore this section defines a symplectomorphism $T^*B/\Lambda|_U \to M|_U$, with the domain's symplectic structure induced from $T^*B$.*

PROOF. Because $U$ is contractible, for any section $s$, $s^*\omega$ is exact: $s^*\omega = -\mathrm{d}\beta$. But then, if $\Phi_\beta$ is the time one flow of $X_\beta$,

$$(\Phi_\beta \circ s)^*\omega = s^*(\omega + \pi^*\mathrm{d}\beta) = \sigma^*\omega + \mathrm{d}\beta = 0.$$

So, we may take the section to be $\Phi_\beta \circ s$.

The choice of $s$ defines an isomorphism $T^*B/\Lambda|_U \to M|_U$ by sending the zero section of $T^*B$ to $s$ and asking the isomorphism be equivariant under the $T^n$ action on fibres. We may lift this map to $\varphi : T^*B|_U \to M|_U$ and we wish to show $\varphi^*\omega = \mathrm{d}\theta_{\mathrm{can}}$.

Because our bundle isomorphism was $T^n$-equivariant, we have $\varphi \circ \alpha = \Phi_\alpha \circ s$ for any one form $\alpha \in \Omega^1(U)$, where $\Phi_\alpha$ again denotes the time one flow of $X_\alpha$. Thus,

$$\begin{aligned}\alpha^*\varphi^*\omega &= s^*\Phi_\alpha^*\omega \\ &= s^*(\omega + \mathrm{d}\pi^*\alpha).\end{aligned}$$

Since $s^*\omega = 0$,

$$\begin{aligned}&= \mathrm{d}s^*\pi^*\alpha \\ &= \mathrm{d}\alpha.\end{aligned}$$

Since $\theta_{\mathrm{can}}$ is tautological,

$$\begin{aligned}&= \mathrm{d}\alpha^*\theta_{\mathrm{can}} \\ &= \alpha^*\mathrm{d}\theta_{\mathrm{can}}.\end{aligned}$$

Since this holds for any $\alpha$, we have $\varphi^*\omega = \mathrm{d}\theta_{\mathrm{can}}$. □

Taking $U \subset B$ as in the above lemma, we may trivialize $\Lambda|_U \cong U \times \mathbb{Z}^n$. Under this isomorphism, the standard basis of $\mathbb{Z}^n$ defines sections of $\Lambda$ $\beta_1, \ldots, \beta_n \in \Omega^1(U)$. Note by the proof of lemma 3.12 that the $\beta_i$ are closed. Because $U$ is contractible, they are in fact exact: $\beta_i = \mathrm{d}I_i$.

Note $T^*U \cong U \times \mathbb{R}^n$ and so $M|_U \cong T^*B/\Lambda \cong U \times T^n$, a symplectomorphism. The projection $\pi_2 : M|_U \to T^n$ can be decomposed into components, and arbitrarily choosing a base point these lift to functions $\theta_1, \ldots, \theta_n : M|_U \to \mathbb{R}$. Note that under the flow $\Phi_i^t$ of $X_{\beta_i}$, we have,

$$\theta_j \mapsto \theta_j + \delta_{i,j}t \pmod 1.$$

We conclude that the coordinates $\theta_i$ on the cotangent bundle correspond to the differentials of the coordinates $I_i$ on the base. Hence pulled back to the total space, the symplectic form $\omega_{\mathrm{can}}$ of $T^*B$ restricted to $U$ is given by,

$$\omega_{\mathrm{can}}|_U = \sum_i \mathrm{d}\theta_i \wedge \mathrm{d}I_i\,.$$

By lemma 3.13, we can apply the symplectomorphism to obtain the same description of $\omega$ on $M|_U$.

**Theorem 3.14: Arnold–Liouville Theorem**

Let $(M^{2n}, \omega, H)$ be a Liouville integrable Hamiltonian system with integrals of motion $F = (F_1 = H, \cdots, F_n) : M \to \mathbb{R}^n$. Suppose $F$ is proper with connected fibres, let $B$ denote the open dense subset of the image of $F$ of regular values of $F$ and let $M'$ denote the set of preimages of regular values. Then $F : M' \to B$ is a Lagrangian torus fibration.

The fibres of $F$ are invariant under the flow of $X_H$, and on each tori the flow is linear in specified coordinates. In particular, locally on $M'$, there are coordinates $(I_1, \ldots, I_n, \theta_1, \ldots, \theta_n)$ so that the symplectic form is standard,

$$\omega = \sum_i \mathrm{d}\theta_i \wedge \mathrm{d}I_i \,,$$

and in these coordinates the flow of $X_H$ is given by,

$$I_i(t) = I_i(0) \quad \text{and} \quad \theta_i(t) = \theta_i(0) + t\frac{\partial H}{\partial I_i} \quad \text{for each } i.$$

Proof. We have proved almost all of this already. We should just check the flow of $X_H$ in the coordinates described above is as claimed. We know that $H$ is constant on the level set of $F$ and hence $\frac{\partial H}{\partial s_i} = 0$ for each $i$. Because we have defined Darboux coordinates for the symplectic form, we can work out explicitly that,

$$X_H = \sum_i \frac{\partial H}{\partial I_i}\frac{\partial}{\partial s_i} - \cancelto{0}{\frac{\partial H}{\partial s_i}}\frac{\partial}{\partial I_i} = \sum_i \frac{\partial H}{\partial I_i}\frac{\partial}{\partial s_i}.$$

It is now immediate that the flow of $X_H$ is as claimed. □

**Definition 3.15**

The coordinates $(I_1, \ldots, I_n, s_1, \ldots, s_n)$ are called *action-angle variables* for the integrable system. The $I_i$'s are action coordinates and the $s_j$'s are angle coordinates. These coordinates also exist locally for proper Lagrangian fibrations. See [**Mein**] for the general case and also for an explanation on conditions under which the action-angle variables are defined globally; in particular we need $B$ simply connected and the fibration to admit a section.

*The Spherical Pendulum.* To see our theorem in action, we now consider a simple but interesting example of an integrable system, namely the spherical pendulum. Suppose we have a rigid massless rod in $\mathbb{R}^3$ of length one with

one end fixed to the origin. On the other end of the rod is a point particle of mass $m$ which experiences only the force of gravity. We wish to study the Hamiltonian dynamics of this particle.

The configuration space of the particle is $S^2$ and the phase space is $T^*S^2$. Let us take spherical coordinates on $S^2$: $\varphi \in [0, 2\pi]$ and $\psi \in (0, \pi)$ so that,

$$x = \sin\psi\cos\varphi, \quad y = \sin\psi\sin\varphi, \quad \text{and} \quad z = \cos\psi.$$

These have conjugate momenta which we call $p_\varphi$ and $p_\psi$. The gravitational potential energy gravity is $mgz$ and the kinetic energy is $\frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{x}^2)$. Converting to spherical coordinates and normalizing $m = g = 1$, the energy should be,

$$H = \frac{1}{2}\left(p_\psi^2 + \frac{1}{\sin^2\psi}p_\varphi^2\right) + \cos\psi.$$

The singularity here is only due to the singularity of the coordinate frame.

Because the system is symmetric under rotations about the $z$-axis, the corresponding angular momentum $p_\varphi$ should be an integral of motion. Equivalently, since $H$ is independent of $\varphi$, we easily compute $\{H, p_\varphi\} = 0$. Thus we have two integrals of motion $(H, p_\varphi)$ in a four dimensional Hamiltonian system, and hence the system is Liouville integrable.

Consider the map $F = (H, p_\varphi) : T^*S^2 \to \mathbb{R}^2$. The image of this map is the region of the plane,

$$\left\{(H, p) : \quad \text{where} \quad |p| \leq \max\left\{\sin(\psi)\sqrt{2H - 2\cos(\psi)} : \psi \in [0, 2\pi]\right\}\right\}.$$

This will look roughly parabolic in $p$ as a function of $H$, but singular at the vertex $(-1, 0)$. The critical values of $F$ are along the boundary of the image and at the point $(1, 0)$. Away from this set, the preimages of $F$ are all tori. Consider the following sketch in Figure 6.

The preimage of $(-1, 0)$ is a single point, corresponding to when the particle is stationary at its stable minimum. The preimage of all other boundary points is a circle corresponding to the particle orbiting at a fixed height. The preimage of the point $(1, 0)$ is a pinched torus; the singular point corresponds to the unstable equilibrium where the particle is stationary with the rod straight upwards.

All other points in the image have a copy of $T^2$ as their inverse image. This is easiest to see for points along the axis $p_\varphi = 0$. If the particle has energy $H$ and no $p_\varphi$ angular momentum, then it will undergo simple harmonic motion in some fixed plane through the $z$-axis of amplitude proportional to $H$. There is an $S^1$ family of such planes based on the angle $\varphi$, and in each plane the simple

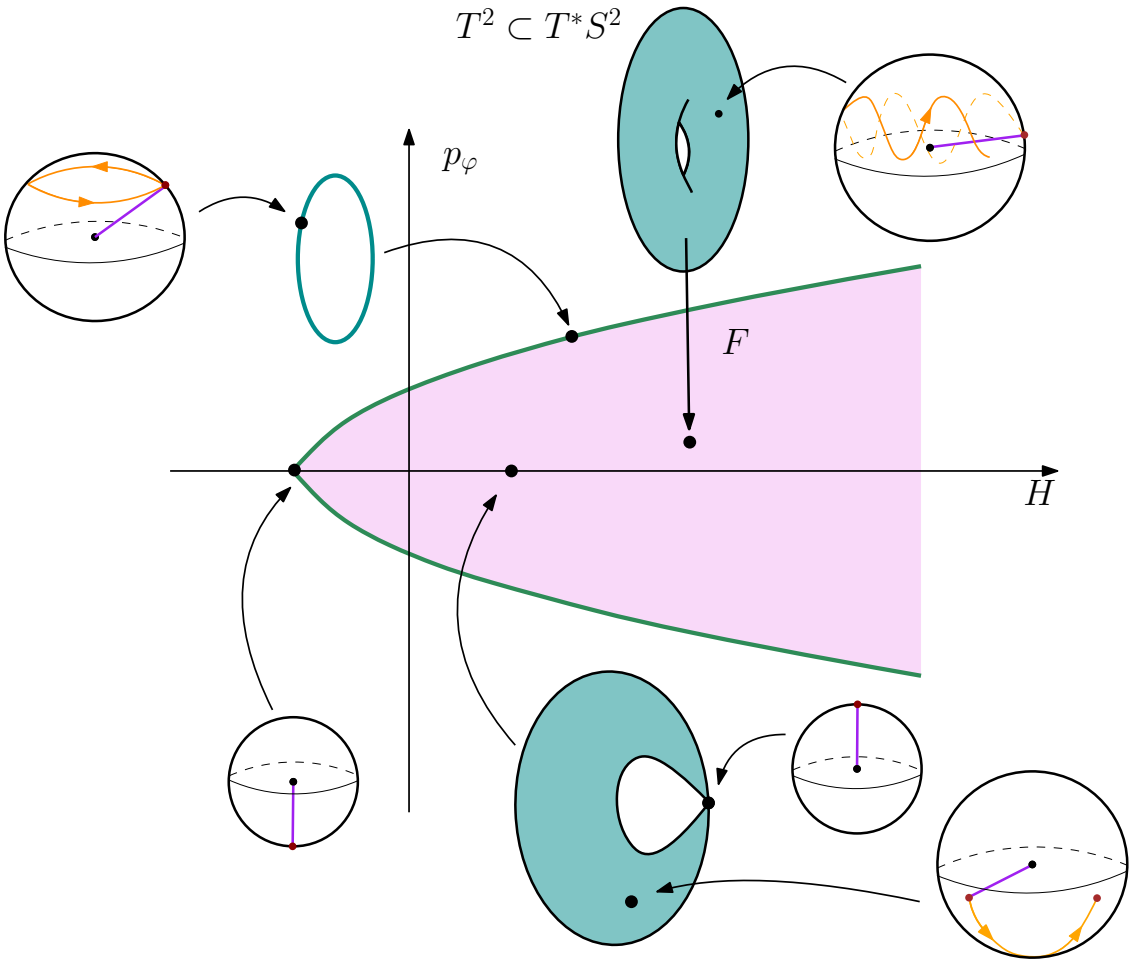



FIGURE 6. The momentum mapping of the spherical pendulum as a Lagrangian torus fibration and its special fibres.

harmonic orbit will contain an $S^1$ family of states. Hence, $F^{-1}(H, 0) \cong T^2$ for $H \neq \pm 1$.

For a generic point in a generic preimage, the motion will look oscillatory as in the top right of Figure 6. It must remain within the torus fibre but in general will not be periodic unless the derivatives of $H$ with respect to the action variables $\frac{\partial H}{\partial I_1}$, $\frac{\partial H}{\partial I_2}$ are rationally linearly dependent.

Note that in this case the set of regular values is not simply connected due to the puncture at $(1, 0)$. Indeed it can be shown the action-angle variable are not defined globally due to the monodromy about this point.

## 4. Contact Manifolds

Contact geometry is often described as the odd dimensional analogue of symplectic geometry. Maybe more accurately, it is the geometry of suitably nice energy hypersurfaces in a symplectic manifold. In many ways the structures of contact and symplectic manifolds are closely related, but contact manifolds have unique geometric features which we only touch upon.

**4.1. Contact Structures.** We begin with fundamental definitions. Rather than a tensor field, the basic object of contact geometry is a distribution.

**Definition 4.1: Contact Manifold**

A *contact manifold* $(Y^{2n+1}, \xi)$ is a $2n+1$ manifold $Y$ equipped with a hyperplane distribution $\xi \subset TM$ which is maximally non-integrable in the sense that there are no integral submanifolds of $\xi$ of dimension bigger than $n$. If locally $\xi$ can be written as the kernel of a one form $\alpha$, then by the Frobenius theorem, maximal non-integrability is equivalent to $\mathrm{d}\alpha$ being non-degenerate on $\xi$, or also the condition,

$$\alpha \wedge (\mathrm{d}\alpha)^n \neq 0.$$

If $\xi$ is co-orientable, then such a one form $\alpha$ is defined globally. In this case we call $\alpha$ a *contact form*. Note $\alpha \wedge (\mathrm{d}\alpha)^n$ will be a volume form so in particular $Y$ must be orientable.

A diffeomorphism $\varphi : (Y, \xi) \to (Y', \xi')$ of contact manifolds is a *contactomorphism* if $\varphi_* \xi = \xi'$. If the contact structures are given by one forms $\alpha, \alpha'$, this is equivalent to asking $\alpha = f \cdot \varphi^* \alpha'$ for some non-vanishing function $f$.

We will almost always assume a contact form exists, but it is highly non-unique since multiplying a contact form by a non-vanishing function gives another contact form (and any two contact forms for $\xi$ are thus related). We call the contact form *positive* or *negative* based on the sign of $\alpha \wedge (\mathrm{d}\alpha)^n$.

Recalling our above comment that contact manifolds are energy hypersufaces, this suggests that contact manifolds have some kind of intrinsic dynamics. Note that any choice $\alpha$ of locally defining one form of $\xi$ determines a line bundle $L$ so that $\mathrm{d}\alpha(L, \cdot) = 0$. A different choice of local contact form will not change $L$. One has $TY = L \oplus \xi$.

**Definition 4.2: Reeb Vector Field**

For a choice of global contact form $\alpha$, the conditions,

$$\mathrm{d}\alpha(R, \cdot) = 0 \quad \text{and} \quad \alpha(R) = 1$$

uniquely define a vector field $R \in \Gamma(TY)$ tangent to $L$. This is called the *Reeb vector field* of $\alpha$. An integral curve of $R$ is called a *characteristic*; these define the *characteristic foliation* of $Y$.

The notion of Lagrangian submanifolds also has a contact analogue.

**Definition 4.3: Legendrian Submanifold**

A *Legendrian submanifold* $S \subset Y$ of a contact manifold $(Y, \xi)$ is a maximal integral submanifold of $\xi$. So, if $Y$ is $2n+1$ dimensional, then a Legendrian submanifold is any $n$ dimensional submanifold tangent to $\xi$.

If $Y$ is three dimensional, a Legendrian submanifold is a 1-manifold. A compact connected Legendrian submanifold of a 3-manifold (usually $\mathbb{R}^3$ or $S^3$) is called a *Legendrian knot*.

An important geometric dichotomy in 3-dimensional contact geometry is the following.

**Definition 4.4**

An *overtwisted disc* $D$ is an embedded disc in a contact 3-manifold $(Y, \xi)$ if its boundary is Legendrian and $TD$ is transverse to $\xi$ except at one point. $Y$ is called *overtwisted* if it has an overtwisted disc, otherwise it is called *tight*.

Eliashberg has shown that up to isotopy, overtwisted contact structures are classified by homotopy classes of tangent plane fields. The existence and classification of tight contact structures is much more subtle.

**4.2. Examples.** Let us see some important examples of contact manifolds.

(1) Consider $\mathbb{R}^{2n-1}$ with coordinates $(x_1, \ldots, x_{n-1}, y_1, \ldots, y_{n-1}, z)$ and its standard contact form,

$$\lambda_{\mathrm{std}} = \mathrm{d}z - \sum_{i=1}^{n-1} y_i \, \mathrm{d}x_i \, .$$

As in Darboux's theorem, this is the local model for all contact manifolds. In $\mathbb{R}^3$, this is $\lambda_{\text{std}} = \mathrm{d}x - y\,\mathrm{d}x$. The associated contact distribution, which we call $\xi_{\text{std}}$, is spanned by $\partial_y$ and $\partial_x + y\partial_z$. The planes of $\xi_{\text{std}}$ will look as depicted in Figure 7 at height $z = 0$ (other heights look the same, just translated).

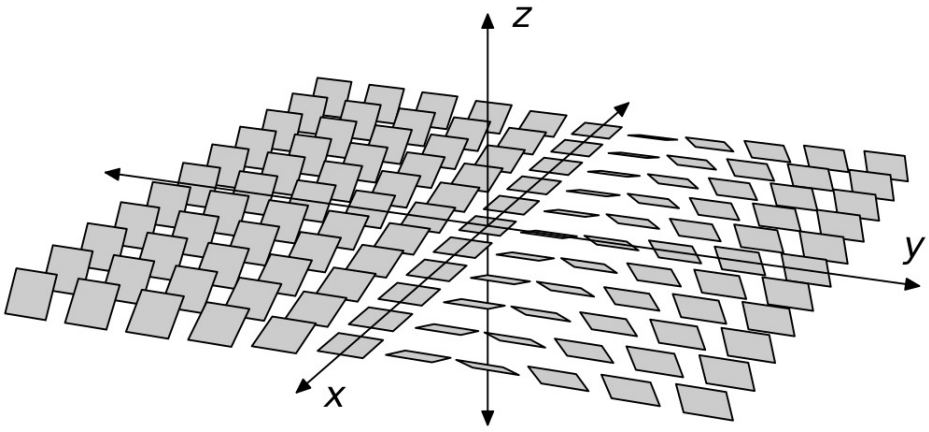



FIGURE 7. The contact distribution $\xi_{\text{std}}$ in $\mathbb{R}^3$.

The Reeb vector field of $\lambda_{\text{std}}$ is just $R = \partial_z$. So the characteristics are just vertical lines.

This contact structure is tight and hence a contact structure being overtwisted is a global feature.

(2) Consider the sphere $S^{2n-1} \subset \mathbb{R}^{2n}$. Let $\mathcal{E}$ be the Euler or radial vector field $\mathcal{E} = \sum x_i \frac{\partial}{\partial x_i} + y_i \frac{\partial}{\partial y_i}$ in $\mathbb{R}^{2n}$. The standard symplectic form has a primitive,

$$\lambda_0 = i_{\mathcal{E}}\omega_{\text{std}} = \sum_{i=1}^{n} x_i\,\mathrm{d}y_i - y_i\,\mathrm{d}x_i\,.$$

It is a simple exercise to check $\lambda_0$ restricts to a contact one form on $S^{2n-1}$. In fact for any hypersurface in $\mathbb{R}^{2n}$ which is starshaped with respect to the origin, the same one-form will induce a contact structure. These are examples of contact type hypersurfaces, which we discuss more shortly. The Reeb vector field of $\lambda_0$ is,

$$R = \sum_{i=1}^{n} x_i \frac{\partial}{\partial y_i} - y_i \frac{\partial}{\partial x_i}.$$

The flow of $R$ is completely periodic, given by $\dot{x} = y, \dot{y} = -x$. The closed characteristics are the circular fibres of the Hopf fibration $S^{2n-1} \to \mathbb{C}P^{n-1}$.

If we remove a point from $(S^{2n+1}, \lambda_0)$, the resulting disk is contactomorphic to $(\mathbb{R}^{2n+1}, \lambda_{\text{std}})$.

(3) We can turn the symplectic cotangent bundle $(T^*M, \lambda_{\text{can}})$ into a contact manifold in two ways: by dimensional expansion or dimensional reduction.

First, we can consider the 1-jet bundle $J^1M = T^*M \times \mathbb{R}$ with contact form,

$$\alpha = \mathrm{d}z - \lambda_{\text{can}}.$$

The associated Reeb vector field is $\partial_z$. In analogy with the Weinstein tubular neighbourhood theorem, any compact Legendrian submanifold has a tubular neighbourhood contactomorphic to a neighbourhood of the zero section in $J^1M$.

The other construction requires a choice of Riemannian metric $g$ on $M$. The we may form $S^*L$, the unit sphere cotangent bundle. Now $\alpha = \lambda_{\text{can}}|_{S^*L}$ defines a contact form on $S^*L$. The Reeb vector field will yield the (co)geodesic flow of $g$ by our analysis of the symplectic case. This is another example of a contact type hypersurface.

(4) Let $(V, \omega)$ be symplectic so that $\frac{1}{2\pi}[\omega]$ lifts to an integral cohomology class. Then we can find a $U(1)$ bundle $p : P \to V$ with $c_1(P) = \frac{1}{2\pi}[\omega]$. The $U(1)$ action on $P$ is generated by a vector field $X$. Consider a connection one form $A \in \Omega^1(P, i\mathbb{R})$; one has that $A(X) = i$ and $A$ is $S^1$-equivariant. By Chern–Weil theory, we may choose $A$ so that $F_A = \mathrm{d}A = -ip^*\omega$.

Then $P$ is contact with contact form $\alpha = -iA$. This form is $S^1$-invariant and $\alpha(X) = 1$. Hence $X$ is the Reeb vector field and its orbits are the fibres of $P$ with period $2\pi$.

$P$ is called the *prequantization bundle* associated to $V$. The name comes from the fact that the creation of $P$ is the crucial step in *geometric quantization*. Suppose $(V, \omega)$ is a symplectic manifold as above and $P$ is its prequantization bundle. Let $E \to V$ be its associated complex line bundle. Pick a *polarization* of $V$, that is, a choice of involutive Lagrangian sub-bundle $\mathcal{F}$ of $(T_{\mathbb{C}}V, \omega_{\mathbb{C}})$. Then the space $\mathscr{H}$ of $L^2$ sections of $E$ which are covariantly constant along $\mathcal{F}$ is

called the *quantum state space* of $V$. This is the outcome of geometric quantization and $\mathscr{H}$ should be thought of as the Hilbert space of wavefunctions in the quantized version of the classical system $(V, \omega)$. For more on this interesting theory, one can read the notes of Bates and Weinstein linked here.

*Legendrian Knots.* Let us say something about Legendrian knots in $(\mathbb{R}^3, \xi_{\text{std}})$. Because $\xi_{\text{std}}$ is non-integrable, any knot is forced to bend around in $\mathbb{R}^3$. We may wish to classify Legendrian knots up to *Legendrian isotopy*, that is an isotopy through Legendrian knots. Since our knots are forced to bend, this will be a finer classification than just the topological classification of knots.

To help visualize a Legendrian knot, we can project the knot onto the $(x, y)$ plane keeping track of crossings; this is called the *Lagrangian projection*. Suppose that $\gamma : [a, b] \to K$ is some loop inside the knot. If $\int_{[a,b]}(\mathrm{d}z - y\,\mathrm{d}x) = 0$ then,

$$z(\gamma(b)) - z(\gamma(a)) = \int_\gamma y\,\mathrm{d}x = -\text{area}(\gamma).$$

So, the signed area of the closed loop must be zero and the relative position of two ends of a crossing is determined by the area of loop bounded by knot between the crossings (in particular we don't need to draw which strand crosses the other, it is totally determined by the diagram). A Legendrian knot is thus the same as a knot projection to the $(x, y)$ plane with closed loops between crossings satisfying some area constraints, since we can figure out the height of the knot in the $z$ axis by integrating $y\,\mathrm{d}x$ along the knot. Note for example this implies an unknot cannot have a knot projection as just an ordinary loop, but it can appear as a figure eight with two regions of equal area.

Recall smoothly knots up to isotopy are determined by knot projections in $\mathbb{R}^2$ up to the three Reidemeister moves pictured in Figure 8. Legendrian knots up to Legendrian isotopy are determined by their knot projection up to the Reidemeister moves RII and RIII preserving the area constraints. If we tried to perform the Reidemeister move RI, at some point the knot would be tangent at a point to the $z$ direction, which is never in the span of $\xi_{\text{std}}$.

One can also consider the *front projection* of a Legendrian knot to the $(x, z)$ plane. This will determine the knot because along the knot $\mathrm{d}z - y\,\mathrm{d}x = 0$ which implies $y = \frac{\mathrm{d}z}{\mathrm{d}x}$. Crossings in this front projection must always be transverse, but there will usually be cusps where a strand enters and exits from a point with opposite tangency. In Figure 9, some basic examples of how these things look are drawn.

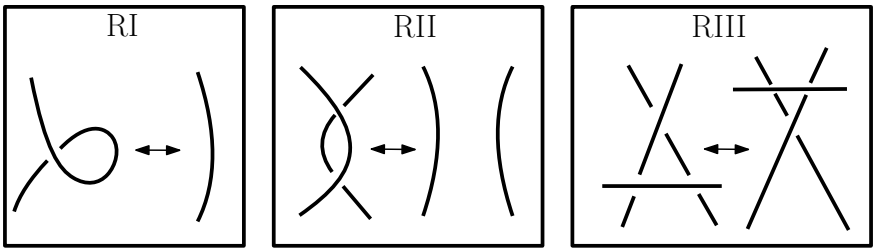



FIGURE 8. The Reidemeister moves

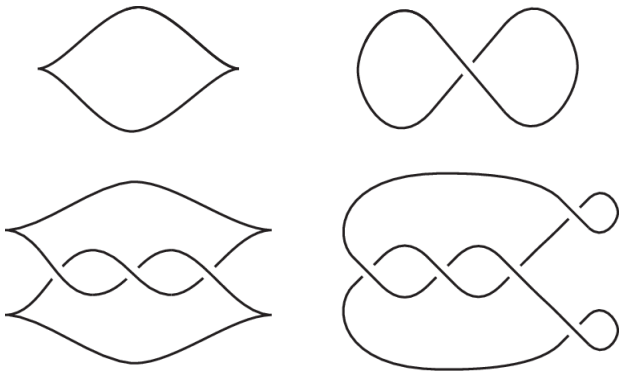


FIGURE 9. Examples of the Legendrian unknot (above) and trefoil knot (below) under front projection (left) and Lagrangian projection (right).

There are some "classical" invariants that help distinguish non-Legendrian isotopic Legendrian knots, namely the rotation number and the Thurston–Bennequin invariant. There are also more modern Floer theoretic invariants coming from the Chekanov–Elisashberg DGA and its Legendrian contact homology (see here).

**4.3. Structure Theorems.** We briefly discuss a few theorems on classification and local structure of contact manifolds. These are very similar to the Darboux and Moser-type results we had for symplectic manifolds.

Given a cooriented contact manifold $(Y, \alpha)$, a *contact isotopy* is a family of maps $\psi : Y \times I \to Y$ so that,

$$\psi_t^* \alpha = e^{f_t} \alpha \quad \text{and} \quad \psi_0 = \text{id}.$$

Contact isotopies are precisely the families of maps generated by time-dependent contact vector fields. Here, a *contact vector field* is a vector field $X$ so that $\mathcal{L}_X \alpha = g\alpha$ for a smooth function $g$.

We will omit the details which just follow from Moser's trick, but we obtain the following two fundamental results.

**Theorem 4.5: Contact Darboux Theorem**

Every co-oriented contact manifold $(Y^{2n-1}, \alpha)$ is locally contactomorphic to $(\mathbb{R}^{2n-1}, \lambda_{\text{std}})$.

**Theorem 4.6: Gray Stability Theorem**

Given a one parameter family of contact forms $\alpha_t$ on a closed manifold $M$, there is an isotopy $\psi_t$ and a family of positive functions $f_t$ so that,

$$\alpha_t = \psi_t^*(f_t \alpha_0),$$

with $\psi_0 = \text{id}$ and $f_0 = 1$.
In particular, smooth families of contact structures $\xi_t$ are isotopic and hence contactomorphic.

There is also an analogue of the Weinstein neighbourhood theorem.

**Theorem 4.7: Legendrian Neighbourhood Theorem**

Let $L$ be a Legendrian submanifold of a co-oriented contact manifold $(Y, \alpha)$. There is a neighbourhood of $L$ contactomorphic to a neighbourhood of the zero section in $J^1 L$.

**4.4. Contact and Symplectic Geometry.** We now flesh out the relationship between symplectic and contact geometry. Recall we defined a Liouville vector field in a symplectic manifold $(M, \omega)$ to be a vector field $X$ so that $\mathcal{L}_X \omega = \omega$. This means that $\omega$ scales exponentially along the flow $\psi_t$ of $X$: $\psi_t^* \omega = e^t \omega$. This implies there are no global Liouville vector fields on a compact manifold since the complete flow must exponentially scale the volume for all time. Another way to see that there are no global Liouville vector fields is that the condition $\mathcal{L}_X \omega = \omega$ equivalently means $\iota_X \omega$ is a primitive of $\omega$, which cannot exist in the compact case where $\omega$ is not null cohomologous. Note that Liouville vector fields always exist locally since we can find them in $\mathbb{R}^n$. We will be especially interested in the case where they exist in a neighbourhood of a compact hypersurface.

We now make the following definition, which initially seems unrelated.

**Definition 4.8**

A compact codimension one hypersurface $Y$ inside a symplectic manifold $(M, \omega)$ is said to be of *contact type* if it is equipped with a contact one form $\alpha$ so that $\mathrm{d}\alpha = \omega|_Y$. Equivalently, $\alpha(v) \neq 0$ for any non-zero $v \in T_y Y$.

PROPOSITION 4.9. *Let $(M, \omega)$ be a symplectic manifold and $Y$ a compact hypersurface. The following are equivalent.*

*(1) $Y$ is of contact type.*

*(2) There is a Liouville vector field defined in a neighbourhood of $Y$ and transverse to $Y$.*

PROOF. First suppose we have a transverse Liouville vector field $V$ defined in a neighbourhood of $Y$. Consider the one form $\iota_V \omega$. As we noted above, this is a primitive for $\omega|_Y$. We claim it is contact. Given $y \in Y$, pick any basis $e_1, \ldots, e_{2n-1}$ for $T_y Y$. Let $v = V_y \in T_y M$. Then by transversality, $v, e_1, \ldots, e_{2n-1}$ is a basis for $T_y M$. By non-degeneracy, we must have,

$$\omega^n(v, e_1, \ldots, e_{2n-1}) \neq 0.$$

So then,

$$\begin{aligned} 0 &\neq \omega^n(v, e_1, \ldots, e_{2n-1}) \\ &= (\iota_v \omega^n)(e_1, \ldots, e_{2n-1}) \\ &= n(\iota_v \omega) \wedge \omega^{n-1}(e_1, \ldots, e_{2n-1}) \\ &= n\alpha \wedge (\mathrm{d}\alpha)^{n-1}(e_1, \ldots, e_{2n-1}). \end{aligned}$$

Hence $\alpha \wedge (\mathrm{d}\alpha)^{n-1}$ is non-vanishing on $\det(TY)$ and so $\alpha$ is contact on $Y$.

Conversely, suppose $\alpha$ is a contact form on $Y$ which is a primitive for $\omega$; we have $\alpha \wedge (\mathrm{d}\alpha)^n \neq 0$. In particular $\alpha \neq 0$. By non-degeneracy, we may find a vector field $X_\alpha$ on $TM|_Y$ so that $\iota_{X_\alpha} \omega = \alpha$ on $TY$.

Define a map $\varphi : \mathbb{R} \times Y \to M$ by $\varphi(t, y) = \exp_y(tX_\alpha(y))$ with respect to some metric. Now,

$$\mathrm{d}\varphi_{(0,y)}(s, v) = sX_y + v \quad \text{for} \quad s \in \mathbb{R},\ v \in T_y Y.$$

Thus, $\varphi^* \omega|_{\{0\} \times Y} = \mathrm{d}s \wedge \alpha + \mathrm{d}\alpha$. We may now apply Moser istopy to find an embedding $\psi : (-\varepsilon, \varepsilon) \times Y \to M$ so that $\psi(0, y) = y$ and $\psi^* \omega = e^s(\mathrm{d}s \wedge \alpha + \mathrm{d}\alpha)$.

Now we take $V = \psi_* \frac{\partial}{\partial s}$ as our Liouville vector field. This is clearly transverse to $Y$. Then,

$$\iota_V \omega(v) = \omega(\psi_* \frac{\partial}{\partial s}, v) = \psi^* \omega(\frac{\partial}{\partial s}, \psi^* v) = e^s \alpha(v).$$

Hence $\mathcal{L}_V \omega = \psi_*(e^s(\mathrm{d}s \wedge \alpha + \mathrm{d}\alpha)) = \omega$, as desired. □

**Definition 4.10**

A compact codimension one hypersurface $Y$ inside a symplectic manifold $(M, \omega)$ is of *restricted contact type* if it is of contact type and its Liouville vector field is defined globally on $M$.

In the case where $M$ is a compact manifold with boundary. We say $M$ is symplectic with *contact type boundary* or sometimes *convex boundary* if $\partial M$ is of contact type. If $\partial M$ is of restricted contact type, we say $M$ is a *Liouville domain*.

Every $2n-1$ dimensional contact manifold $(Y, \alpha)$ admits a natural inclusion into a $n$ dimensional symplectic manifold as a hypersurface of restricted contact type. Consider the product manifold $M = \mathbb{R} \times Y$. Let $s$ be the $\mathbb{R}$ coordinate on $M$. Now define a form on $M$ by,

$$\omega = \mathrm{d}(e^s \alpha) = e^s(\mathrm{d}s \wedge \alpha + \mathrm{d}\alpha).$$

Then one has $\mathrm{d}\omega = 0$ and,

$$\omega^n = n! e^{ns}(\mathrm{d}s \wedge \alpha \wedge (\mathrm{d}\alpha)^{n-1}) \neq 0.$$

Hence, $(M, \omega)$ is symplectic. This space is called the *symplectization* of $(Y, \alpha)$. Note that $\frac{\partial}{\partial s}$ is a global Liouville vector field on $\mathbb{R} \times Y$. Every slice $\{s\} \times Y$ is a hypersurface of restricted contact type with the contact form $e^s \alpha$ with respect to the Liouville vector field $\frac{\partial}{\partial s}$.

Exercise 4.11. A submanifold $L$ of $Y$ is Legendrian if and only if $\mathbb{R} \times L$ is Lagrangian in the symplectic completion $\mathbb{R} \times Y$. ▲

Remark 4.12. This construction gives a bit of motivation to the proof of Proposition 4.9. We identified a tubular neighbourhood of $Y \subset M$ with a neighbourhood of $\{0\} \times Y$ in the symplectization of $Y$. Then we used Moser isotopy to give a symplectomorphism of a tubular neighbourhood of $Y$ in the two spaces preserving $Y$ and the pushforward of the Liouville vector field in the symplectization gives one in $M$. △

Example 4.13. Here are a few examples of contact type hypersurfaces, Liouville domains, and symplectizations.

(1) As discussed before, the Euler vector field $\mathcal{E}$ on $\mathbb{R}^{2n}$ is Liouville. It makes any domain in $\mathbb{R}^{2n}$ which is starshaped with respect to the origin (i.e. with boundary transverse to $\mathcal{E}$) into a Liouville domain. Any starshaped or in particular convex hypersurface in $\mathbb{R}^{2n}$ will be of contact type.

(2) Consider a cotangent bundle $T^*M$. In local standard coordinates $(q_1, \ldots, q_n, p_1, \ldots, p_n)$ consider the vector field,

$$X = \sum_{i=1}^{n} p_i \frac{\partial}{\partial p_i}.$$

This will be Liouville for $(T^*M, \omega_{\text{can}})$. The flow of $X$ is just exponential growth preserving the fibres. This makes the cotangent disk bundle $D^*M$ with respect to any metric on (compact) $M$ a Liouville domain with contact type bundle the sphere bundle $S^*M$.

(3) Take $(\mathbb{R}^{2n-1}, \lambda_{\text{std}})$. Its symplectization is $(\mathbb{R}^{2n}, \omega_{\text{std}})$.

(4) The symplectization of $(S^{2n-1}, \lambda_0)$ is symplectomorphic to $(\mathbb{R}^{2n} \setminus 0, \omega_{\text{std}})$.

(5) The symplectization of the cotangent sphere bundle $S^*M$ with respect to any metric and its standard contact form is the complement of the zero section in the cotangent bundle with its standard symplectic form.

(6) Here is a procedure to generate a non-contact hypersurface. Consider $M = N \times \mathbb{R}^2$ for $(N, \omega_0)$ closed symplectic, and put the symplectic form $\omega_0 \oplus \omega_{\text{std}}$ on $M$. Then suppose the level set $S = N \times S^1$ were contact-type with contact one form $\alpha$. Under the inclusion $j : S \hookrightarrow M$, one has $\mathrm{d}\alpha = j^*(\omega_0 \oplus \omega_{\text{std}})$. Letting $i : N \to S$ be the inclusion along the upper half of the circle, one has,

$$\mathrm{d}(i^*\alpha) = i^*j^*(\omega_0 \oplus \omega_{\text{std}}) = \omega_0.$$

But this implies the symplectic form on $N$ is exact, which cannot be true for $N$ closed. ▲

Suppose $(M, \omega)$ is a symplectic manifold with Hamiltonian $H$. Suppose $H$ has a regular level set $Y$ which is of contact type with contact primitive $\alpha$. Note that this is an open condition in the sense that the existence of a Liouville vector field implies nearby level sets of $H$ are also of contact type.

How does $M$ inform the geometry of $Y$? As we have seen above, $Y$ is a coisotropic submanifold of $M$: $TY^\omega \subset TY$. The symplectic subdistribution

of $TY$ will be the contact distribution $\xi$. The symplectic complement $TY^\omega$ is one dimensional, spanned by the Reeb vector field $R$. But note that the Hamiltonian vector field $X_H$ is tangent to $Y$ and symplectically orthogonal since $\omega(V, X_H) = V(H) = 0$ for any vector field tangent to $Y$ along which $H$ is constant. We conclude that $R$ and $X_H$ are colinear! Thus, up to reparameterizing, the Reeb dynamics are the same as the Hamiltonian dynamics. In fact, any other Hamiltonian function $H'$ with the same level set will have $\mathrm{d}H'|_Y = f\,\mathrm{d}H|_Y$ and hence $X_{H'} = fX_H$. And so we can modify the Hamiltonian so its flow is exactly our Reeb field (alternatively, we could modify the contact form to agree with the Hamiltonian flow). This up to reparameterization also agrees with the characteristic flow of $Y$.

An often studied object in symplectic dyamics are periodic orbits of some Hamiltonian, or equivalently closed characeristics. One might hope that Hamiltonian systems always admit periodic orbits on all of their regular level sets. This cannot be true for example, due to Zehnder, consider $T^4 = (\mathbb{R}/2\pi\mathbb{Z})^4$ with symplectic form,

$$\omega = \mathrm{d}x_1 \wedge \mathrm{d}x_2 + \varepsilon\,\mathrm{d}x_2 \wedge \mathrm{d}x_3 + \mathrm{d}x_3 \wedge \mathrm{d}x_4\,,$$

for $\varepsilon$ irrational. Then the Hamiltonian $H = \sin x_4$ has,

$$X_H = \cos x_4 \left( \frac{\partial}{\partial x_3} + \varepsilon \frac{\partial}{\partial x_2} \right),$$

which, since $\varepsilon$ is irrational, is never periodic on its regular level sets with $\cos(x_4) \neq 0$.

Weinstein proved that in $(\mathbb{R}^{2n}, \omega_{\mathrm{std}})$, all convex compact hypersurfaces have closed characteristics. Contact type hypersurfaces are a natural symplectic generalization of this family, leading to his famous conjecture.

Conjecture 4.14 (Weinstein Conjecture). *Every compact contact manifold $(Y, \alpha)$ with Reeb vector field $R$ admits a periodic orbit of $R$.*

This conjecture was proven in dimension three by Taubes using Seiberg–Witten theory and ideas closely related to Hutchings' theory of embedded contact homology. The conjecture in higher dimensions is still largely open outside of some known special cases.

In general, a much more precise statement is conjectured. Namely that the Reeb vector field on a $2n-1$ dimensional compact contact manifold either possesses exactly $n$ or infinitely many (embedded) periodic orbits. This is often called the "$n$ or infinity conjecture." This result has been proven in many cases in dimension 3, i.e. the "2 or infinity theorem" of Cristofaro-Gardiner–Hryniewicz–Hutchings–Liu. In higher dimensions, some limited cases are known.

## 5. Symplectic and Hamiltonian Actions

We now study symplectic manifolds equipped with some symmetry given by a Lie group action which is compatible in some sense. In the special case of Hamiltonian actions there will be a new construction called a momentum map which yields new examples of symplectic manifolds through symplectic reduction.

We also prove two important theorems related to Hamiltonian actions, namely the Duistermaat–Heckmann localization formula and the Atiyah–Guillemin–Sternberg convexity theorem.

### 5.1. Circle Actions.

To warm up, we begin with $S^1$-actions.

**Definition 5.1**

A *Hamiltonian $S^1$-action* on a symplectic manifold $(M, \omega)$ is a 1-parameter family of symplectomorphisms $\psi_t$ of $M$ which is 1-periodic and is the flow of a autonomous Hamiltonian vector field $H : M \to \mathbb{R}$. In this case, $H$ is called the *momentum map* of $\psi_t$ (or sometimes moment map).

More generally, a *symplectic $S^1$-action* is one given by the 1-periodic flow of a symplectic vector field.

Remark 5.2. We will see later that momentum maps correspond to measuring conserved quantities (like momenta) obtained from the symmetry of a Lie group action. The alternate name moment map is a misonomer (it is not the same as the concept of a moment in physics) based on a mistranslation from French.

The following is a basic version of the Marsden–Weinsten theorem for symplectic quotients which we treat later.

Proposition 5.3. *Suppose $S^1$ has a Hamiltonian action with moment map $H$ on $(M, \omega)$. If $S^1$ acts freely on a level set $H^{-1}(\lambda)$, then the quotient manifold,*

$$B_\lambda = H^{-1}(\lambda)/S^1,$$

*has a symplectic form $\tau_\lambda$ which is characterized by the fact it agrees with $\omega|_{H^{-1}(\lambda)}$ when pulled back by the $S^1$-action. We call $(B_\lambda, \tau_\lambda)$ the* Marsden–Weinstein quotient *or* symplectic quotient *at $\lambda \in \mathbb{R}$.*

PROOF. As we have discussed before, $Q = H^{-1}(\lambda)$ is coisotropic. Further, the $S^1$ orbits on $Q$ belong to the characteristic foliation $TQ^\omega$. Thus $\omega$ induces a non-degenerate 2-form $\widetilde{\omega}_p$ on $T_pQ/T_pQ^\omega$.

Because our action is Hamiltonian, $\widetilde{\omega}$ is preserved under the flow of $H$. Thus,

$$\widetilde{w}_p : \frac{T_pQ}{T_pQ^\omega} \times \frac{T_pQ}{T_pQ^\omega} \to \mathbb{R}$$

descends to a non-degenerate two-from $\tau_\lambda$ on $B_\lambda$. Since its pullback to $H^{-1}(\lambda)$ is closed, $\tau_\lambda$ must be closed and hence symplectic. □

EXAMPLE 5.4. Consider $S^2$ with the height function from its usual embedding into $\mathbb{R}^3$. This is the momentum map for the Hamiltonian $S^1$ action given by rotation around the sphere in each fixed $z$ plane. In this case, the image of the momentum map is a closed interval and away from the endpoints, there are regular values with associated symplectic quotient a point. △

EXAMPLE 5.5. Consider the action of $S^1$ on $(\mathbb{C}^n, \omega_{\text{std}})$ by multiplication by a complex unit $e^{2\pi i\theta}$. This has a momentum mapping $H(z) = \pi|z|^2$. The symplectic quotient at $\lambda = \pi$ gives a copy of $\mathbb{C}P^n$ equipped with the Fubini–Study form.

PROOF. Recall on a coordinate chart $\mathbb{C}^n \to \mathbb{C}P^n$ of points $(z_1, \ldots, z_n) \mapsto [1 : z_1 : \ldots : z_n]$, the Fubini–Study form was given by,

$$\omega_{FS} = \frac{i}{2}\partial\overline{\partial} \log\left(\sum_{i=1}^{n} |z_i|^2 + 1\right).$$

Let $\pi : \mathbb{C}^{n+1} \setminus \{0\} \to \mathbb{C}P^n$ be the Hopf projection map. On the above coordinate chart, this map takes the form,

$$(z_0, \ldots, z_n) \mapsto \left(\frac{z_1}{z_0}, \ldots, \frac{z_n}{z_0}\right) = [z_0 : \cdots : z_n].$$

In this chart we compute,

$$\begin{aligned}
\pi^*\omega_{FS} &= \frac{i}{2}\partial\overline{\partial}\log\left(\sum_{i=1}^{n}\frac{|z_i|^2}{|z_0|^2}+1\right)\\
&= \frac{i}{2}\partial\overline{\partial}\log\left(\frac{1}{|z_0|^2}\sum_{i=0}^{n}|z_i|^2\right)\\
&= \frac{i}{2}\partial\overline{\partial}\log\left(\sum_{i=0}^{n}|z_i|^2\right)-\frac{i}{2}\partial\overline{\partial}\log\big(|z_0|^2\big)\\
&= \frac{i}{2}\partial\overline{\partial}\log\left(\sum_{i=0}^{n}|z_i|^2\right)-\frac{i}{2}\partial\left(\frac{1}{\overline{z_0}}\right)\mathrm{d}\overline{z_0}\\
&= \frac{i}{2}\partial\overline{\partial}\log\left(\sum_{i=0}^{n}|z_i|^2\right).
\end{aligned}$$

Let us call this form defined on the last line $\rho \in \Omega^2(\mathbb{C}^{n+1}\setminus\{0\})$. Using the same argument on any other chart, we have $\pi^*\omega_{FS} = \rho$.

Expanding out we have,

$$\begin{aligned}
\rho &= \frac{i}{2}\partial\overline{\partial}\log\big(\|\vec{\mathbf{z}}\|^2\big)\\
&= \frac{i}{2}\partial\sum_{i=0}^{n}\frac{z_i}{\|\vec{\mathbf{z}}\|^2}\,\mathrm{d}\overline{z_i}\\
&= \frac{i}{2}\sum_{i=0}^{n}\left(\frac{1}{\|\vec{\mathbf{z}}\|^2}\,\mathrm{d}z_i\wedge\mathrm{d}\overline{z_i}-\sum_{j=0}^{n}\frac{z_i\overline{z_j}}{\|\vec{\mathbf{z}}\|^4}\,\mathrm{d}z_i\wedge\mathrm{d}\overline{z_j}\right).
\end{aligned}$$

Hence,

$$\rho|_{S^{2n+1}} = \frac{i}{2}\sum_{i=0}^{n}\left(\mathrm{d}z_i\wedge\mathrm{d}\overline{z_i}-\sum_{j=0}^{n}z_i\overline{z_j}\,\mathrm{d}z_i\wedge\mathrm{d}\overline{z_j}\right).$$

Expand out $z_i = z_i + iy_i$ and using skew symmetry,

$$\begin{aligned}
&= \sum_{i=0}^{n}\mathrm{d}x_i\wedge\mathrm{d}y_i\\
&= \omega_{\mathrm{std}}|_{S^{2n+1}}.
\end{aligned}$$

Hence on $H^{-1}(\pi)$, the pullback of $\omega_{FS}$ agrees with $\omega_{\mathrm{std}}$. By our above proposition, this implies the induced symplectic form on the reduction at $\lambda = \pi$ is $(\mathbb{C}P^n, \omega|_{FS})$. □

In the case of $\mathbb{C}P^1$, the Fubini–Study form is given by,

$$\omega_{FS} = \frac{\mathrm{d}x \wedge \mathrm{d}y}{(1+x^2+y^2)^2},$$

where $z = x + iy$ is the usual coordinate on the Riemann sphere. ▲

One can ask which symplectic $S^1$-actions are Hamiltonian.

**Theorem 5.6: Ono**

Let $(M, \omega)$ be a closed connected symplectic manifold so that wedging with $\omega^{n-1}$ induces an isomorphism between $H^1(M;\mathbb{R})$ and $H^{2n-1}(M;\mathbb{R})$. Then an $S^1$-action on $M$ is Hamiltonian if and only if it has fixed points.

Proof. If the action is Hamiltonian with momentum map $H$, then $H$ has fixed points by Morse theory which give rise to fixed points of the circle action.

Let $X$ be a symplectic vector field giving rise to an $S^1$-action given by its flow $\psi_t$.The orbits of the action all represent the same homology class $\alpha_X \in H_1(M;\mathbb{Z})$ since $M$ is connected.

If we assume this action is not Hamiltonian, then $\iota_X\omega$ cannot be exact. By our assumption, $\iota_X\omega^n$ must also be non-exact. Rescaling so $\int_M \omega^n = 1$, we now show $[\iota_X\omega^n] = PD(\alpha_X)$ which implies $\alpha_X$ cannot be null-homologous and so $\psi_t$ cannot have fixed points.

Consider an $S^1$-orbit $\gamma(t)$ and a $2n$-form $\sigma$ supported in a neighbourhood of $\gamma$ with integral one; by averaging, we can assume $\sigma$ is $S^1$-invariant. We must have $\int_M \sigma - \omega^n = 0$ and so $\sigma - \omega^n = \mathrm{d}\beta$ for $\beta \in \Omega^{2n-1}(M)$. Again by averaging, $\beta$ can be taken to be $S^1$-invariant. In particular $\mathcal{L}_X\beta = 0$. Hence,

$$\iota_X(\sigma - \omega^n) = \iota_X \mathrm{d}\beta = -\mathrm{d}\iota_X\beta.$$

Thus $\iota_X\sigma$ is cohomologous to $\iota_X\omega^n$. The fact $\iota_X\sigma$ is Poincaré dual to $\alpha_X$ is standard just by considering in local neighbourhoods. □

**5.2. Hamiltonian Actions and Momentum Maps.** We now generalize to consider actions of arbitrary Lie groups $G$ on a symplectic manifold $(M, \omega)$. We define a *symplectic $G$-action* to be a smooth $G$0action by symplectomorphisms in the sense that there is a smooth group homomorphism,

$$G \to \mathrm{Symp}(M, \omega) \quad \text{given by} \quad g \mapsto \psi_g.$$

Infinitesimally, this determines a Lie algebra homomorphism from the Lie algebra $\mathfrak{g}$ of $G$ to the space of symplectic vector fields on $M$, which we write as $\xi \mapsto X_\xi$.

Recall the adjoint action Ad of a Lie group on itself is given by conjugation. Taking the differential of this map, conjugating by a fixed group element $g$, gives an adjoint action also denoted $\mathrm{Ad}_g$ of $G$ on $\mathfrak{g}$. The adjoint action induces an infinitesimal action $\mathrm{ad}_\xi$ of $\mathfrak{g}$ on itself. This adjoint action is given by $\mathrm{ad}_\xi(\eta) = [\xi, \eta]$.

For a symplectic action $g \mapsto \psi_g$ giving an infinitesimal action $\xi \mapsto X_\xi$, one has,

$$X_{\mathrm{Ad}_{g^{-1}}\xi} = \psi_g^* X_\xi \quad \text{and} \quad X_{[\xi,\eta]} = -[X_\xi, X_\eta].$$

**Definition 5.7: Hamiltonian Action**

A symplectic $G$-action is *weakly Hamiltonian* if the corresponding vector fields $X_\xi$ are all Hamiltonian. This determines up to a constant a Hamiltonian $H_\xi$ so that $\iota_{X_\xi}\omega = -\mathrm{d}H_\xi$. Defining on a basis and extending linearly, we can pick these constants so that the map $\xi \mapsto H_\xi$ is linear. We call our action *Hamiltonian* if the map,

$$\mathfrak{g} \to C^\infty(M), \quad \xi \mapsto H_\xi,$$

can be chosen to be $G$-equivariant with respect to the adjoint action on $\mathfrak{g}$ and the action on $M$.

This map can equivalently be regarded as an element $\mu \in C^\infty(M) \otimes \mathfrak{g}^*$ or as a $G$-equivariant smooth map,

$$\mu : M \to \mathfrak{g}^* \quad \text{where} \quad H_\xi(p) = \langle \mu(p), \xi \rangle.$$

This is called the *momentum mapping* of the Hamiltonian $G$-action. The fact a $G$-action is Hamiltonian can be stated fully in terms of the momentum map which can be defined abstractly as a map $\mu : M \to \mathfrak{g}^*$ such that:

(1) $\mu$ determines the Hamiltonian functions of the action,

$$\mathrm{d}\langle \mu(p), \xi \rangle = -\iota_{X_\xi}\omega.$$

(2) $\mu$ satisfies an equivariance condition,

$$\mu \circ \psi_g = \mathrm{Ad}^*_{g^{-1}} \circ \mu,$$

for all $g \in G$. Here, $\mathrm{Ad}^*_g$ is the coadjoint action on $\mathfrak{g}^*$.

LEMMA 5.8. *Suppose $\psi_g$ is a symplectic $G$-action of a compact Lie group $G$ with infinitesimal action $X_\xi$. Suppose $\mathfrak{g} \to C^\infty(M)$ given by $\xi \mapsto H_\xi$ is a linear map with $\iota_{X_\xi}\omega = -\mathrm{d}H$. Consider the following statements.*

*(1) The map $\xi \mapsto H_\xi$ is $G$-equivariant:*

$$H_\xi \circ \psi_g = H_{\mathrm{Ad}_{g^{-1}}\xi}.$$

*(2) The map $\xi \mapsto H_\xi$ is a Lie algebra homomorphism with respect to the Poisson bracket:*

$$\{H_\xi, H_\eta\} = H_{[\xi,\eta]}.$$

*Then (1) implies (2) and if $G$ is connected they are equivalent.*

Proof. To show (1) implies (2), simply differentiate the identity,

$$H_\xi \circ \psi_{\exp(t\eta)} = H_{\exp(-t\eta)\xi\exp(t\eta)} \quad \text{at} \quad t = 0.$$

Now suppose $G$ is connected and assume (2) holds. As above, $X_{\mathrm{Ad}_{g^{-1}}\xi} = \psi_g^* X_\xi$ and so $H_\xi \circ \psi_g$ and $H_{g^{-1}\xi g}$ give the same Hamiltonian vector field. In particular, they differ by a constant. From (2),

$$\begin{aligned} H_{g^{-1}[\xi,\eta]g} &= \{H_{g^{-1}\xi g}, H_{g^{1-}\eta g}\} \\ &= \{H_\xi \circ \psi_g, H_\eta \circ \psi_g\} \\ &= \{H_\xi, H_\eta\} \circ \psi_g \\ &= H_{[\xi,\eta]} \circ \psi_g. \end{aligned}$$

Consider any $g_1 \in G$ and $\xi \in \mathfrak{g}$. By connectedness, there is a smooth path $g_t$ between $g_0 = e$ and $g_1$. This gives a path in the Lie algebra $\eta_t = \dot{g}_t g_t^{-1}$. One has,

$$\frac{\mathrm{d}}{\mathrm{d}t}\psi_g = X_\eta \circ \psi_g \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Ad}_{g^{-1}}\xi = \mathrm{Ad}_{g^{-1}}[\xi,\eta].$$

We then compute,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}(H_\xi \circ \psi_g - H_{g^{-1}\xi g}) &= \mathrm{d}(H_\xi \circ \psi_g)\, \psi_g^* X_\eta + H_{g^{-1}[\xi,\eta]g} \\ &= -\omega(X_{H_\xi \circ \psi_g}, \psi_g^* X_\eta) + H_{[\xi,\eta]} \circ \psi_g \\ &= -\omega(X_{H_\xi \circ \psi_g}, X_{H_\eta \circ \psi_g}) + \{H_\xi, H_\eta\} \circ \psi_g \\ &= -\{H_\xi \circ \psi_g, H_\eta \circ \psi_g\} + \{H_\xi, H_\eta\} \circ \psi_g \\ &= 0. \end{aligned}$$

As desired. □

Suppose we have a weakly Hamiltonian action. We can define $\tau \in \Lambda^2\mathfrak{g}^*$ by,

$$\tau(\xi,\eta) = \{H_\xi, H_\eta\} - H_{[\xi,\eta]}.$$

One has, applying Lie algebra cohomology theory,

LEMMA 5.9. *The weakly Hamiltonian action is Hamiltonian if and only if* $[\tau] = 0$ *in* $H^2(\mathfrak{g})$.

PROOF. See [McS1, p. 204] □

One can also show that when $H^1(\mathfrak{g}) = 0$, every symplectic action is weakly Hamiltonian and for a Hamiltonian action the choice of Hamiltonian is unique [daS, §26]. Here are some important special cases:

- If $G$ is a semi-simple group, so that $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$, then $H^1(\mathfrak{g}) = H^2(\mathfrak{g}) = 0$. Thus every symplectic action is uniquely Hamiltonian. See here p. 22.
- If $G$ is compact or $M$ is compact, then any weakly Hamiltonian action is Hamiltonian. See [Mein, pp. 62–63]
- If $(M, \omega = \mathrm{d}\theta)$ is exact, then any $G$-action preserving $\theta$ is Hamiltonian [Mein, p. 65]. The momentum map is given by,

$$\langle \mu, \xi \rangle = \iota_{X_\xi} \theta.$$

EXERCISE 5.10. Suppose $(M, \omega)$ has a Hamiltonian $G$-action with momentum map $\mu$. Suppose there is a Lie group homomorphism $\varphi : H \to G$. Show that precomposing with $\varphi$ gives a Hamiltonian $H$-action with momentum map given by post-composing with the induced map $\mathfrak{g}^* \to \mathfrak{h}^*$. ▲

**5.3. Examples of Momentum Maps.** We now discuss some important motivating examples of Hamiltonian $G$-actions and their momentum maps. As we noted above, the name momentum map comes from the underlying physics as conserved quantities like momenta that arise in Hamiltonian dynamics are expressed as the momentum maps of some $G$-action. The close relation between conserved momenta and group symmetries is expressed by Noether's famous theorem. In the physics literature, this is often stated as "every symmetry of a system has a corresponding conservation law."

**Theorem 5.11: Noether's Theorem**

Let $(M, \omega, H)$ be a Hamiltonian system and suppose there is a Hamiltonian $G$-action on $(M, \omega)$ for a Lie group $G$ with momentum map $\mu$. Then $H$ is $G$-equivariant if and only if $\mu$ is an integral of motion.

PROOF. Let $X_H$ be the Hamiltonian vector field. For any $\xi \in \mathfrak{g}$, let $G_\xi = \langle \mu, \xi \rangle$ be the Hamiltonian generator of the action. We have,

$$\begin{aligned}
\mathcal{L}_{X_H} G_\xi &= \iota_{X_H} \mathrm{d} G_\xi \\
&= -\iota_{X_H} \iota_{X_\xi} \omega \\
&= \iota_{X_\xi} \iota_{X_H} \omega \\
&= -\iota_{X_\xi} \mathrm{d} H \\
&= -\mathcal{L}_{X_\xi} H.
\end{aligned}$$

Hence $G_\xi$ is preserved under the flow of $X_H$ if and only if $H$ is invariant under the symplectic vector field $X_\xi$. Doing this for all $\xi$, $\mu$ is constant under the flow of $X_H$ if and only if $H$ is constant under the Hamiltonian action.

Maybe a more compact way to write the above computation using Hamilton's equation is,

$$\frac{\mathrm{d}}{\mathrm{d}t} G_\xi = \{H, G_\xi\} = -\{G_\xi, H\} = -\frac{\mathrm{d}}{\mathrm{d}t} \psi^*_{\exp(t\xi)} H.$$

Which yields the same result □

With this general philosophy in mind, let us now see a selection of examples.

**Linear Momentum:** Consider $(\mathbb{R}^{2n}, \omega_{\mathrm{std}})$. Let $G = \mathbb{R}^n$ act by translations,

$$\psi_a(x, y) = (x + a, y).$$

This action is Hamiltonian and is given for $\xi \in \mathbb{R}^n$ by,

$$H_\xi(x, y) = \langle \xi, y \rangle.$$

Using the usual inner product on $\mathbb{R}^n$ to identify $\mathbb{R}^n$ as an abelian Lie algebra with its dual, we obtain the momentum map,

$$\mu(x, y) = y.$$

Thus our momentum map, identifying $\mathbb{R}^{2n} = T^*\mathbb{R}^n$, returns the linear momentum of a point in phase space.

**Angular Momentum:** Consider $(\mathbb{R}^6, \omega_{\mathrm{std}})$. Let $G = SO(3)$ act diagonally,

$$\psi_A(x, y) = (Ax, Ay).$$

This action preserves the standard primitive of $\omega_{\mathrm{std}}$ and hence is Hamiltonian, generated by

$$H_\xi(x, y) = \langle y, \xi x \rangle \quad \text{for} \quad \xi \in \mathfrak{so}(3).$$

We can identify $\mathbb{R}^3$ with $\mathfrak{so}(3)$ so that $\xi \in \mathbb{R}^3$ corresponds to the skew-symmetric matrix,

$$\xi \mapsto A_\xi := \begin{pmatrix} 0 & -\xi_3 & \xi_2 \\ \xi_3 & 0 & -\xi_1 \\ -\xi_2 & \xi_1 & 0 \end{pmatrix}.$$

The action of this matrix $A_\xi$ on a vector $x \in \mathbb{R}^3$ is the same as taking the cross product $\xi \times x$.

There is a Killing form on $\mathfrak{so}(3)$ given by $\operatorname{tr}\left(A_\xi^{\intercal} A_\eta\right) = 2\langle \xi, \eta \rangle$. This allows us to identify $\mathfrak{so}(3)$ with its dual. The Hamiltonian generators then have the form,

$$H_\xi(x, y) = \langle y, \xi \times x \rangle = \langle x \times y, \xi \rangle,$$

and we can write the momentum map as,

$$\mu : \mathbb{R}^6 \to \mathfrak{so}(3), \quad \mu(x, y) = A_{x \times y}.$$

Recall that angular momentum in 3 dimensional classical physics is given by $\vec{\mathbf{L}} = \vec{\mathbf{r}} \times \vec{\mathbf{p}}$, the cross product of position and momentum. Hence we can identify our momentum map with the map $\mu : \mathbb{R}^6 \to \mathbb{R}^3$ that gives the angular momentum of a point in phase space.

If a Hamiltonian function $H$ depends only on $|x|, |y|$, and $\langle x, y \rangle$ (like for a central force), then by Noether's theorem, angular momentum is conserved. This is a source of many integrable systems.

More generally, on $(T^*\mathbb{R}^n, \omega_{\text{can}})$, we can let $G = SO(n)$ act diagonally and this will be Hamiltonian. The momentum map $\mu : T^*\mathbb{R}^n \to \mathfrak{o}(n)^*$ is given by,

$$\mu(q, p)_{ij} = \frac{1}{2}(q_i p_j - q_j p_i).$$

**Unitary Representation:** Consider $(\mathbb{C}^n, \omega_{\text{std}})$ with the standard action of $U(n)$. Given a skew-Hermitian matrix $\xi \in \mathfrak{u}(n)$, one has,

$$H_\xi(z) = -\frac{i}{2} z^* \xi z,$$

where $z^*$ denotes the conjugate transpose of the vector $z$. We can identify $\mathfrak{u}(n)$ and its dual with the Killing form $\langle \xi_1, \xi_2 \rangle = \operatorname{tr}(\xi_1^* \xi_2)$. The momentum map is then determined by $H_\xi(z) = \operatorname{tr}(\mu(z)^* \xi)$. So we obtain,

$$\mu(z) = \frac{i}{2} z z^*.$$

From the above exercise 5.10, we conclude that any unitary representation $G \to U(n)$ defines a Hamiltonian action of $G$ on $\mathbb{C}^n$. For example, we had an earlier $S^1$ action which arises in this way from the diagonal embedding $S^1 \to U(n)$. The momentum map of this should be the composition of $\mu$ with the projection $\mathfrak{u}(n)^* \to \mathfrak{u}(1)^*$. The resulting momentum map is $\frac{1}{2}\|z\|^2$, which agrees with what we wrote above.

More generally, we can consider the diagonal embedding $T^n \to U(n)$ and obtain a Hamiltonian action. The momentum map of this is,

$$\mu : \mathbb{C}^n \to \mathbb{R}^n, \quad \mu(z) = \frac{1}{2}(|z_1|^2, \ldots, |z_n|^2).$$

**Projective Representation:** Acting by $U(n+1)$ on $\mathbb{C}^{n+1}$ will induce an action on $(\mathbb{C}P^n, \omega_{FS})$. This will also be Hamiltonian. Its momentum mapping is given by,

$$\langle \mu[z_0 : \cdots : z_n], \xi \rangle = -\frac{\langle z, i\xi z \rangle}{2|z|^2}.$$

In this way, any projective representation of $U(n+1)$ also gives a Hamiltonian action on $\mathbb{C}P^n$.

**Coadjoint Orbits:** Recall we saw that for a Lie algebra $\mathfrak{g} = \mathrm{Lie}(G)$, its dual $\mathfrak{g}^*$ was a Poisson manifold whose symplectic leaves were the orbits of the coadjoint action of $G$. The symplectic form on such an orbit $\mathcal{O}$ is given by,

$$\omega_\eta(\mathrm{ad}(\xi)^*\eta, \mathrm{ad}(\xi')\eta) = \langle \eta, [\xi, \xi'] \rangle.$$

This is well defined since $T_\eta \mathcal{O} = \{\mathrm{ad}(\xi)^*\eta : \eta \in \mathfrak{g}\}$.

The group $G$ acts on $\mathcal{O}$ by the coadjoint action,

$$\psi_g(\eta) = \mathrm{Ad}(g^{-1})^*\eta.$$

Infinitesimally, this is given by the vector field $X_\xi(\eta) = -\mathrm{ad}(\xi)^*\eta$. This is Hamiltonian, and the associated Hamiltonian function is $H_\xi(\eta) = \langle \eta, \xi \rangle$. The momentum map $\mu : \mathcal{O} \to \mathfrak{g}^*$ is then just the inclusion.

**Connections on a Riemann Surface:** Here is a wonderful example in infinite dimensions originally due to Atiyah and Bott. Let $\Sigma$ be a Riemann surface and let $P \to \Sigma$ be a principal $G$-bundle for some compact Lie group $G$. We can consider connections on this principal bundle. They form an affine space $\mathcal{A}$ modelled on $\Omega^1(\Sigma, \mathfrak{g}_P)$. Here $\mathfrak{g}_P$ is the adjoint bundle obtained as an associated vector bundle $P \times_G \mathfrak{g}$ from the adjoint action of $G$ on $\mathfrak{g}$. Locally, a connection is given by a

Lie algebra valued one-form, although globally it may not have that description due to the fact $P$ is non-trivial.

There is an Ad-invariant inner product $\mathfrak{g}$ (the Killing form when $\mathfrak{g}$ is semi-simple) which will induce a metric on $\mathfrak{g}_P$. We can then define a two-form on $\mathcal{A}$ by,

$$\Omega(\alpha, \beta) = \int_\Sigma \langle \alpha \wedge \beta \rangle.$$

Here we take the inner product of the elements of $\mathfrak{g}_P$ and wedge the forms to get a real valued two-form that we can integrate. This two-form is non-degenerate and hence gives $\mathcal{A}$ a symplectic structure. Picking a metric on $\Sigma$, the Hodge star operator defines a complex structure $\alpha \mapsto \star\alpha$. Hence $\mathcal{A}$ is Kähler.

The space of connections is acted on by the *gauge group* $\mathcal{G} = \Gamma(\Sigma, G_P)$. Here, $G_P$ is the adjoint bundle $P \times_G G$, where $G$ acts on itself by the adjoint action. The group $\mathcal{G}$ acts on connection one-forms by,

$$g^* A = g^{-1}\,\mathrm{d}g + g^{-1} A g.$$

We claim this defines a Hamiltonian $\mathcal{G}$-action on $\mathcal{A}$.

The Lie algebra of $\mathcal{G}$ is the space of sections $\Gamma(\Sigma, \mathfrak{g}_p)$. Given such a section $\xi$, the infinitesimal action gives a vector field,

$$X_\xi(A) = -\mathrm{d}_A \xi = -\mathrm{d}\xi - [A, \xi].$$

This is the exterior covariant derivative induced by a connection one-form $A$. Note that $\Omega^2(\Sigma, \mathfrak{g}_P)$ and $\Omega^0(\Sigma, \mathfrak{g}_P)$ are formally dual via,

$$\langle \xi, F \rangle = \int_\Sigma \langle \xi \wedge F \rangle.$$

The corresponding Hamiltonian function is,

$$H_\xi(A) = \int_\Sigma \langle F_A \wedge \xi \rangle \quad \text{where} \quad F_A = \mathrm{d}_A^2 = \mathrm{d}A + \frac{1}{2}[A \wedge A].$$

$F_A \in \Omega^2(\Sigma, \mathfrak{g}_P)$ is the curvature of the connection $A$. To see this is the Hamiltonian note for $B \in \Omega^1(\Sigma, \mathfrak{g}_P)$,

$$\frac{\mathrm{d}}{\mathrm{d}t} F_{A+tB}\bigg|_{t=0} = \mathrm{d}B + [A \wedge B] = \mathrm{d}_A B.$$

Hence, integrating by parts,

$$\mathrm{d}H_\xi(B) = \int_\Sigma \langle \mathrm{d}_A B \wedge \xi \rangle = \int_\Sigma \langle B \wedge \mathrm{d}_A \xi \rangle = \omega(X_\xi(A), B).$$

Thus we can identify the momentum map with,

$$\mu : \mathcal{A} \to \Omega^2(\Sigma, \mathfrak{g}) \quad \text{given by} \quad \mu(A) = F_A.$$

So the momentum mapping is the curvature.

**5.4. Marsden–Weinstein–Meyer Theorem.** Suppose that $(M, \omega)$ is symplectic with a Hamiltonian $G$-action and moment map $\mu$. If $H$ is any $G$-invariant Hamiltonian, then Noether's theorem tells us that the dynamics of $H$ leave $\mu$ invariant. This suggests we should be able to simplify the problem by restricting to a level set of $\mu$ and modding out by the $G$-action under which both $H$ and $\mu$ are invariant; in other words, this restricted quotient should be symplectic. Indeed there is such a general procedure which we now construct.

We begin with a general lemma in differential geometry. We will say a submanifold $i : N \hookrightarrow (M, \omega)$ is *constant rank* if $\ker(i^*\omega)_x$ has dimension independent of $x \in N$.

PROPOSITION 5.12. *Let $N$ be a manifold with a closed 2-form $\sigma$ of constant rank. Then* $\ker \sigma$ *is integrable.*

PROOF. Consider vector field $X_1, X_2$ in the kernel of $\sigma$. Then,

$$\mathcal{L}_{X_i}\sigma = \iota_{X_i}\mathrm{d}\sigma + \mathrm{d}\iota_{X_i}\sigma = 0.$$

Thus,

$$\iota_{[X_1,X_2]}\sigma = \mathcal{L}_{X_1}\iota_{X_2}\sigma - \iota_{X_2}\mathcal{L}_{X_1}\sigma = 0.$$

By the Frobenius theorem, we are done. □

Thus, if $N$ is a constant rank submanifold, then $\ker \omega|_N$ will define a foliation, which we call the *null foliation* by *isotropic leaves*. Write $\pi : N \to B$ for the submersion whose fibres are the leaves and $B$ is the space of leaves. Note $\iota_X\omega|_N = \mathcal{L}_X\omega|_N = 0$ for all vertical vector fields $X$. Hence, we can find a 2-form on $B$ which pulls back under $\pi$ to $\omega$.

LEMMA 5.13. *Let $(M, \omega)$ be symplectic and let $x \in M$ have orbit $\mathcal{O}$ under a Hamiltonian $G$-action with moment map $\mu$. Then,*

$$\ker(\mathrm{d}_x\mu) = (T_x\mathcal{O})^\omega,$$
$$\mathrm{im}(\mathrm{d}_x\mu) = \mathrm{ann}(\mathfrak{g}_x).$$

*Here $\mathrm{d}_x\mu$ is the differential of the map and $\mathfrak{g}_x$ is the stabilizer in $\mathfrak{g}$ of the infinitesimal action on $M$.*

PROOF. We have for $\xi \in \mathfrak{g}$ and $V \in T_xM$,

$$\omega_x(X_\xi(x), V) = -\iota_V \mathrm{d}\langle \mu(x), \xi \rangle = -\langle \mathrm{d}_x\mu(V), \xi \rangle.$$

We conclude $\ker(\mathrm{d}_x\mu)$ is precisely the symplectic complement of the set of vectors $\{X_\xi(x) : \xi \in \mathfrak{g}\}$, which is $T_x\mathcal{O}$.

We also see from the computation that $\xi \in \mathfrak{g}_x$, meaning $X_\xi(x) = 0$, implies $\langle \mathrm{d}_x\mu(V), \xi \rangle = 0$ for all $V \in T_xM$. Thus $\mathrm{im}(\mathrm{d}_x\mu) \subset \mathrm{ann}(\xi)$. Checking dimensions,

$$\begin{aligned}
\dim(\mathrm{im}\, \mathrm{d}_x\mu) &= \dim M - \dim(\ker \mathrm{d}_x\mu) \\
&= \dim M - \dim(T_x\mathcal{O})^\omega \\
&= \dim T_x\mathcal{O} \\
&= \dim \mathcal{O}.
\end{aligned}$$

By orbit stabilizer,

$$\begin{aligned}
&= \dim G - \dim G_x \\
&= \dim \mathrm{ann}(\mathfrak{g}_x).
\end{aligned}$$

Hence the inclusion above is an equality. □

PROPOSITION 5.14. *An element $\eta \in \mathfrak{g}^*$ is a regular value of $\mu$ if and only if the stabilizer $G_x$ is discrete for all $x \in \mu^{-1}(\eta)$. In this case, $\mu^{-1}(\eta)$ will be a constant rank submanifold; the isotropic leaves of its null foliation are the orbits of $G_\eta$ in $\mu^{-1}(\eta)$ (here $G_\eta$ denotes the stabilizer of $\eta$ under the coadjoint action of $G$).*

PROOF. Suppose $\eta$ is a regular value. Then for $\mu(x) = \eta$, $\mathrm{d}_x\mu$ is surjective and by the above lemma, $\mathrm{ann}(\mathfrak{g}_x) = \mathfrak{g}^*$ i.e. $\mathfrak{g}_x = \{0\}$. Exponentiating, this implies $G_x$ is discrete. Conversely if $G_x$ is discrete for all $x \in \mu^{-1}(\eta)$, then each $\mathfrak{g}_x = \{0\}$ and so $\mathrm{d}_x\mu$ must be surjective and $\eta$ is a regular value.

Consider the inclusion $i : \mu^{-1}(\eta) \hookrightarrow M$. We know from general theory of submersions, $T_x\mu^{-1}(\eta) = \ker \mathrm{d}_x\mu$. We have for $\mu(x) = \eta$ with $x$ in the $G$-orbit $\mathcal{O}_G(x)$,

$$\begin{aligned}
\ker(i^*\omega)_x &= \ker(\mathrm{d}_x\mu) \cap \ker(\omega|_{\mu^{-1}(\eta)})_x \\
&= T_x\mu^{-1}(\eta) \cap (T_x\mu^{-1}(\eta))^\omega \\
&= T_x\mu^{-1}(\eta) \cap T_x\mathcal{O}_G(x)
\end{aligned}$$

The part of the tangent space of $x$ tangent to $x$'s orbit contained within the tangent space of $\mu^{-1}(\eta)$ will be precisely the part of the tangent space tangent to $x$'s orbit under the stabilizer of $\mu^{-1}(\eta)$. By equivariance of the moment map, this will be the stabilizer of $\eta$ under the coadjoint action of $G$.

$$= T_x\mathcal{O}_{G_\eta}(x).$$

Then, $\dim \mathcal{O}_{G_\eta}(x) = \dim G_\eta$ (since the stabilizer of $x$ is discrete). Hence $\mu^{-1}(\eta)$ is constant rank. The tangent space to a leaf is $T_x\mathcal{O}_{G_\eta}(x)$ and hence the leaf is just the orbit of $G_\eta$. □

**Theorem 5.15: Marsden–Weinstein–Meyer**

Let $(M, \omega)$ be a Hamiltonian $G$-space with moment map $\mu$. Suppose $\eta$ is a regular value of $\mu$ and $G$ acts freely and properly on $\mu^{-1}(\eta)$. Then $\mu^{-1}(\eta)$ is a submanifold of $M$ with isotropic leaves the orbits of $G_\eta$. The quotient,

$$M_\mu := \mu^{-1}(\eta)/G_\eta,$$

is a smooth manifold with a unique symplectic form $\omega_\eta$ satisfying $i^*\omega = \pi^*\omega_\eta$ for the inclusion and projection maps,

$$i : \mu^{-1}(\eta) \to M \quad \text{and} \quad \pi : \mu^{-1}(\eta) \to M_\mu.$$

In particular for the case $\mu = 0$, $G_0 = G$, $\mu^{-1}(0)$ is coistropic, its leaves are the $G$-orbits, and we obtain a smooth quotient manifold,

$$M/\!\!/G := \mu^{-1}(0)/G.$$

Proof. Considering the previous proposition, $M_\mu$ is just the leaf space since the leaves are the orbits of $G_\eta$. The free and proper assumption ensures $M_\mu$ is a smooth manifold. As we noted above, since vector fields tangent to the leaves are symplectic, we can then find a unique form $\omega_\eta$ so that $\pi^*\omega_\eta = \omega|_{\mu^{-1}(\eta)}$.

The fact $\mu^{-1}(0)$ is coisotropic follows from the computation,

$$(T_x\mu^{-1}(0))^\omega = T_x\mathcal{O}_G(x) \subset T_x\mu^{-1}(0),$$

since the $G$ action preserves $\mu^{-1}(0)$ by equivariance. □

**Definition 5.16**

This procedure is called *symplectic reduction*. The space $M/\!\!/G$ is called the *Marsden–Weinstein quotient*, the *symplectic quotient*, or the *symplectic reduction*.

REMARK 5.17. If one drops the assumption that $G$ acts freely and properly on $\mu^{-1}(\eta)$, then $M_\mu$ will be a symplectic orbifold. If we don't assume $\eta$ is a regular value, then the space may be something even more singular. Quotients of this form are studied in *geometric invariant theory* (GIT). $\triangle$

EXAMPLE 5.18. Let us look at a few examples of symplectic reduction.

(1) As a special case of unitary representations, consider $U(k)$ acting on the space of $n \times k$ matrices $\mathbb{C}^{n\times k}$ by right multiplication. Identifying $\mathfrak{u}(k)$ with its dual via the Killing form, the momentum map of this $U(k)$-action is,

$$\mu(U) = \frac{i}{2}U^*U.$$

Thus the preimage $\mu^{-1}(\frac{i}{2}I)$ is the space of unitary $k$-frames and the Marsden–Weinstein quotient $\mu^{-1}(iI/2)/U(k)$ is the complex Grassmannian $\mathbb{C}Gr(k,n)$.

(2) Consider the motion of a particle in $\mathbb{R}^2$ described in phase space $T^*\mathbb{R}^2$ with rotationally symmetric Hamiltonian,

$$H(q,p) = \frac{1}{2}\|p\|^2 + V(\|q\|).$$

This has an $S^1$ action by rotation with momentum map $\mu(q,p) = p_1q_2 - q_2p_1$. Working in polar coordinates $(r,\theta)$, the Hamiltonian becomes,

$$H(r,\theta,p_r,p_\theta) = \frac{1}{2}(p_r^2 + \frac{1}{r^2}p_\theta^2) + V(r).$$

The momentum map is $\mu = p_\theta$. The symplectic form is $\omega = \mathrm{d}r \wedge p_r + \mathrm{d}\theta \wedge p_\theta$ and non-zero points are a regular value of $\mu$, since $S^1$ acts freely there. Under the inclusion $i_\eta : \mu^{-1}(\eta) \to T^*\mathbb{R}^2$, we have $i_\eta^*\omega = \mathrm{d}r \wedge p_r$. Hence $M_\eta \cong T^*\mathbb{R}_{>0}$. On the symplectic reduction, the Hamiltonian descends to a function,

$$H_\eta(r,p_r) = \frac{1}{2}p_r^2 + V(r) + \frac{\eta^2}{2r^2}.$$

The second two terms in this Hamiltonian are called the effective potential, written $V_{\text{eff}}(r)$. If $E$ is the total energy of some orbit (i.e. the value of $H$ on the orbit), conservation of energy implies $\dot{r}^2 = 2(E - V_{\text{eff}}(r))$. This can be solved to obtain,

$$t - t_0 = \int_{r_0}^{r} \frac{\mathrm{d}r}{\sqrt{2(E - V_{\text{eff}}(r))}}.$$

Since $r^2\dot{\theta} = \mu$,

$$\theta - \theta_0 = \int_{r_0}^{r} \frac{\eta \, \mathrm{d}r}{r^2\sqrt{2(E - V_{\text{eff}}(r))}}.$$

When $V(r) = 1/r$, this can solved to show the motion $\theta(t)$ follows elliptic orbits or hyperbolic trajectories. This describes the orbital motion of a planet about a star due to gravity. This example suggests how symmetry and momentum mappings might simplify physical calculation.

(3) For the space of $G$-connections $\mathcal{A}$ with Hamiltonian action by the gauge group $\mathcal{G}$ and momentum map the curvature, the Marsden–Weinstein quotient is,

$$\mathcal{M}^{\text{flat}} = \mathcal{A}/\!\!/\mathcal{G} = \mathcal{A}^{\text{flat}}/\mathcal{G}.$$

This space consists of the flat connections modulo gauge transformations. It is call the *moduli space of flat connections*. This is a very important object in gauge theory, for example it shows up in instanton Floer homology. In general 0 is not a regular value of the curvature and this space may be singular. ▲

An important example we will return to shortly is that of toric manifolds.

**Definition 5.19**

A closed symplectic manifold $(M^{2n}, \omega)$ is called a *toric manifold* if it has an effective Hamiltonian $T^n$-action. Here effective means that every non-identity element of $T^n$ acts non-trivially on some point in $M$, or equivalently the intersection of the stabilizers of all points in $M$ is $\{e\}$.

The action being effective means that the kernel of the homomorphism $T^n \to \mathrm{Symp}(M)$ is trivial. Infinitesimally, this means that the for each element $\xi \in \mathbb{R}^n$, the vector field $X_\xi$ it generates should be non-zero (on an open dense set). In particular, there should be a point $p$ so that the infinitesimal action $\mathbb{R}^n \to T_pM$ is injective. Dually, this means there is a $p$ so that the differential of the momentum map $\mathrm{d}\mu_p : T_pM \to \mathbb{R}^n$ is surjective.

Example 5.20. Let $k < n$ and let $T^k$ act on $\mathbb{C}^n$ via some inclusion $T^n \subset T^k$. Suppose the momentum map $\mu : \mathbb{C}^n \to \mathbb{R}^k$ is proper and $\eta \in \mathbb{R}^k$ be a regular value so that $T^k$ acts freely on $\mu^{-1}(\eta)$. We claim $\mu^{-1}(\eta)/T^k$ is a toric manifold with respect to the natural $T^{n-k}$ action.

Consider the diagonal action of $S^1$ on $\mathbb{C}^{n+1}$. By the above example, this will induce an action of $T^n$ on the symplection reduction $(\mathbb{C}P^n, \omega_{FS})$ given by,

$$(e^{i\theta_1}, \cdots, e^{i\theta_n}) \cdot [z_0 : \cdots : z_n] = [z_0 : z_1 e^{i\theta_1} : \cdots : z_n e^{i\theta_n}].$$

Thus $\mathbb{C}P^n$ is a toric manifold. This Hamiltonian action has momentum map,

$$\mu([z_0 : \cdots : z_n]) = \frac{1}{2(|z_0|^2 + \cdots + |z_n|^2)}(|z_1|^2, \ldots, |z_n|^2).$$

▲

**5.5. Torus Actions and Convexity.** We now focus on Hamiltonian torus actions, which have many nice properties. Consider such a torus action $T^m \to \mathrm{Symp}(M^{2n}, \omega)$ given by $\theta \mapsto \psi_\theta$. Suppose the associated momentum map is $\mu : M \to \mathbb{R}^m$. We have a Hamiltonian function $H_\theta = \langle \theta, \mu \rangle$ generating the flow:

$$\frac{\mathrm{d}}{\mathrm{d}t}\psi_{t\theta} = X_{H_\theta} \circ \psi_{t\theta}.$$

**Theorem 5.21: Atiyah–Guillemin–Sternberg**

Let $(M^{2n}, \omega)$ be a closed connected symplectic manifold with a Hamiltonian $T^m$-action $\theta \mapsto \psi_\theta$ and momentum map $\mu : M \to \mathbb{R}^m$. The image of $\mu$ is a convex polytope. In particular, the fixed points of the action form disjoint connected symplectic submanifolds on which $\mu$ is constant and $\mu(M)$ is the convex hull of the image of the fixed points.

Example 5.22. We begin with some examples. We just discussed $\mathbb{C}P^n$ as a toric manifold given by the action,

$$(e^{i\theta_1}, \cdots, e^{i\theta_n}) \cdot [z_0 : \cdots : z_n] = [z_0 : z_1 e^{i\theta_1} : \cdots : z_n e^{i\theta_n}].$$

The momentum map of this is,

$$\mu([z_0 : \cdots : z_n]) = \frac{\pi}{\|z\|^2}(|z_1|^2, \ldots, |z_n|^2).$$

The image of $\mu$ is thus the simplex,

$$\Delta = \left\{(x_1, \ldots, x_n) \in \mathbb{R}^n : 0 \leq \sum_i x_i \leq \pi\right\}.$$

There are $n+1$ isolated fixed points $[0 : \cdots : 0 : 1 : 0 : \cdots : 0]$. Under $\mu$, these are sent to the vertices of $\Delta$, and $\Delta$ is clearly the convex hull of these points. ▲

*Morse–Bott Theory.* To prove this theorem, we need to know about Morse–Bott functions. We assume the reader is familiar with Morse theory; the theory is speedily explained in my algebraic topology notes. One can also consult [**Hutchings**] for an introduction to Morse and Morse–Bott theory.

Morse–Bott theory is a generalize of Morse theory meant to accommodate functions whose critical points come in smooth families (usually parameterized by some Lie group acting on the manifold and preserving the set of critical points).

**Definition 5.23**

We say a smooth function $f : M \to \mathbb{R}$ is *Morse–Bott* if its set of critical points is a union of smooth *critical submanifolds* so that at any critical point $x$ in a critical submanifold $S$, the Hessian of $f$ at $x$ decomposes as a direct sum of a definite matrix and $T_xS$ ( which will be the 0-eigenspace), i.e. the Hessian is invertible on the normal bundle fibre $N_xS$.

For a critical submanifold $S$ of a Morse–Bott function $f$, we assign an *index interval* $[n^-(S), n^+(S)]$, where $n^-(S)$ is the dimension of the negative eigenspace of the Hessian and $n^+(S) = n^-(S) + \dim(S)$. The dimension of the positive eigenspace will be $\dim M - n^+(S)$.

The standard basic results of Morse theory generalize easily to the Morse–Bott case. The important results are summarized below and the proofs are left as exercises to the reader who is familiar with the Morse case.

PROPOSITION 5.24 (Morse–Bott Lemma). *Let $f$ be a Morse–Bott function on $M$ and $z$ a critical point of $f$ belonging to a $k$ dimensional critical submanifold $Z$. Then there are coordinates $(x, y) = (x_1, \ldots, x_k, y_1, \ldots, y_{n-k})$ near $z$ so that:*

*(1) The critical submanifold $Z$ is locally cut out by the plane $y = 0$.*

*(2) There is a quadratic form $q(x, y)$, non-degenerate in terms of $y$, so that $f(x, y) = f(z) + q(z, y)$.*

For the following result, for $f : M \to \mathbb{R}$, let $M_a = f^{-1}((-\infty, a])$.

PROPOSITION 5.25. *Let $f$ be a Morse–Bott function on $M$.*

*(1) Suppose $f^{-1}(a, b)$ contains no critical points of $f$. There there are homotopy equivalence $f^{-1}(a) \cong f^{-1}(b)$ and $M_a \cong M_b$.*

(2) *Suppose $f^{-1}(a,b)$ contains one critical manifold $Z$. Then there is a homotopy equivalence,*

$$M_b \cong M_a \cup D\nu_-(Z),$$

*where $D\nu_-(Z)$ is the disc bundle of the negative normal bundle of $Z$ (i.e. the tangent directions along which the Hessian is negative definite).*

COROLLARY 5.26. *Suppose $f : M \to \mathbb{R}$ is Morse–Bott with no critical manifolds $Z$ with indices $n^-(Z) = 1$ or $n^+(Z) = n-1$. Then there is one connected critical manifold of index $n^- = 0$ and each non-empty level set of $f$ is connected.*

PROOF. Let $S$ denote the union of the local minima of $f$ (i.e. index $n^- = 0$ critical manifolds). Under our hypothesis, the complement of the stable manifold of $S$ is codimension at least two. Hence $W^s(S)$ is connected and so $S$ is connected.

To prove the level sets are connected, one can work inductively passing through critical points. When we pass a critical point corresponding to a critical manifold $S$, the homotopy type of the level set will change by attaching a sphere bundle over $S$. Under the index assumption, this is codimension at least two and can not disconnect the level set. □

*Proving Convexity.*

LEMMA 5.27. *There is a compatible almost complex structure $J$ on $(M,\omega)$ invariant under the torus action: $\psi_\theta^* J = J$.*

PROOF. We can pick any metric $g_0$ on $M$ and average over $\psi_\theta^* g_0$ to find an $T^m$-invariant metric. We can use the homotopy equivalence map we constructed from metrics to compatible almost complex structures to obtain $J$. □

LEMMA 5.28. *The fixed points set of a subgroup $G \subset T^m$ is a symplectic submanifold.*

PROOF. Pick a compatible triple $(\omega, g, J)$ which is $T^m$-invariant by the above lemma. Let $x$ be a fixed point of $G$. The differentials $(\mathrm{d}\psi_\theta)_x$ for $\theta \in G$ define a unitary $G$-action on $(T_xM, \omega_x, J_x)$. We can consider the exponential map with respect to $g$ which will satisfy the equivariance condition:

$$\exp_x(\mathrm{d}\psi_\theta(\xi)) = \psi_\theta(\exp_x(\xi)).$$

Thus, fixed points of $\psi_\theta$ near $x$ are those coming from exponentiating a fixed point of $\mathrm{d}\psi_\theta$:

$$T_x\mathrm{Fix}(G) = \bigcap_{\theta\in G} \ker(I - (\mathrm{d}\psi_\theta)_x).$$

Recall $(\mathrm{d}\psi_\theta)_x$ was unitary and so its 1-eigenspace is $J_x$-invariant. Hence $T_x\mathrm{Fix}(G)$ is a non-degenerate subspace with respect to $\omega_x$. □

We now show the generating Hamiltonians $H_\theta = \langle \mu, \theta\rangle$ are Morse–Bott with even index critical submanifolds.

PROPOSITION 5.29. *For each $\theta \in \mathbb{R}^n$, $H_\theta : M \to \mathbb{R}$ is a Morse–Bott function with even dimensional critical manifolds of even index. The set of critical points of $H_\theta$ is a symplectic submanifold.*

PROOF. Note that,

$$\mathrm{Crit}(H_\theta) = \bigcap_{\tau\in T_\theta} \mathrm{Fix}(\psi_\tau).$$

Here, $T_\theta$ is the subtorus given as the closure of $\{t\theta + k : t \in \mathbb{R}, k \in \mathbb{Z}^m\}/\mathbb{Z}^m$.

Without loss of generality, the components of $\theta$ are independent over $\mathbb{Q}$ so that $T_\theta = T^m$; otherwise we can just restrict to consider the action of $T_\theta$. Then $\mathrm{Crit}(H_\theta)$ is the intersection of the fixed points of $\psi_\tau$ for all $\tau \in T^m$. By the preceding lemma, this is thus symplectic.

To check $H_\theta$ is Morse–Bott, consider the Hessian of $H_\theta$ at $x$, which we denote $S_x$. Note the Hamiltonian vector field has differential $\mathrm{d}X_H(x) = -J_xS_x$ and so the differential of the action is given by $\mathrm{d}\psi_{t\theta} = \exp(-tJ_xS_x)$. Thus,

$$\ker S_x = \bigcap_{\tau\in T^m} \ker(I - \mathrm{d}\psi_\tau) = T_x\mathrm{Crit}(H_\theta).$$

Thus the Hessian is non-degenerate on the normla bundle of a critical submanifold. Since $\mathrm{d}\psi_\theta$ are unitary, $S_x$ must commute with $J_x$ and so the eigenspaces of $S_x$ are $J_x$-invariant and even dimensional. Hence the positive and negative eigenspace of $S_x$ are even and so the critical manifolds have even index. □

PROOF. (of Theorem 5.5) To be completed. In the mean time, it can be found in [**McS1**]. □

*Delzant Polytopes.* If we place slightly more restrictions on our torus action, we can give a kind of explicit classification of these actions in terms of the image of their momentum maps. In particular, we want to consider toric manifolds. As defined above, these are symplectic manifolds $(M^{2n}, \omega)$ with effective Hamiltonian $T^n$-actions.

**Definition 5.30: Delzant Polytope**

A convex polytope $\Delta \subset \mathbb{R}^n$ is *Delzant* if:

(a) Each edge is parallel to an integer vector.

(b) Each vertex has exactly $n$ incident edges, and their $n$ associated integer vectors, in lowest terms, are a basis for $\mathbb{Z}^n$.

**Theorem 5.31: Delzant**

If $(M, \omega)$ is toric with momentum map $\mu$, then the image of $\mu$ is a Delzant polytope $\Delta$. Moreover, this map determines a bijection between toric symplectic $2n$-manifolds up to $T^n$ equivariant symplectomorphism and Delzant polytopes in $\mathbb{R}^n$ up to translation.

We will not fully prove this but we will explain the *Delzant construction* which associates to a Delzant polytope a toric manifold.

Given a Delzant polytope $\Delta$, define $\widetilde{M} = \Delta \times (\mathbb{R}/2\pi\mathbb{Z})^n$. If $\mu_1, \dots, \mu_n$ are the integer edge vectors of $\Delta$ and $\theta_1, \dots, \theta_n$ are the standard coordinates on the torus, we can define a symplectic form on $\widetilde{M}$:

$$\widetilde{\omega} = \frac{1}{2\pi} \sum_{i=1}^{n} \mathrm{d}\mu_i \wedge \mathrm{d}\theta_i \,.$$

This has torus action,

$$(t_1, \dots, t_n) \cdot (\mu_1, \dots, \mu_n, \theta_1, \dots, \theta_n) = (\mu_1, \dots, \mu_n, \theta_1 + 2\pi t_1, \dots, \theta_n + 2\pi\theta_n).$$

This is close to what we want, but we need to take $\widetilde{M}$ from a manifold with corners, to a closed manifold. The way we do this is by forming a quotient $M = \widetilde{M}/\!\sim$, where over a codimension $i$ face of $\Delta$ we mod out by the action of the $T^i$ subgroup of $T^n$ generated by the integer normal vectors of the face. If $\Delta$ is Delzant, this manifold will be smooth and inherit a symplectic form from $\widetilde{\omega}$. The area of the Delzant polytope will be equal to the symplectic volume of the toric manifold.

We won't go into the details of how this quotient works but instead give a few examples.

EXAMPLE 5.32. Here are a few specific Delzant polytopes and the corresponding toric manifolds.

(1) If $\Delta = [a, b]$, then $\widetilde{M}$ is a cylinder of length $b - a$. The quotient identifies the boundary circles to points and we obtain $M = S^2$ with area $b-a$. The toric structure on $S^2$ is just the usual rotaitonal circle action with momentum map the height function.

(2) Consider $\Delta = [a, b] \times [c, d]$. Now we form $\widetilde{M}$ as a trivial torus bundle over $\Delta$. On the horizontal edges we collapse $(\mu_1, \mu_2, \theta_1, \theta_2) \sim (\mu_1, \mu_2, \theta_1, \theta_2 + t_2)$, on the vertical edge we collapse $(\mu_1, \mu_2, \theta_1, \theta_2) \sim (\mu_1, \mu_2, \theta_1 + t_1, \theta_2)$, and at the corner we collapse $(\mu_1, \mu_2, \theta_1, \theta_2) \sim (\mu_1, \mu_2, \theta_1 + t_1, \theta_2 + t_2)$. The resulting singular torus fibration $M$ is just $S^2 \times S^2$, the product of two copies of the last example.

(3) Consider the following triangular polytope $\Delta$ as shown on the left of Figure 10.

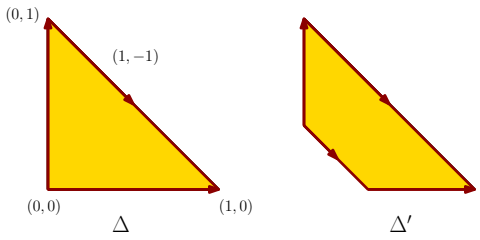



FIGURE 10. The Delzant polytopes of $\mathbb{C}P^2$ and $\mathbb{C}P^2 \# \overline{\mathbb{C}P^2}$

In this case, $\widetilde{M} = \Delta \times T^2$. Along the vertical edge, we want to identify $(0, \mu_2, \theta_1, \theta_2) \sim (0, \mu_2, \theta_1 + t, \theta_2)$. Along the horizontal edge, we want to identify $(0, \mu_2, \theta_1, \theta_2) \sim (\mu_1, 0, \theta_1, \theta_2 + t)$. And along the diagonal edge, we want to identify $(\mu_1, \mu_2, \theta_1, \theta_2) \sim (0, \mu_2, \theta_1 + t, \theta_2 + t)$.

If we just perform the horizontal and vertical identification, we can identify the resulting space $M'$ with the closed ball $B(1)$ of radius $\pi^{-1/2}$. To see this, set $\mu_i = \pi|z_i|^2$, then $B(1) = \{z \in \mathbb{C}^2 | \mu_1 + \mu_2 \leq 1\}$ and the vertical edge of the polytope is identified with $z_1 = 0$ and the horizontal edge with $z_2 = 0$. In this way, over each diagonal line tangent to the hypotenuse the space $M'$ is a $T^2$ fibration which degenerates to a circle at the endpoints and this defines a copy of $S^3$, i.e. a level set of the radial function in $B(1)$. The point $(0, 0)$ is the centre of the sphere and the diagonal edge is the boundary.

Now the final quotient of the diagonal edge will perform the identification $(z_1, z_2) \sim (e^{i\theta} z_1, e^{i\theta} z_2)$ on the boundary where $\pi(|z_1|^2 + |z_2|^2) = 1$.

This has the effect of taking $\mathbb{C}^2$ and quotienting by the complex Hopf fibration at the sphere at infinity to make it a copy of $\mathbb{C}P^1$. Thus the overall space will be $\mathbb{C}P^2$.

(4) Consider the polytope $\Delta'$ formed from the previous example with a bottom corner cut out as in the right of Figure 10. What we have done is remove a ball around the origin and quotient the boundary sphere by the same Hopf map. The result is topologically the connected sum $\mathbb{C}P^2\#\overline{\mathbb{C}P^2}$. This is the process of "blowing up" and cutting the corner off any Delzant polytope will have that same effect. We will discuss blowups in more detail in the next section.

**5.6. Duistermaat–Heckman Localization.** This result is a kind of physics-motivated result. It says that some kind of stationary phase approximation is in fact an exact equality. This is a special case of the Atiyah–Bott fixed point localization formula for $T^n$-equivariant cohomology and we will explain the more general result in our proof. The Duistermaat–Heckman location formula has applications to computing the cohomology rings of symplectic quotients.

Let $(M, \omega)$ be a symplectic with a circle action generated by a Morse function $H : M \to \mathbb{R}$. At a fixed point $p$ of this action, we obtain a tangent action of $S^1$ on $T_pM$. This will be conjugate to a product of circle action $z_j \mapsto e^{-2\pi i k_j t} z_j$ on copies of $\mathbb{C}$. The integers $k_j$ are called the *weights of the action* at $p$ and we write $e(p)$ for the product of the weights at $p$. This will be related by some constant to the determinant of the Hessian of $H$ at $p$.

**Theorem 5.33: Duistermaat–Heckman Localization Formula**

Let $(M, \omega)$ be a symplectic with a circle action generated by a Morse function $H : M \to \mathbb{R}$. For any $\hbar \in \mathbb{C}$,

$$\int_M e^{-\hbar H} \frac{\omega^n}{n!} = \sum_{p \in M^{S^1}} \frac{e^{-\hbar H(p)}}{\hbar^n e(p)}.$$

Here $M^{S^1}$ denotes the (finite) fixed point set of the $S^1$ action.

Remark 5.34. There is a simple generalization of this to a torus action and a general (by theorem 5.5 Morse–Bott) momentum map $\mu$. In this case, the formula becomes for $X \in \mathfrak{t} = \mathbb{C}^n$,

$$\int_M e^{i\langle \mu, X\rangle} \frac{(i\omega)^n}{n!} = \sum_{F \in \mathcal{F}} e^{i\langle \mu(F), X\rangle} \int_F \frac{e^{i\omega}}{e_F(X)}.$$

Here, $\mathcal{F}$ denotes the set of critical submanifolds of the Morse–Bott function $\mu$ and $e_F$ is the $T^n$-equivariant Euler class of the normal bundle $NF$. We will explain the meaning of $e_F$ below. Explicitly, $NF$ will split $T$-equivariantly as a direct sum of complex line bundle $\nu_j$ on which $T$ acts with weight $\beta_j \in \mathfrak{t}^*$ and then,

$$e_F(X) = \prod_j [c_1(\nu_j) + \beta_j(X)].$$

This more general formula follows from exactly the same proof as the basic case using the Atiyah–Bott formula applied to $\omega + \langle \mu, X \rangle$. △

REMARK 5.35. For more on this theorem and some commentary on its interpretation in terms of stationary phases see the book by Jeffrey et al.[**?**].

*Equivariant Cohomology.* To prove the Duistermaat–Heckman formula, we need to understand equivariant cohomology. We now give a topological perspective on the Borel model of equivariant cohomology. One can also define equivariant de Rham theory which gives another model of equivariant cohomology (see [**BGV**, Ch. 7]).

Let $X$ be a topological space with a continuous action of $G$ a topological group. One defines the *homotopy quotient* as the fibre product,

$$X_G = X \times_G EG.$$

**Definition 5.36: Equivariant Cohomology**

The (Borel) *equivariant cohomology* of $X$ is the cohomology of the homotopy quotient,

$$H^*_G(X) := H^*(X_G).$$

EXAMPLE 5.37. If the $G$-action on $X$ is free, then $X_G = X/G$ and the equivariant cohomology of $X$ is just the cohomology of $X/G$. One can think of equivariant cohomology as a replacement for the cohomology of the quotient, where we replace the quotient with a better behaved space when the action is not free.

If the $G$-action on $X$ is trivial, then $X_G = X \times BG$ and $H^*_G(X) = H^*(X \times BG)$. In particular, the $G$-equivariant cohomology of a point is $H^*(BG)$. ▲

Suppose $f : X \to Y$ is a $G$-equivariant continuous map between $G$ spaces. This induces a map $X_G \to Y_G$ and hence a pullback map on equivariant cohomology $f^* : H^*_G(X) \to H^*_G(X)$.

For any $G$-space $X$, there is a $G$-equivariant map $X \to \mathrm{pt}$. This induces a map $H^*_G(\mathrm{pt}) \to H^*_G(X)$. This is compatible with pullbacks. Hence, $G$-equivariant cohomology naturally has the structure of a module over $H^*(BG)$.

Given a subgroup $H \subset G$, a $G$-space is naturally restricts to an $H$-space structure and there is a a fibration $G/H \hookrightarrow X_H \to X_G$.

One also has a fibration $X \hookrightarrow X_G \to BG$. Hence the equivariant cohomology can be computed from the cohomology of $X$ and of $BG$ using a Serre spectral sequence.

**Definition 5.38: Equivariant Characteristic Classes**

Suppose $\xi : E \to X$ is a complex vector bundle which is $G$-equivariant. This gives rise to a vector bundle $\xi_G : E_G \to X_G$. The *equivariant Chern classes* of $\xi$ are given by,

$$c^G_k(\xi) = c_k(\xi_G) \in H^{2k}_G(X).$$

Analogous definitions hold for other equivariant characteristic classes.

Example 5.39. Consider the trivial bundle $\mathbb{C}$ over a point. Let $G = S^1$ act by $e^{2\pi i t\ell}$ for some $\ell \in \mathbb{Z}$. The $G$-equivariantized version of this bundle will be $L^{\otimes \ell} \to BS^1$, where $L$ is the Hopf bundle on $\mathbb{C}P^\infty$. One has $H^*(\mathbb{C}P^\infty) = \mathbb{Z}[x]$ where $x$ in degree two is defined by $\langle x, [\mathbb{C}P^1]\rangle = 1$. One can compute that $c_1^{S^1}$ of our bundle is $-\ell x$. ▲

Now we explain the Atiyah–Bott localization formula for torus actions, which will imply the Duistermaat–Heckman formula. From now on, we suppose $G$ is a connected Lie group acting smoothly on a compact oriented manifold.

Consider the projection $\pi : X_G \to BG$. There is a $G$-equivariant pushforward map $\pi_! : H^*_G(X) \to H^*(BG)$ given by,

$$\pi_!(\alpha) = \langle [X], \alpha \rangle \in H^*_G(\mathrm{pt}).$$

This generalizes the wrong-way pushforward map on ordinary cohomology.

**Theorem 5.40: Atiyah–Bott Fixed Point Localization Formula**

Let $M$ be a smooth compact oriented manifold with a torus-action. Let $M^T$ denote the fixed point set of the $T$-action, which will be a union of submanifolds of varying dimension, and $i : M^T \hookrightarrow M$ be the inclusion.

Given $\alpha \in H_T^*(M)$ one has,

$$\int_M \alpha = \int_{X^T} \frac{i^*\alpha}{e_T(\nu_i)},$$

where $\nu_i$ is the normal bundle of the embedding $i$ and $e_T$ is the $T$-equivariant Euler class.

PROOF. (Sketch) We present Atiyah and Bott's algebro-topological argument. There is a more geometric approach due to Berline–Vergne [**BGV**, Ch. 7] and a physics-inspired argument of Witten and Bismut involving some asymptotic parameter.

Consider the following commutative diagram of spaces and the induced maps on cohomology.

$$\begin{array}{ccc} X & \xhookleftarrow{i} & X^T \\ {\scriptstyle\pi}\downarrow & \swarrow{\scriptstyle\widetilde{\pi}} & \\ \mathrm{pt} & & \end{array} \qquad \begin{array}{ccc} H_T^*(X) & \xleftarrow{i_!} & H_T^*(X^T) \\ {\scriptstyle\pi_!}\downarrow & \swarrow{\scriptstyle\widetilde{\pi}_!} & \\ H^*(BT) & & \end{array}$$

Consider some class $\alpha \in H_T^*(X)$, as in the statement of the theorem. We can tensor with $\mathbb{Q}(\lambda)$, where $\lambda$ represents the generators of $H^*(BT)$. A localization result of Borel and Atiyah–Bott implies that after tensoring, $i_!$ and $i^*$ are isomorphisms. Hence we may find $\beta \in H_T^*(X^T)$ so that $\pi_!\alpha = \widetilde{\pi}_!\beta$. Then,

$$i^*\alpha = i^*i_!\beta = e_T(\nu)\beta.$$

The second equality follows from a Gysin sequence. Since $i^*$ is an isomorphism, after dividing through by the Euler class, we conclude. □

*Proving Duistermaat–Heckman.* The symplectic localization formula is now a simple consequence of the Atiyah–Bott formula we have just proved. We first need to pass from equivariant cohomology classes to differential forms.

Cartan defines a de Rham model for equivariant cohomology as follows. Define,

$$\Omega_G^*(M) = (\Omega^*(M) \otimes S(\mathfrak{g}^*))^G.$$

Here $S(\mathfrak{g}^*)$ is the space of real valued polynomials on the Lie algebra $\mathfrak{g}$ which inherits a $G$ action via the coadjoint action on $\mathfrak{g}^*$. This is graded by the grading on $\Omega^*(M)$ plus twice the usual grading on polynomials. One can define a differential $D : \Omega^*(M) \to \Omega^*(M)$. Given an element $f \in \Omega^*(M)$ and an element $X$ in the Lie algebra,

$$(Df)(X) = \mathrm{d}f(X) - \iota_{X^\sharp} f(X),$$

where $X^\sharp$ is the vector field generated by the action of $X$.

It turns out $D^2 = 0$ and the resulting cohomology of this complex recomputes equivariant cohomology (with $\mathbb{R}$ coefficients). All the machinery of equivariant cohomology and the localization formula will hold with differential forms. Now the proof of Duistermaat–Heckman is simple.

PROOF. (of Theorem 5.33) Just apply Theorem 5.40 to the form $e^{\omega - \hbar H}$. Note that $D(\omega - \hbar H) = -u(\iota_{X_H}\omega + \mathrm{d}H) = 0$, so we obtain an equivariant cohomology class. The fixed points are isolated and so the Atiyah–Bott formula yields,

$$\int_M e^{-\hbar H}\frac{\omega^n}{n!} = \int_M e^{\omega - \hbar H} = \sum_{p \in M^{S^1}} \frac{e^{-\hbar H(p)}}{e_T(p)}.$$

To go from $e_T$ to $e$, defined above as a product of weights, we just need to multiply by $\hbar^n$. □

## 6. Symplectic Constructions

We now describe how to construct many more examples of symplectic manifolds. One key idea will be a symplectic analogue of the complex blowup construction. From this, we will be able to describe the Gompf fibre connected sum. We will also discuss some important classes of open symplectic manifolds and Liouville domains related to complex geometry (Weinstein and Stein domains).

**6.1. Complex Blowup.** Let us recall the basic blowup construction in complex and algebraic geometry.

**Definition 6.1**

The *blowup* of $\mathbb{C}^n$ at the origin is the subspace $\widetilde{\mathbb{C}}^n \subset \mathbb{C}P^{n-1} \times \mathbb{C}^n$ consisting of all pairs $(\ell, z)$ where $z$ is a point in the line $\ell$. This is naturally a complex manifold of dimension $n$.

The projection $\pi : \widetilde{\mathbb{C}}^n \to \mathbb{C}^n$ is called the *blowdown* map. Over $\mathbb{C}^n \setminus \{0\}$, it is a biholomorphism. In contrast, $Z = \pi^{-1}(0)$ is a copy of $\mathbb{C}P^{n-1}$ called the *exceptional divisor*.

One can think of the space as built by replacing the origin by the set of all lines through the origin, i.e. a copy of $\mathbb{C}P^{n-1}$ (this is the "blowup"). The projection $\widetilde{\mathbb{C}}^n \to \mathbb{C}P^{n-1}$ exhibits the blowup as the total space of the tuatological line

bundle on $\mathbb{C}P^{n-1}$. From this perspective, the exceptional divisor $Z$ is the zero section of the bundle.

Now suppose $M$ is any complex $n$-manifold and $p \in M$. Let $Z$ be the complex projectivization of $T_pM$, i.e. the set of lines in $T_pM$. We can then define the *blow up of $M$ at $p$* to be the manifold $\widetilde{M} = (M \setminus \{p\}) \cup Z$, where we attach the boundary of $M \setminus \{p\}$, a copy of $S^{2n-1}$, to $Z$, a copy of $\mathbb{C}P^{n-1}$ via the generalized Hopf fibration. More precisely, we biholomorphically identify a neighbourhood of $p$ with a neighbourhood of the origin in $\mathbb{C}^n$ and apply the above blowup construction. Again we have a blowdown map $\pi : \widetilde{M} \to M$, which is a biholomorphism away from $p$ and $Z = \pi^{-1}(p)$ is called the *exceptional divisor*.

Suppose $\widetilde{M}^{2n}$ is complex with a holomorphically embedded submanifold $Z$, a copy of $\mathbb{C}P^{n-1}$ whose normal bundle is isomorphic to the tautological bundle. Then we can collapse $Z$ to a point and obtain a manifold $M$, the *blowdown* of $\widetilde{M}$. This construction is clearly inverse to the blowup construction.

The blown up space $\widetilde{M}$ is naturally a complex manifold. Note, topologically it was formed by a removing a $2n$-ball from $M$ and attaching along the boundary a copy of $\mathbb{C}P^n$ with a $2n$-ball removed (since $\mathbb{C}P^n \setminus B^{2n}$ is diffeomorphic to $\mathbb{C}P^{n-1}$). Keeping track of orientations, we conclude that $\widetilde{M}$ is oriented-diffeomorphic to $M \# \overline{\mathbb{C}P^n}$.

LEMMA 6.2. *Every non-constant connected holomorphic curve $u : \Sigma \to M$ has a unique holomorphic lift $\widetilde{u} : \Sigma \to \widetilde{M}$ so that $\pi \circ \widetilde{u} = u$.*

PROOF. Without loss of generality, $u$ is a holomorphic curve $U \to \mathbb{C}$, for $U$ a neighbourhood of 0 in $\mathbb{C}$, with $u(0) = 0$ and never otherwise equal to 0. We simply set,

$$\widetilde{u}(0) = ([u_1^{(k)}(0) : \cdots : u_n^{(k)}(0)], 0) \in \widetilde{\mathbb{C}}^n,$$

where $k$ is the smallest integer so that the $k$th derivative $u^{(k)}$ does not vanish at the origin. $\widetilde{u}$ is determined every else by $u$. It can be checked in coordinate charts that $\widetilde{u}$ is smooth at zero. This extension is unique just by imposing continuity. □

EXAMPLE 6.3. Here is one example we will need later. We say a complex surface $X$ is *elliptic* if there is a holomorphic map $f : X \to \mathbb{C}P^1$ with generic fibre an elliptic curve, i.e. a 2-torus with a linear complex structure. Here is an important elliptic surface.

Consider the space $E(1)$ formed by blowing up $\mathbb{C}P^2$ at nine points. To be concrete, we can blow up at the nine intersection points of transverse smooth cubic curves $\Gamma_1, \Gamma_2$. Let $\Gamma_i$ be the zero set of homogeneous cubric polynomials $P_i$ For each $w = [a : b] \in \mathbb{C}P^1$, we can find a unique cubic curve $\Gamma_w$ which is the zero set of $aP_1 + bP_2$. Each pair of curves in the family $\Gamma_w$ intersects transversely in the same nine points.

Thus we can lift to curves $\widetilde{\Gamma}_w$ in $E(1)$ which must all be disjoint. There is exactly one cubic $\Gamma_w$ intersecting every other point of $\mathbb{C}P^2$ and so each point of $E(1)$ lie in exactly one of the curves $\Gamma_w$. This defines a holomorphic map $V \to \mathbb{C}P^1$ sending a point $z$ to $w$ with $z \in \Gamma_w$. Finitely many of the fibres of this map will be singular, but generically the fibre is a torus since non-degenerate cubics in $\mathbb{C}P^2$ have genus one (genus-degree formula). Thus $E(1)$ is an elliptic surface, often called a *rational elliptic surface*. ▲

To define symplectic blowups, we will need to say a bit about Kähler forms on the blowup. Let us return to $\widetilde{\mathbb{C}}^n$ identified with the total space of the tualogical bundle and let $p : \widetilde{\mathbb{C}}^n \to \mathbb{C}P^{n-1}$ be the projection to the base. We can define a family of Kähler forms on $\widetilde{\mathbb{C}}^n$ by,

$$\widetilde{\omega}_\lambda = \pi^* \omega_{\text{std}} + \lambda^2 p^* \omega_{FS}.$$

We will need the following notation,

$$B(\varepsilon) := \left\{ z \in \mathbb{C}^n : |z| \leq \varepsilon \right\}$$

$$L(\varepsilon) := \pi^{-1}(B(\varepsilon)) = \left\{ ([\ell], z) \in \widetilde{\mathbb{C}}^n : |z| \leq \varepsilon \right\}. \tag{6.1}$$

Note $L(\varepsilon)$ is a closed neighbourhood of the exceptional divisor $Z$. It turns out the complement $L(\varepsilon) \setminus Z$ equipped with $\widetilde{\omega}_\lambda$ is symplectomorphic to the spherical shell $B(\sqrt{\lambda^2 + \varepsilon^2}) \setminus B(\varepsilon)$ inside $(\mathbb{C}^n, \omega_{\text{std}})$. The proof is in [**MS**, Lemma 7.1.11] and it is just a tedious computation.

**6.2. Symplectic Blowups and Blowdowns.** The idea of a symplectic blowup is to identify a neighbourhood of a symplectic manifold with a complex manifold and then perform a complex blowup and equip it with an appropriate symplectic form.

Another perspective on this construction is that the blowup removes a symplectically embedded ball and collapses the bounding sphere to projective space via the Hopf fibration. Inversely, one can blowdown by removing an exceptional divisor, a symplectically embedded $\mathbb{C}P^{n-1}$, and gluing in a ball. These two operations have a kind of trade off: blowups increase the complexity of the topology of the space while reducing its volume, and blowdowns simplify the topology while increasing the volume.

**Definition 6.4**

Let $(M, \omega)$ be symplectic with a symplectically embedded $Z \cong \mathbb{C}P^{n-1}$ whose normal bundle is isomorphic to the tautological bundle. By the symplectic neighbourhood theorem, there is a neighbourhood $\mathcal{N}(Z)$ of $Z$ symplectomorphic to $(L(\varepsilon), \widetilde{\omega}_\lambda)$ as defined in (6.1). The *symplectic blowdown* of $M$ is given by removing $\mathcal{N}(Z)$ and replacing it with $B(\sqrt{\lambda^2 + \varepsilon^2})$. Up to symplectomorphism this is independent of $\varepsilon$ and the isomorphism of $\mathcal{N}(Z)$ with $L(\varepsilon)$.

Given a symplectic embedding $\psi : B^{2n}(\lambda) \to M$, we can extend $\psi$ to an embedding of $B(\sqrt{\lambda^2 + \varepsilon^2})$ for some $\varepsilon > 0$. We can then define the symplectic blowup $(\widetilde{M}_\psi, \widetilde{\omega}_\psi)$ of $(M, \omega)$ of *weight* $\lambda$ to be the space,

$$\widetilde{M}_\psi = (M \setminus \psi(B(\lambda))) \sqcup L(\varepsilon)/\sim,$$

where we identify $([z], z] \in L(\varepsilon) \setminus Z$ with $\psi(\sqrt{1 + \lambda^2/|z|^2} z) \in M \setminus \psi(B(\lambda))$. The form $\widetilde{\omega}_\psi$ is given by $\omega$ off og $\psi(B(\lambda))$ and $\widetilde{\omega}_\lambda$ on $L(\varepsilon)$.

Implying a misnomer[1], the blowdown will *increase* the volume of the space by $\mathrm{vol}_{\omega_{\mathrm{std}}} B(\lambda)$ while the blowup with *decrease* the volume by the same amount. An illustrative but not really accurate picture, showing a two-dimensional *real* rather than complex blowup, is given in figure 11.

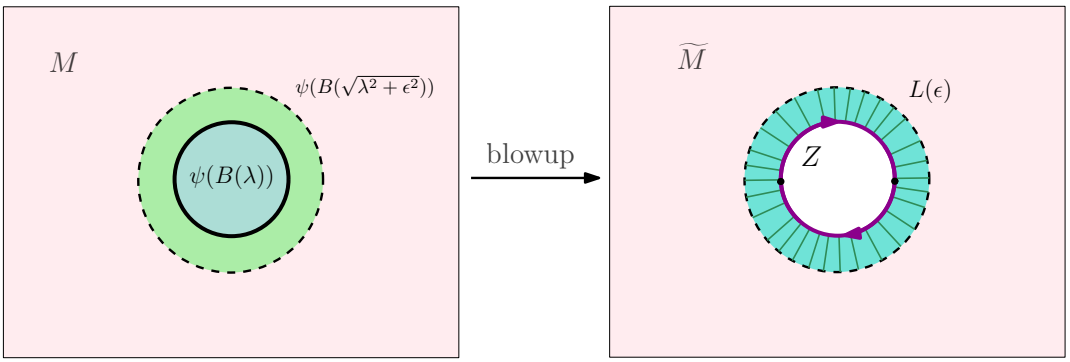



FIGURE 11. The blowup of $M$ as a excision of a ball

There is more one can say about symplectic blowups and blowdowns. Indeed one can consider blow ups associated to *normalized symplectic embeddings* that satisfy some compatibility with an almost-complex structure and these

[1]Actually, I am just parroting [**McS1**]; this doesn't seem like a misnomer to me. The blowup "blows up" a ball, eliminating it and replacing it with some data of its normal bundle.

will satisfy some general existence and uniqueness statements for the obtained symplectic forms up to isotopy of symplectic embeddings.

REMARK 6.5 (Symplectic Cut). Symplectic blowups can be seen as a special case of the following construction, called a symplectic cut. Suppose $(M, \omega)$ has a Hamiltonian $S^1$-action with momentum map $H$. Let $c$ be a regular value of $H$ so that $S^1$ acts freely on its preimage. We can define the symplectic manifold,

$$(M \times \mathbb{C}, \omega \oplus \mathrm{d}x \wedge \mathrm{d}y),$$

with Hamiltonian $\widetilde{H} = H - x^2 - y^2 - c$. This will be a momentum map for the product $S^1$-action,

$$e^{i\theta} \cdot (x, z) = (e^{i\theta} \cdot x, e^{-i\theta} z).$$

Note that $\widetilde{H}^{-1}(0)$ consists of pairs $(x, z)$ so that $H(x) = |z|^2 + c$. I.e. it consists of a copy of $H^{-1}(c)$ and $S^1 \times H^{-1}(b)$ for any $b > c$. We now form the *symplectic cut* as a symplectic reduction,

$$\widetilde{M_c} = (M \times \mathbb{C}) /\!\!/ S^1 = \widetilde{H}^{-1}(0)/S^1.$$

A dense open subset of this reduction is identified with $H^{-1}(c, \infty)$ and its complement is the codimension two symplectic reduction $H^{-1}(c)/S^1$. Thus we may see this construction as cutting off $M$ below $H^{-1}(c)$ and then closing up the boundary by taking a symplectic quotient.

Note if $M = \mathbb{C}^n$ and $H(z) = |z|^2$, then $\widetilde{M_\lambda}$ is the same as a weight $\lambda$ symplectic blowup of $\mathbb{C}^n$ at the origin. Explicitly, we cut out the ball $H^{-1}([0, \lambda))$ and replace it with $H^{-1}(\lambda)/S^1 = \mathbb{C}P^{n-1}$, just as in the blowup. The construction of symplectic cuts was purely local and thus it suffices to have a local Hamiltonian circle action. In this way, we can consider any symplectic blowup as a local symplectic cut. $\triangle$

**6.3. Blowups and Momentum Maps.** Consider the natural action of $U(n)$ on $\mathbb{C}^n \times \mathbb{C}P^{n-1}$. This restricts to an action on $\widetilde{\mathbb{C}}^n$. This action is actually Hamiltonian with respect to the form $\widetilde{\omega}_\lambda$, as we now explain. Recall the action on $\mathbb{C}^n$ is Hamiltonian with generating Hamiltonians,

$$H_\xi(z) = \frac{i}{2} z^* \xi z,$$

for $\xi \in \mathfrak{u}(n)$. And the action on $\mathbb{C}P^{n-1}$ is Hamiltonian with generating Hamiltonians,

$$H_\xi([z]) = \frac{i}{2|z|^2} z^* \xi z.$$

Combining these, once concludes the generating Hamiltonians for the action on $(\widetilde{\mathbb{C}}^n, \widetilde{\omega}_\lambda)$ are,

$$H_\xi(z,[w]) = \frac{i}{2} z^* \xi z + \frac{i\lambda^2}{2|w|^2} w^* \xi w.$$

The corresponding momentum map under the identification of $\mathfrak{u}(n)$ and its dual is,

$$\mu(z,[w]) = \frac{i}{2} z z^* + \frac{i\lambda^2}{2|w|^2} w w^*$$

Similarly, recall we had a $T^n$ action on $\mathbb{C}^n$ with generating Hamiltonians and momentum map,

$$H_\theta(z) = \frac{1}{2}\sum_{i=1}^n \theta_i |z_i|^2, \quad \mu(z) = \frac{1}{2}(|z_1|^2, \ldots, |z_n|^2).$$

This action gives one on $\mathbb{C}P^{n-1}$ and then on $\widetilde{\mathbb{C}}^n$. This is also Hamiltonian, generated by,

$$H_\theta(z,[w]) = \frac{1}{2}\sum_{i=1}^n \theta_i |z_i|^2 + \frac{\lambda^2}{2|w|^2}\sum_{j=1}^n \theta_j |w_j|^2.$$

And the momentum map will be,

$$\mu(z,[w]) = \frac{1}{2}\left(|z_1|^2 + \lambda^2 \frac{|w_1|^2}{|w|^2}, \ldots, |z_n|^2 + \lambda^2 \frac{|w_n|^2}{|w|^2}\right).$$

Suppose $x = (x_1, \ldots, x_n)$ is the image of this point $(z,[w])$ under the momentum map. One has,

$$\sum_{i=1}^n x_i = \frac{1}{2}\sum_{i=1}^n |z_i|^2 + \lambda^2 \frac{|w_i|^2}{|w|^2} = \frac{|z|^2}{2} + \frac{\lambda^2}{2} \geq \frac{\lambda^2}{2}.$$

We conclude,

$$\operatorname{im}\mu = \left\{ x \in \mathbb{R}^n_{\geq 0} : 2\sum x_i \geq \lambda^2 \right\}.$$

The effect of this is that the image of the momentum map of $\mu$ looks like the image of the momentum map for the $T^n$ action on $\mathbb{C}^n$ but with a simplex near zero removed.

More generally, suppose $M$ has a $T^n$-action with fixed point $p$ and we form the weight $\lambda$ blowup $\widetilde{M}$ at $p$ with exceptional divisor $Z$. The blowup inherits a $T^n$-action with momentum map $\widetilde{\mu}$. The momentum polytope $\widetilde{\mu}(\widetilde{M})$ of the momentum map $\widetilde{\mu}$ will be the old momentum polytope $\mu(M)$ but with the corner $\mu(B(\lambda))$ removed and replaced with a boundary face $\widetilde{\mu}(Z)$, which is the

codimension-one momentum polytope of the $T^n$-action restricted to $Z$. An illustration is given in Figure 12.

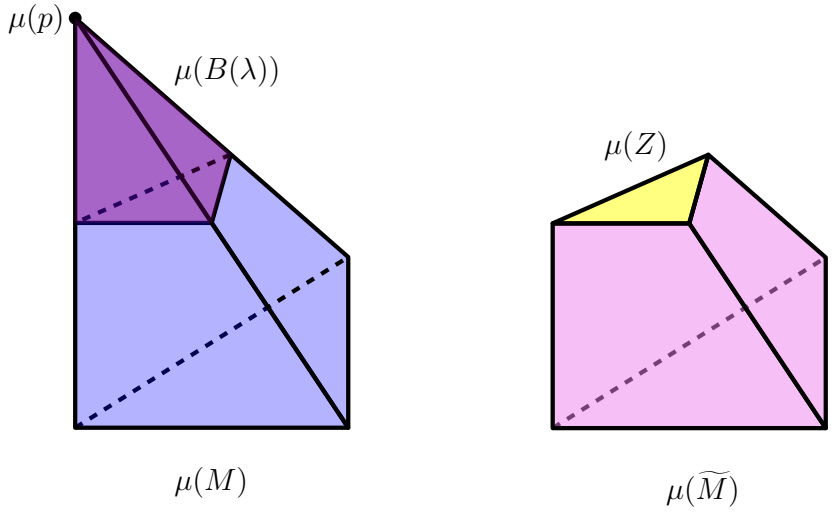



FIGURE 12. The momentum polytope before and after blowing up.

**6.4. Fibre Connected Sums.** Connected sums in the usual sense of differential topology do not play nicely with the category of symplectic manifolds of dimension greater than two. If one could form the connected sum, a spherical shell surrounding the sphere on which we glue the two manifolds would be glued by some symplectomorphism interchanging the two sides of the shell. But such a thing would allow us to define a symplectic structure on $S^{2n}$ by gluing collar neighbourhoods of the boundaries of two disks and hence cannot exist for $n > 1$.

In fact in dimension other than 2 and 6 one cannot even put a symplectic structure on a connect sum which on each summand is isotopic through non-degenerate forms to the original forms [**McS1**, Lemma 7.2.2].

We will instead be interested in the fibre connect sum construction of Gompf which works in greater generality and yields exciting new symplectic constructions.

Suppose $(M_1, \omega_1)$ and $(M_2, \omega_2)$ are $2n$-dimensional symplectic manifolds and $(Q, \tau)$ a $2n-2$ dimensional compact symplectic manifold. Further suppose there are embeddings $\iota_i : Q \to M_i$ for $i = 1, 2$ with trivial normal bundle. By the symplectic neighbourhood theorem, we can find symplectic embeddings,

$$f_i : Q \times B^2(\varepsilon) \to M_i, \quad f_i^* \omega_i = \tau \oplus \mathrm{d}x \wedge \mathrm{d}y \,,$$

which on $Q \times \{0\}$ act by $\iota_i$.

Define the annulus $A(\delta, \varepsilon) = B^2(\varepsilon) \setminus \text{int}(B^2(\delta))$ for $\delta < \varepsilon$. We can find an oriented area-preserving diffeomorphism $\varphi : A(\delta, \varepsilon) \to A(\delta, \varepsilon)$ that swaps the inner and outer boundary of the annulus (this is exactly what we couldn't do in higher dimensions to define a connect sum).

**Definition 6.6**

The *fibre connected sum* of $(M_1, \omega_1)$ and $(M_2, \omega_2)$ along $Q$ is,

$$M_1 \#_Q M_2 = (M_1 \setminus \mathcal{N}(Q_1)) \cup_\varphi (M_2 \setminus \mathcal{N}(Q_2)).$$

Here, $\mathcal{N}(Q_i) = f_i(Q \times B^2(\delta))$ and we glue along $\varphi$ via the identification $f_2(q, z) = f_1(q, \varphi(z))$ for $z \in A(\delta, \varepsilon)$.

The symplectic forms $\omega_i$ on $M_i$ agree on the overlap $Q \times A(\delta, \varepsilon)$ and hence glue to give a symplectic structure on the fibre connected sum.

This construction is independent of $\delta$ up to symplectomorphism, but certainly may depend on the isotopy class of the embeddings $f_i$.

Example 6.7. Gompf's fibre sum was generalized by McCarthy and Wolfson. We can still form the same connected sum if $Q$ does not have trivial normal bundle but instead the normal bundles in $M_1$ and $M_2$ are inverses (i.e. their Euler classes differ by a sign). Note that the complement of a hyperplane $\mathbb{C}P^{n-1}$ inside $\mathbb{C}^n$ is a ball and the hyperplane $\mathbb{C}P^{n-1}$ has normal bundle given by its tautological bundle with self intersection number one. If $M$ is a symplectic manifold with an embedded copy $Z$ of $\mathbb{C}P^{n-1}$ with self-intersection number minus one, then we can take the fibre connected sum of $(M, Z)$ and $(\mathbb{C}P^n, \mathbb{C}P^{n-1})$. The resulting space $M \#_Z \mathbb{C}P^n$ is just the blowdown of $M$ along the exceptional divisor $Z$. ▲

Example 6.8. Recall example 6.3 and consider the elliptic surface $E(1)$ with $T$ a torus fibre. $T$ will have trivial normal bundle and for any symplectic torus $T'$ in $(M, \omega)$ with trivial normal bundle, we can form $M' = M \#_{T'} E(1)$ (after rescaling $E(1)$ so the tori have the same area).

By Seifert–Van Kampen, $\pi_1(M') = \pi_1(M)/\langle i_*(\pi_1(T'))\rangle$, where $i : T' \to M$ is the inclusion. ▲

We will now see how to build many symplectic manifolds using this new construction.

**Theorem 6.9: Gompf**

Every finitely presented group $G$ is the fundamental group of some compact symplectic 4-manifold.

PROOF. Take some presentation for $G$ as generated by $g_1, \ldots, g_k$ with relations $r_1, \ldots, r_\ell$. Begin with $F$ a closed genus $k$ Riemann surface and choose a standard set of oriented simple closed curves $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_k$ to generate $H_1(F)$.

Now choose curves $\gamma_1, \ldots, \gamma_\ell$ representing the classes $r_i(\alpha_1, \ldots, \alpha_k)$. For $j = 1, \ldots, k$ set $\gamma_{\ell+j} = \beta_j$. Then we have,

$$G = \frac{\pi_1(F)}{\langle \gamma_1, \ldots, \gamma_{\ell+k} \rangle}.$$

Thus we just need to alter $F$ to kill off the loops $\gamma_i$ to obtain a space with fundamental group $G$.

We suppose $F$ has a closed 1-form $\rho$ which is a volume form on each loop $\gamma_i$. The existence of such a $\rho$ is proven in [**McS1**, Ex. 7.2.9]. Now form $X = F \times T^2$ with its product symplectic form $\omega$. Consider a non-contractible oriented simple closed curve $\alpha$ in $T^2$, and a 1-form $\theta$ which is positive in the positive tangent direction to $\alpha$. Let $T_i = \gamma_i \times \alpha$ be the immersed torus in $X$ for $i = 1, \ldots, \ell + k$. This is Lagrangian with respect to $\omega$ and symplectic with respect to $\rho \wedge \theta$.

For $\rho, \theta$ small in size, $\omega' = \omega + \rho \wedge \theta$ will be symplectic on $X$. And the tori $T_i$ will be symplectically immersed, as will the torus $\{z\} \times T^2$ for a fixed point $z$. By a perturbation, we can make all these tori symplectically embedded and disjoint (since they are codimension 2). The tori also all have trivial normal bundle.

Now form $X \#_T (\ell + k + 1) E(1)$, the connected sum of $X$ with $\ell + k + 1$ copies of the rational elliptic surface $E(1)$ from example 6.3 attached along the above tori. Each connected sum will change the fundamental group by killing the homotopy class of the attaching tori. Hence,

$$\pi_1(X \#_T (\ell + k + 1) E(1)) = \pi_1(F \times T^2) / \left\langle \pi_1(\gamma_i \times \alpha), \pi_1(\{z\} \times T^2) \right\rangle = G.$$

Thus this is the desired symplectic manifold. □

**6.5. Some More Results in Symplectic Topology.** We review a few more miscellaneous theorems and constructions related to constructing symplectic manifold. We describe the ideas of some proofs without really going into any detail.

*Telescope Construction.* A reasonable question to ask is about whether every non-degenerate two-form on a manifold is homotopic to a symplectic form. In his thesis, Gromov resolved this in the affirmative for open symplectic manifolds and proved a somewhat stronger statement. By open manifold, we mean there is no closed connected component of the manifold.

**Theorem 6.10: Gromov**

Let $M$ be an open $2n$-dimensional manifold, $\tau \in \Omega^2(M)$ a non-degenerate 2-form, and $a \in H^2(M;\mathbb{R})$.

(1) There is a smooth family $\tau_t$ of non-degenerate forms with $\tau_0 = \tau$ and $\tau_1$ in class $a$.
(2) If $\tau_t$ is a smooth non-degenerate family between symplectic forms $\tau_0$ and $\tau_1$ both in class $a$, then there is a smooth family of symplectic forms $\omega_t$ in class $a$ with $\omega_0 = \tau_0$ and $\omega_1 = \tau_1$.

The proof strategy is as follows. Assume $a = 0$. We consider some Morse function $f$ with finitely many critical points, no index $2n$ critical points, a single critical point of index 0, and so all critical points lie in distinct level sets. Then we construct the family $\tau_t$ so that $\tau_1$ is exact on the sub-level set $M^c = f^{-1}(-\infty, c]$. The crux of the theorem is based on extending the isotopy over a critical level. To do this, Gromov uses an idea called the *telescope construction*. Many generalizations of this idea in constructions of geometric structures has been important, and broadly this set of ideas is referred to as *h-principles*.

*Donaldson Submanifolds.* There is an interesting existence result for symplectic submanifolds that Donaldson has proved for integral symplectic manifolds, generalizing a much simpler construction in Kähler geometry (just look at the zero section of a line bundle).

**Theorem 6.11: Donaldson**

Let $(M, \omega)$ be a closed $2n$-dimensional symplectic manifold with an integral symplectic form. Let $a$ be an integer lift of $[\omega]$ and $k$ a sufficiently large integer. Then there exists a connected codimension two symplectic submanifold $Z_k$ in $M$ whose fundamental class is Poincaré dual to $ka$. If $k$ is large enough, this submanifold can be chosen so that the map

induced by inclusion $\pi_i(Z_k) \to \pi_i(M)$ is a bijection for $i \leq n-2$ and surjective for $i = n-1$.

The idea of the proof is now explained. Find a complex line bundle with first Chern class $a$ and choose a Hermitian structure and connection on $D$. Now we find a section $s : M \to L^{\otimes k}$ so that $|\overline{\partial} s| < |\partial s|$ on the zero locus of $s$. Given such a section (proving it exists is highly non-trivial analysis), it is a simple lemma [**McS1**, 7.4.2] to show that $s \pitchfork 0$ and the zero locus $Z$ is symplectic representing the class $a$ (the fact it represents class $a$ is just obstruction theory).

*Symplectic Lefschetz Fibrations.* Consider the following definition, important in 4-dimensional symplectic geometry.

**Definition 6.12**

Let $(M, w)$ be a closed symplectic 4-manifold. A *symplectic Lefschetz fibration on* $M$ is a smooth map $\pi : M \to \mathbb{C}P^1$ and a compatible almost complex structure $J$ so that:

(i) $\pi$ has finitely many critical values $t_1, \ldots, t_\ell \in \mathbb{C}P^1$ and each has a unique critical point $x_i \in \pi^{-1}(t_i)$, which must be non-degenerate.

(ii) Near each critical point $x_i$, $J$ is integrable and $\pi$ is holomorphic.

(iii) For every regular point $x$, $\omega$ is non-degenerate on the tangent space of the fibre $\ker \mathrm{d}\pi(x)$.

Over the regular values, $\pi$ will be a symplectic fibration with fibres being Riemann surfaces. We have the following result related to Donaldson's above theorem on submanifolds.

**Theorem 6.13: Donaldson**

Let $(M, \omega)$ be a compact symplectic 4-manifold with an integral symplectic form. Let $m$ be the (integer) symplectic volume of $M$. Then for sufficiently large $k$, the blowup $M \# mk^2 \overline{\mathbb{C}P^2}$ has the structure of a symplectic Lefschetz fibration $\pi$.

The generic fibre of $\pi$ is a Riemann surface $\Sigma_k$ with genus given by,

$$2g(\Sigma_k) - 2 = mk^2 + k\langle c_1(TM), [\omega]\rangle,$$

and the number of singular fibre is,

$$\ell = 3mk^2 + 2k\langle c_1(TM), [\omega]\rangle + \chi(M).$$

The idea of the proof, similar to the previous one, is to find two sections $s_1, s_2$ of $L^{\otimes k} \to M$ satisfying $|\overline{\partial} s_i| < |\partial s_i|$ that intersect simply and transversally. There are $mk^2$ intersection points and we can blowup $M$ at all of them. The resulting space has a projection $\pi$ to $\mathbb{C}P^1$ via $\pi(x) = [s_0(x) : s_1(x)]$ away from the blowup points and the identity on the exceptional divisors. Over $[t_0 : t_1] \in \mathbb{C}P^1$ the fibre will be the zero set of $t_1 s_0(x) - t_0 s_1(x)$. The critical points of the fibration are those for which this zero locus is not a smooth curve.

**6.6. Weinstein Domains.** Recall we earlier defined Liouville domains to be compact symplectic manifolds with boundary, whose boundary is of restricted contact type. This is the same as having a global Liouville vector field which is outward facing and transverse to the boundary. We now expand on this concept a bite more, and describe an important class of Liouville domains called Weinstein domains.

Given a Liouville domain $(M, \omega)$ with boundary $\partial M$ and Liouville vector field $X$. We can form the *symplectic completion* by gluing half of the symplectization of $\partial M$ to $M$ along the boundary. Explicitly, we form the non-compact manifold,

$$\widehat{M} := M \cup_{\partial M} [0, \infty) \times \partial M,$$

with gluing along $\{0\} \times \partial M$. If $\mathcal{L}_X \omega = \lambda$ and $\alpha = \lambda_{\partial M}$, then there is a collar neighbourhood $(-\varepsilon, 0] \times \partial M$ of the boundary so that,

$$\lambda|_{\{t\} \times \partial M} = e^t \alpha.$$

Hence we may define a symplectic form on $\widehat{M}$ given by $\omega$ on $M$ and $e^t \alpha$ on $[0, \infty) \times \partial M$, where $t$ is the coordinate projecting to $[0, \infty)$. By above, these smoothly glue along $\partial M$ to define a symplectic structure on $\widehat{M}$.

The symplectic completion is a special case of a *Liouville manifold*, which is a non-compact exact symplectic manifold with a complete Liouville vector field that is exhausted by Liouville subdomains.

Given a Liouville domain, we define its *skeleton* to be the set of points contained in the image of the flow of $X$ for all negative times. Equivalently, these are the points which never reach the boundary $M$ under the flow of $X$. Given a Liouville manifold, we can define its skeleton to be the union of the skeleta of the domains in its Liouville exhaustion. A Liouville manifold is *finite-type* if its skeleta is compact; in this case it will correspond to a symplectic completion of a Liouville domain.

EXAMPLE 6.14. Let $S$ be a starshaped domain about the origin in $\mathbb{R}^{2n}$. Recall this is Liouville using the Euler vector field. Then its symplectic completion $\widehat{S}$ is symplectomorphic to $\mathbb{R}^{2n}$.

Let $S^*M$ be the cotangent sphere bundle with respect to some metric on a compact manifold $M$; this is Liouville using the fibrewise Euler vector field. The symplectic completion $\widehat{S^*M}$ is symplectomorphic to $T^*M$. ▲

**Definition 6.15**

A Liouville domain $(M, \omega)$ with Liouville vector field $X$ is called a *Weinstein domain* if $X$ is *gradient-like*. I.e. if there is a Morse function $f : M \to \mathbb{R}$ so that,

$$X(f) \geq \delta(\|\mathrm{d}f\|^2 + \|X\|^2),$$

for some choice of metric and $\delta > 0$. Note in particular, this condition means $X$ and $f$ have the same critical points and $X(f) > 0$ otherwise (hence the name gradient-like). We additionally assume $f$ has no critical points on the boundary and $f$ is maximized precisely along $\partial M$.

A *Weinstein manifold* is a Liouville manifold $(M, \omega)$ with Liouville vector field $X$ so that $X$ is gradient-like for a Morse function $f : M \to \mathbb{R}$ which is proper and bounded below. A Weinstein manifold is *finite-type* if $f$ has finitely many critical points, in which case $M$ is the completion of a Weinstein domain.

The gradient-like condition means the regular level sets of $f$ are all restricted contact-type hypersurfaces. So $M$ possesses something like a singular fibration by contact-type level sets.

EXAMPLE 6.16. Again consider a starshaped domain about the origin in $\mathbb{R}^{2n}$. This is Weinstein since the Euler vector field is the gradient vector field of the radial Morse function $f(x) = \frac{1}{2}\|x\|^2$.

Similarly, consider a cotangent sphere bundle $S^*M$. The fibrewise Euler vector field $X$ is the gradient flow of the fibrewise norm $f(x, v) = \frac{1}{2}\|v\|_g^2$ induced from the metric $g$ on $M$. This function is only Morse–Bott, but we can pick a sufficiently small Morse function $\rho$ on the base. Now taking $f(x, v) = \frac{1}{2}\|v\|_g^2 + \rho(x)$ and $X = p\frac{\partial}{\partial p} + X_{\lambda(\nabla\rho)}$, $X$ will be gradient-like for $f$. This gives a Weinstein structure on $S^*M$. ▲

There are also notions of *Liouville and Weinstein cobordisms* $M$ which have boundary $\partial_- M \sqcup \partial_+ M$ so that the Liouville vector field is inwardly transverse

on $\partial_- M$ and outwardly transverse on $\partial_+ M$. If $\partial_- M$ is empty, we reobtain a Liouville/Weinstein manifold.

Not all Liouville domains are Weinstein. In fact being a Weinstein domain imposes a very stringent condition on the domain's topology.

PROPOSITION 6.17. *Let $(M, \omega)$ be a $2n$-dimensional Weinstein domain with Liouville vector field $X$ which is gradient-like for a Morse function $f$. Then the critical points of $f$ have index at most $n$. In particular, $M$ has the homotopy type of an $n$-dimensional CW complex.*

PROOF. All the usual concepts of Morse theory apply equally well to gradient-like vector fields. Let $p$ be a critical point of $f$. Consider the stable manifold $W^s(p)$ of $p$ under the flow of $X$. Note that this manifold is invariant under the forward flow of $X$. But because $X$ is Liouville, its forward flow by time $t$ scales the symplectic form by a factor of $e^t$. Hence if $\varphi_t$ is the flow and $\varphi_t(q) = p$, then $\omega_q = e^{-t}\varphi_t^* \omega_q$. Sending $t \to \infty$, $\omega_q \to 0$. Thus the symplectic form vanishes along the stable manifold and so $W^s(p)$ is isotropic. This means it must have dimension at most $n$. But by Morse theory, its dimension is the index of $p$ as a critical point of $f$.

Then by standard Morse theory, $f$ will determine a cell structure on $M$ with cells of dimension at most $n$. □

One clear restriction this imposes is that a Weinstein domain of dimension at least four must have connected boundary. There are Liouville domains that have disconnected boundary and hence are not Weinstein (finding these is non-trivial because the Liouville vector field has to be outwardly transverse on all boundary components).

One of the major features of Weinstein domains is their close relationship with an important class of complex manifolds.

**Definition 6.18**

A *plurisubharmonic function* on a complex manifold $(M, J)$ is a real-valued function $f$ so that $i\partial\overline{\partial} f > 0$. This is equivalent to saying $i\partial\overline{\partial} f$ defines a symplectic $(1,1)$ form which is $J$-compatible.

A *Stein manifold* is a connected complex manifold $M$ without boundary that admits a plurisubharmonic function which is bounded below and proper. A *Stein domain* is a complex manifold $M$ with boundary

admitting a plurisubharmonic function $f$ with no critical points on the boundary and which is maximized precisely along $\partial M$.

It is a classic result that the Stein manifolds are precisely the closed connected complex submanifolds of complex space $\mathbb{C}^n$. Indeed every such submanifold is Stein using $\|z\|^2$ as the plurisubharmonic function. In particular, smooth affine varieties are all Stein.

Note that the form $\omega := i\partial\overline{\partial}f$ is symplectic on a Stein manifold $(M, J)$. Moreover $\omega$ and $J$ are compatible and determine a metric $g$ on $M$. Then the gradient flow of $f$ with respect to $g$ defines a vector field $X$ which can be shown to be Liouville for $\omega$. After altering $f$ by some convex transformation of $\mathbb{R}$, we can ensure $X$ is complete. Hence every Stein manifold/domain is naturally a Weinstein manifold/domain.

Surprisingly, Cieliebak and Eliashberg have proved a result in the other direction. Given a Weinstein manifold $(M, \omega)$ with Morse function $f$, $M$ has an integrable almost complex structure $J$ and there is a diffeomorphism $\alpha : M \to M$ so that $\alpha_* f$ is a proper plurisubharmonic function on $(M, J)$.

Because every Stein manifold is Weinstein, we also conclude every $n$ (complex) dimensional Stein manifold has the homotopy type of an $n$-dimensional CW complex. This can be deduced from some kind of linear maximum principle, but the above argument gives a simple and symplectic proof of this fact. This is a basic illustration of how the relation between Stein and Weinstein manifold allows symplectic geometry to tell us something non-trivial about affine complex geometry. For more on this, one should consult the book of Cieliebak and Eliashberg [**CE**].

On $(\mathbb{R}^{2n}, \omega_{\text{std}})$, consider the Liouville vector field,

$$X_k = \sum_{j=1}^{n-k} \left( \frac{1}{2} x_j \partial_{x_j} + \frac{1}{2} y_j \partial_{y_j} \right) + \sum_{j=n-k+1}^{n} \left( 2x_j \partial_{x_j} - y_j \partial_{y_j} \right).$$

This is the gradient of the Morse function,

$$f_k = \sum_{j=1}^{n-k} \left( \frac{1}{4} x_j^2 + \frac{1}{4} y_j^2 \right) + \sum_{j=n-k+1}^{n} \left( x_j^2 - \frac{1}{2} y_j^2 \right).$$

We can consider a tubular neighbourhood of the unit disk in the $(y_{n-k+1}, \ldots, y_n)$ plane,

$$H_k^\varepsilon = \left\{ \sum_{j=1}^{n} x_j^2 + \sum_{j=1}^{n-k} y_j^2 < \varepsilon \right\} \cap \left\{ \sum_{j=n-k+1}^{n} y_j^2 \leq 1 \right\}.$$

This is naturally identified with a $k$-handle $D^k \times D^{2n-k}$. Moreover, it is transverse to $X_k$, and $X_k$ will flow in along the boundary $\partial_- H_k^\varepsilon := \partial D^k \times D^{2n-k}$ and out along the boundary $\partial_+ H_k^\varepsilon := D^k \times \partial D^{2n-k}$. Moreover, the sphere $S^{k-1} = \partial D^k \times \{0\}$ will be isotropic in $\partial_- H_k^\varepsilon$ with a symplectic trivialization of its normal bundle.

Suppose $M^{2n}$ is some Liouville domain whose boundary contains an isotropic sphere $S$ of dimension $k-1$ with symplectically trivial normal bundle. Then one can attach our handle along this sphere to form $M \cup_S H_k^\varepsilon$. We call $H_k^\varepsilon$ a *Weinstein handle* and call this procedure *Weinstein handle attaching*. The key result is the following.

PROPOSITION 6.19. *Let $M$ be a Liouville domain with a $(k-1)$-dimensional isotropic sphere $S$ with symplectically trivial normal bundle. Then the manifold $M'$ obtained by attaching the $k$-handle $H_k^\varepsilon$ along $S$ has the structure of a Liouville domain. If $M$ is Weinstein, then so is $M'$.*

The key idea here is just to continue the Liouville vector field along the handle using the vector field $X_k$ given above. In the Weinstein case, we extend our Morse function using $f_k$ given above.

It turns out that every Weinstein domain admits a *Weinstein handle decomposition*. I.e. it can be obtained from a ball by a finite sequence of Weinstein handle attachments. This opens up the study of Weinstein manifolds into more topological methods that ask how the geometry/topology of the manifold changes under Weinstein handle attaching.

A $2n$-dimensional Weinstein manifold (resp. a Stein manifold) is called *subcritical* if it has the homotopy type of an $(n-1)$-dimensional CW complex. I.e. it can be built from Weinstein handles of dimension at most $n-1$. These subcritical manifolds have special properties, for example Cieliebak showed they must have vanishing symplectic homology (an extension of Floer homology to non-closed symplectic manifolds).

## 7. Pseudoholomorphic Curves and Their Moduli Spaces

Now we begin the study of pseudoholomorphic curves in earnest. These objects are fundamental to many facets of modern symplectic geometry, including all of Floer theory. We will start by describing some basic properties of holomorphic curves before embarking on a study of the moduli spaces they compose. We will also sketch a proof of the Gromov non-squeezing theorem and a couple other results in symplectic topology.

**7.1. Basic Properties of Pseudoholomorphic Curves.** Let $(M, \omega)$ be a symplectic manifold equipped with an almost-complex structure $J$ which we usually assume to be tame or compatible. Recall, for $(\Sigma, j)$ a Riemann surface, we define a $J$-holomorphic curve to be a map $u : \Sigma \to M$ whose differential is complex linear:

$$J \circ \mathrm{d}u = \mathrm{d}u \circ j.$$

We will define the Cauchy–Riemann operator,

$$\overline{\partial}_J(u) = \frac{1}{2}(\mathrm{d}u + J \circ \mathrm{d}u \circ j) \in \Omega^{0,1}(\Sigma, u^*TM).$$

$u$ being $J$-holomorphic is equivalent to $\overline{\partial}_J u = 0$.

Consider the space $\mathcal{B} = C^\infty(\Sigma, M)$ with infinite-dimensional vector bundle over it $\mathcal{E}$ with fibre at $u$, $\mathcal{E}_u = \Omega^{0,1}(\Sigma, u^*TM)$. Then $\overline{\partial}_J$ can be regarded as a section of the $\mathcal{E}$ whose zero set is the collection of $J$-holomorphic curves.

The $J$-holomorphic curve equation is elliptic and after Sobolev completion of $\mathcal{B}$ and $\mathcal{E}$, $\overline{\partial}_J$ will be a Fredholm section.

We have a notion of energy for our maps, which allows us to think of $J$-holomorphic maps as some kind of minimal surface. We define a norm on linear maps $L : T_z\Sigma \to T_xM$ by,

$$|L|_J = \frac{1}{|\xi|}\sqrt{|L(\xi)|^2 + |L(j\xi)|^2},$$

where $\xi \in T_z\Sigma \setminus 0$ and we use the norm induced on $T_xM$ by our tame or compatible almost-complex structure $J$.

We then define the *energy* of $u : \Sigma \to M$ to be,

$$E(u) := \frac{1}{2}\int_\Sigma |\mathrm{d}u|_J^2 \,\mathrm{dvol}\,.$$

This is independent of the volume form on $\Sigma$.

Lemma 7.1. *If $J$ is $\omega$-tame then, for any $J$-holomorphic $u$,*

$$E(u) = \int_\Sigma u^*\omega.$$

*If $J$ is $\omega$-compatible, then for any smooth $u$,*

$$E(u) = \int_\Sigma |\overline{\partial}_J u|_J^2 \,\mathrm{dvol} + u^*\omega.$$

PROOF. We work locally with $\Sigma$ an open neighbourhood of $\mathbb{C}$ with coordinates $z = s + it$. Then,

$$\begin{aligned}\frac{1}{2}|\mathrm{d}u|_J^2 \,\mathrm{dvol} &= \frac{1}{2}(|\partial_s u|_J^2 + |\partial_t u|_J^2)\,\mathrm{d}s \wedge \mathrm{d}t \\ &= \frac{1}{2}(|\partial_s u + J\partial_t u|_J^2 - \langle \partial_s u, J\partial_t u\rangle_J)\,\mathrm{d}s \wedge \mathrm{d}t \\ &= |\overline{\partial}_J u|_J^2 \,\mathrm{dvol} + \frac{1}{2}(\omega(\partial_s u, \partial_t u) + \omega(J\partial_s u, J\partial_t u))\,\mathrm{d}s \wedge \mathrm{d}t\,.\end{aligned}$$

The result follows. □

Note this implies that the energy is purely topological when $u$ is $J$-holomorphic for $J$ $\omega$-tame and $\Sigma$ is closed (or $u$ has fixed Laganragian voundary conditions). It also implies that when $J$ is $\omega$-compatible, $J$-holomorphic curves are harmonic maps.

Here are some simple properties we should know.

**Theorem 7.2: Unique Continuation**

Suppose $u, v$ are $C^1$ functions from a ball around the origin in $\mathbb{C}$ to $\mathbb{C}^n$ satisfying the $J$-holomorphic curve equation,

$$\partial_s u + J(u)\partial_t u = 0,$$

for some $C^1$ almost-complex structure $J : \mathbb{C}^n \to GL(\mathbb{R}^{2n})$. If $u - v$ vanishes to infinite order at zero, i.e.,

$$\int_{|z|\leq r} |u(z) - v(z)| = O(r^k) \quad \text{for all} \quad k > 0,$$

then $u = v$.

This result can be deduced from some Sobolev theory using bounds given by the ellipticity of the holomorphic curve equation (see [**McS2**, §E.4]).

COROLLARY 7.3. *Let $\Sigma$ be a connected Riemann surface and $J$ a $C^1$ almost complex structure on $M$. If $u, v$ are $J$-holomorphic curves which agree to infinite order at a point, then $u = v$.*

PROOF. Working locally, the set of points with $u = v$ must be open. But it is also closed. Hence all of $M$. □

We define a *critical point* of $u : \Sigma \to M$ if $u$ is $J$-holomorphic to be a points $z \in \Sigma$ so that $\mathrm{d}u\,(z) = 0$. The image of a critical point is a *critical value*, otherwise points in $u(\Sigma)$ are called *regular*.

**Theorem 7.4: Critical Points of Holomorphic Curves**

Let $\Sigma$ be a closed Riemann surface and $J$ a $C^1$ almost complex structure on $M$. If $u$ is $J$-holomorphic and non-constant, then the set of preimages of critical values is finite. For each point $x \in M$, its preimage under $u$ is finite.

**Theorem 7.5: Curve Intersection**

Let $\Sigma_0, \Sigma_1$ be closed connected Riemann surfaces and $J$ be a $C^2$ almost complex structure on $M$. suppose we have $J$-holomorphic maps $u_0, u_1$ from $\Sigma_0, \Sigma_1$ so that their images are not the same and they are non-constant. Then $u_0^{-1}(u_1(\Sigma_1))$ is at most countable and its only accumulation points can be the critical points of $u_0$.

**7.2. Somewhere Injectivity.** We now discuss a fundamental condition we can impose on our curves, which will be essential to having regularity in our moduli spaces of pseudoholomorphic curves.

**Definition 7.6: Simple and Somewhere Injective Curves**

We say a $J$-holomorphic curve $u : \Sigma \to M$ is *multiply-covered* if it factors through a holomorphic branched multiple cover of Riemann surfaces. That is, if one can write $u = u' \circ \varphi$ where $u' : \Sigma' \to M$ is $J$-holomorphic and $\varphi : \Sigma \to \Sigma'$ is a holomorphic map of degree greater than one. If a curve is not multiply-covered, we say it is *simple*.

A $J$-holomorphic curve $u : \Sigma \to M$ is *somewhere injective* if there is a point $z \in \Sigma$ so that $\mathrm{d}u\,(z) \neq 0$ and $z$ is the only point in the preimage of $u(z)$. Such a point is called an *injective point* of $u$.

Suppose $u = u' \circ \varphi : \Sigma \to M$ is multiply covered. Then the only points at which $u$ could be an injective map are the branch points of $\varphi$. But at every such point $z$, $\varphi$ locally looks like $z^d$ where $d = \deg(\varphi)$ and hence $\mathrm{d}u\,(z) = 0$. Hence $u$ is not somewhere injective.

From this we conclude that every somewhere injective curve is simple. In fact if $\Sigma$ is closed, the converse holds so that somewhere injectivity characterizes the simple curves.

**Theorem 7.7: Somewhere Injective Curves**

Let $J$ be a $C^2$ almost complex structure on $M$ and $\Sigma$ a closed Riemann surface. Any simple $J$-holomorphic curve $u : \Sigma \to M$ is somewhere injective. The set of non-injective points of $u$ is countable and its accumulation points can only be at critical points of $u$.

Note this result requires that $\Sigma$ be closed, for example it is not true for $\Sigma$ a disk. In the case of the theorem, the set of injective points will in fact be cofinite. This follows from the Micallef–White theorem which we state shortly.

Proof. (Sketch) We claim that any non-constant $J$-holomorphic curve $u : \Sigma \to M$ factors as $u = v \circ \varphi$ where $\varphi : \Sigma \to \Sigma'$ is a holomorphic map of positive degree and $v : \Sigma' \to M$ is injective on an open dense set. The result then clearly follows.

To prove this, we can construct $\Sigma'$ explicitly from the image of $u$ away from critical points and points where branches of $u$ meet. Around a branch meeting point, the image of $u$ is diffeomorphic to a finite set of disjoint punctured disks, so we can fill in the points where branches meet in the natural way to remove the non-injectivity. We can fill in the critical points too to obtain a closed Riemann surface $\Sigma'$. We may then take $v : \Sigma' \to M$ to be the inclusion on injective points and the natural map on non-injective points. Then $\varphi : \Sigma \to \Sigma'$ is just given by $u$ away from the singularities and then extends by a removable singularities theorem.

The details of this require some knowledge of what $J$-holomorphic curves look like near singularities. We can use the curve intersection theorem above and the following lemma, which follows from Micallef–White.

Lemma 7.8 (Branching). *Let $u : \Sigma \to M$ be a non-constant $J$-holomorphic curves with a critical point $z_0$. Then there is a neighbourhood $U$ of $z_0$ biholomorphic to the unit disk in $\mathbb{C}$ so that $u(z) = v(z^k)$ on $U$ where $k \in \mathbb{N}$ and $v : D \to M$ is an injective $J$-holomorphic map with no critical points except possibly at the origin.*

From this we deduce we can fill in $\Sigma'$ at a critical point. □

Remark 7.9. To illustrate why somewhere injectivity is important for regularity, we now sketch an argument that we will detail later.

For a pair of a map $u : \Sigma \to M$ and a compatible almost complex structure $J$, we can define the map $\mathcal{F}(u) = \overline{\partial}_J(u)$. After suitable Sobolev completion, this defines a section of a Banach bundle.

This section has a differential which in a trivialization of the fibres defines a map,

$$D\mathcal{F}_{(u,J)} : W^{k,p}(u^*TM) \times C^\ell(\mathrm{End}(TM, J, \omega)) \to W^{k-1,p}(\Lambda^{0,1}T^*\Sigma \otimes_J u^*TM),$$

where we work with $C^\ell$ almost-complex structures. This differential is given by,

$$D\mathcal{F}_{(u,J)}(\xi, Y) = D_u\xi + \frac{1}{2}Y(u) \circ \mathrm{d}u \circ j.$$

In order that the universal moduli space of pseudoholomorphic curves (i.e. the zero set of $\mathcal{F}$) be smooth at a pair $(u, J)$ with $\mathcal{F}(u, J) = 0$, we need $D\mathcal{F}_{(u,J)}$ to be surjective (this is just the implicit function theorem and transversality). Since $D_u$ is Fredholm, it is enough to prove $D\mathcal{F}_{(u,J)}$ has dense image. If not, we could find $\eta \in W^{k-1,p^*}(\Lambda^{0,1}T^*\Sigma \otimes_J u^*TM)$ which is non-zero and so that,

$$\int_\Sigma \langle \eta, D_u\xi \rangle \mathrm{dvol} = 0 \quad \text{and} \quad \int_\Sigma \langle \eta, Y(u) \circ \mathrm{d}u \circ j \rangle \mathrm{dvol} = 0.$$

The firs equation implies $D_u^*\eta = 0$. By elliptic regularity, we conclude $\eta$ would be zero if it vanishes on an open set. But if $z$ is an injective point of $u$ (and hence a neighbourhood of $z$ is too) then we can always find a $Y$ near $z$ to pair positively with $\eta$. Hence $\eta$ vanishes near $z$ and thus on all of $\Sigma$. Thus somewhere injectivity of $u$ guarantees regularity of the moduli space at $u$. $\triangle$

**7.3. Properties of Four-Dimensional Curves.** We now explore a couple special properties of four dimensional $J$-holomorphic curves. These come in useful for special applications of these curves to three and four dimensional theories, like the Gromov–Taubes invariants and embedded contact homology.

**Definition 7.10**

Let $(M, J)$ be an almost complex 4-manifold. Suppose $u_i : \Sigma_i \to M$ are closed $J$-holomorphic curves for $i = 1, 2$. A *transverse intersection of $u_1$ and $u_2$* is a pair $(z_0, z_1) \in \Sigma_0 \times \Sigma_1$ so that $u_0(z_0) = u_1(z_1) = x$ and $T_xM = \mathrm{im}\,\mathrm{d}_{z_0}u_0 \oplus \mathrm{im}\,\mathrm{d}_{z_1}u_1$. If $u_0 = u_1$, we call such a pair a *transverse self-intersection.*

The number of intersection points of a pair of curves is denoted,

$$\delta(u_0, u_1) = \left|\left\{(z_0, z_1) \in \Sigma_0, \Sigma_1 : u_0(z_0) = u_1(z_1)\right\}\right|.$$

For a single curve, we count unordered distinct intersection pairs,

$$\delta(u) = \frac{1}{2}\Big|\big\{(z_0, z_1) \in \Sigma_0, \Sigma_1 : z_0 \neq z_1 \text{ and } u_0(z_0) = u_1(z_1)\big\}\Big|.$$

These intersection numbers will be finite for simple curves. Moreover, as the following result shows, simple holomorphic intersections always contribute positively to the homological intersection number.

**Theorem 7.11: Intersection Positivity**

Let $(M, J)$ be an almost-complex 4-manifold and $A_0, A_1 \in H_2(M; \mathbb{Z})$ homology classes represented by simple $J$-holomorphic curves $u_0, u_1$. Suppose the union of the curves is simple so that their intersection is discrete. Then the intersection number of the curves is controlled by the homological intersection product,

$$\delta(u_0, u_1) \leq A_0 \cdot A_1.$$

One has equality if and only if $u_0 \pitchfork u_1$.

Note that we require the curves representing $u_0, u_1$ to be distinct. It may be that a single curve $u$ can have negative self intersection, for example the exceptional divisor of a blowup in a complex manifold. However, the above result then implies that such a curve must be the unique holomorphic curve in its homology class.

A related result for self-intersections of a simple holomorphic curve is now stated.

**Theorem 7.12: Adjunction Inequality**

Let $(M, J)$ be an almost-complex 4-manifold and $A \in H_2(M; \mathbb{Z})$ a homology class represented by a $J$-holomorphic curve $u : \Sigma \to M$. We have an inequality,

$$2\delta(u) - \chi(\Sigma) \leq A \cdot A - \langle c_1(TM), A\rangle.$$

One has equality if and only if $u$ is an immersion with transverse self-intersections.

Remark 7.13. One can see this as arising from the exact sequence of bundles,

$$0 \to T\Sigma \to u^*TM \to N\Sigma \to 0.$$

The first Chern number of $u^*TM$ is the sum of $c_1(T\Sigma) = \chi(\Sigma)$ and $c_1(N\Sigma)$. The intersection number $A \cdot A$ is given both by looking at self-intersections in

the normal bundle (i.e. from $c_1(N\Sigma)$) and the literal intersection indices of the curve (which is bounded below by $2\delta(u)$ by intersection positivity). $\triangle$

Both results and other important results follow from the Micallef–White theorem. This result describes the structure of $J$-holomorphic curves in local coordinates.

**Theorem 7.14: Micallef–White**

Let $u : \Sigma \to M$ be $J$-holomorphic curve and $\Omega \subset M$ a precompact open set. Suppose $x_0$ is in the image of $\Omega$ away from the boundary and $\{z_1, \ldots, z_n\} = u^{-1}(x_0) \cap \Omega$.

Then there are a $C^2$ charts $\varphi_i : (U_i, z_i) \to (\mathbb{C}, 0)$ on disjoint opens $U_i \subset \Omega$, unitary $n \times n$ matrices $L_i$, and a $C^1$ chart $\Psi : (V, x_0) \to (\mathbb{C}^n, 0)$ defined on $V \subset M$ an open neighbourhood of $x_0$. These satisfy:

(i) $\Psi$ preserves the complex structure: $\Psi_* J - J_0$.

(ii) The coordinate functions $u_i = \Psi \circ u \circ \varphi_i^{-1}$ are polynomials.

(iii) The matrices satisfy,

$$L_i u_i(z) = (z^{k_i}, z^{k_i} p_i(z)),$$

where $k_i$ is a positive integer and $p_i : \mathbb{C} \to \mathbb{C}^{n-1}$ is a polynomial with $p_i(0) = 0$.

Given two $J$-holomorphic curves $u_0, u_1$ with isolated intersection $x = u_0(z_0) = u_1(z_1)$, one can pick a contractible neighbourhood $U$ of $x$ and deform the curves on $u_i^{-1}(U)$ to intersect transversally. One defines the *local intersection number* $i(u_0, u_1, x)$ of the $u_i$'s at $x$ to be the oriented intersection count of the deformed curves on $U$. This extends naturally to a notion of local self-intersection number.

This Micallef–White theorem implies the following intuitive result from which intersection positivity and the adjunction inequality readily follows.

PROPOSITION 7.15. *Let $(M, J)$ be an almost-complex 4-manifold and $u : \Sigma \to M$ a simple $J$-holomorphic curve. Every self-intersection $x$ of $u$ is isolated and has local intersection number $i(u_0, u_1, x) \geq 1$ with equality if and only if $u_0 \pitchfork u_1$ at $x$.*

The inequalities in intersection positivity and the adjunction formula just become basic topological equivalences if we replace $\delta(u_0, u_1)$ and $\delta(u)$ with the sum of local intersection numbers $i(u_0, u_1)$ and $i(u, u)$. The inequalities then follow from the above proposition applied to $u_1 \cup u_2$ or $u$, since the sum of

local intersection numbers is at least as large as the actual count of intersection points.

**7.4. Moduli Spaces of Curves, Transversality, and Regularity.** $J$-holomorphic curves on their own are interesting geometric objects. But to get a fuller geometric understanding of when these curves exist and how plentiful they are, we want to shift to describing moduli spaces parameterizing collections of $J$-holomorphic curves (possibly modulo some symmetry).

We will want these moduli spaces to have some nice structure and regularity, in particular they should be smooth manifolds with a natural compactification. In general, as we sketched previously, one needs to restrict to the space of simple curves to have a smooth structure on one's moduli space. Another option to achieve regularity is to allow the almost-complex structure $J$ to vary parameterized by $\Sigma$, and/or to make a Hamiltonian perturbation of the pseudoholomorphic curve equation. This second approach is standard in the construction of Hamiltonian Floer theory, but here we will focus on the first and usually restrict to simple curves.

**Definition 7.16**

Let $(M, \omega)$ be a compact symplectic manifold and $(\Sigma, j)$ a compact Riemann surface. For any homology class $A \in H_2(M; \mathbb{Z})$ we can define the *moduli space of $J$-holomorphic curves* representing $A$,

$$\mathcal{M}(A, \Sigma; J) := \{u \in C^\infty(\Sigma, M) : \overline{\partial}_J u = 0,\ [u] = A\}.$$

We can also restrict to the moduli space of simple curves,

$$\mathcal{M}^*(A, \Sigma; J) := \{u \in \mathcal{M}(A, \Sigma; J) :\ u \text{ is simple}\}.$$

For the case $\Sigma = \mathbb{C}P^1$, we just write $\mathcal{M}(A; J)$ and $\mathcal{M}^*(A; J)$.

Let $\mathcal{B}$ be the space of smooth maps $\Sigma \to M$ representing the homology class $A$. This is a Fréchet manifold with tangent space $T_u\mathcal{B} = \Gamma(u^*TM)$. Over this space is a infinite-dimensional vector bundle $\mathcal{E} \to \mathcal{B}$ with fibre $\mathcal{E}_u = \Omega^{0,1}(\Sigma, u^*TM)$. Then the Cauchy–Riemann operator defines a section $\mathcal{F}(u) = (u, \overline{\partial}_J u)$. The space $\mathcal{M}(A, \Sigma; J)$ is the zero section of $\mathcal{F}$ and $\mathcal{M}^*$ is its intersection with the open subset $\mathcal{B}^* \subset \mathcal{B}$ of somewhere-injective maps.

To prove (generic) regularity, we need that $\mathcal{F}$ intersects the zero section of our bundle transversally, i.e. its differential at a holomorphic curve should span the vertical direction. For $u \in \mathcal{M}^*(A, \Sigma; J)$, we define the *vertical differential*,

$$D_u : \Gamma(\Sigma, u^*TM) \to \Omega^{0,1}(\Sigma, u^*TM),$$

to be the composition of $\mathrm{d}\mathcal{F}_u : T_u\mathcal{B} \to T_{(u,0)}\mathcal{E}$ with the projection,

$$\pi_u : T_{(u,0)}\mathcal{E} = T_u\mathcal{B} \oplus \mathcal{E}_u \to \mathcal{E}_u.$$

Thus for transversality at $u$, we need $D_u$ to be surjective.

Take local coordinates $z = s+it$ on $\Sigma$ so that the $J$-holomorphic curve equation becomes $\partial_s u + J(u)\partial_t u = 0$. Considering a tangent vector, i.e. a vector field $\xi$ along $u$ and differentiating this equation in the $\xi$ direction gives,

$$D_u\xi = \eta\,\mathrm{d}s - J(u)\eta\,\mathrm{d}t \quad \text{where} \quad \eta = \frac{1}{2}(\partial_s\xi + J(u)\partial_t\xi + \partial_\xi J(u)\partial_t u).$$

Using that $u$ is $J$-holomorphic, this simplifies to,

$$D_u\xi = \overline{\partial}_J\xi - \frac{1}{2}(J\partial_\xi J)\partial_J u.$$

We can extend $D_u$ to an operator for arbitrary smooth maps $\Sigma \to M$. If $\nabla$ is the Levi–Civita connection of the metric induced by the tame pair $(\omega, J)$, then we set,

$$D_u : \Gamma(\Sigma, u^*TM) \to \Omega^{0,1}(\Sigma, u^*TM),$$

to be given by,

$$D_u\xi = \frac{1}{2}(\nabla\xi + J(u)\nabla\xi \circ j) - \frac{1}{2}J(u)\nabla_\xi J\partial_J u.$$

This agrees with the projection of $\mathrm{d}\mathcal{F}_u$ to the vertical direction determined by the connection. In particular, it is equal to $D_u$ as defined above when $u$ is $J$-holomorphic. As above, $D_u$ is the sum of a complex-linear Cauchy–Riemann operator and a zeroth order complex anti-linear operator. In particular it is Fredholm; in the surjective case, its index will describe the local dimension of the moduli space cut out by $\overline{\partial}_J$. We will often consider the same operator extended to a Sobolev completion $W^{k,p}$ for $k \geq 1$ (so differentiating is defined) and $p > 2$ (so that Sobolev embedding means our functions are $C^0$).

We now begin in earnest the process of proving regularity, for the space of simple curves and generic tame almost complex structures. We will also want regularity for the space of curves over a one-parameter family of complex structures, which will allow us to describe cobordisms between moduli spaces of $J$-holomorphic curves for different $J$.

**Definition 7.17**

Let $(\Sigma, j)$ be a compact Riemann surface and $A \in H_2(M;\mathbb{Z})$. An almost complex structure $J$ on $M$ is *regular* if the operator $D_u$ is surjective for all $u \in \mathcal{M}^*(A, \Sigma; J)$. The set of all regular structures is denoted $\mathcal{J}_{\mathrm{reg}}(A, \Sigma)$.

Here are the basic results we are working towards.

**Theorem 7.18: Regularity of the Moduli Space of Simple $J$-Holomorphic Curves**

Let $(\Sigma, j)$ be a compact Riemann surface and $A \in H_2(M;\mathbb{Z})$ for $(M,\omega)$ closed symplectic.

(i) If $J \in \mathcal{J}_{\text{reg}}(A,\Sigma)$, then the moduli space $\mathcal{M}^*(A,\Sigma;J)$ is a naturally oriented smooth manifold of dimension,

$$\dim \mathcal{M}^*(A,\Sigma;J) = 2(2-2g) + 2c_1(A). \tag{7.1}$$

(ii) The set $\mathcal{J}_{\text{reg}}(A,\Sigma)$ is a countable intersection of open dense sets (second category) in the space of compatible or tame almost-complex structures on $M$.

Given two almost-complex structures $J_0, J_1 \in \mathcal{J}_{\text{reg}}(A,\Sigma)$, consider a smooth homotopy $J_\lambda$ between them. This homotopy defines a fibration of moduli spaces over $[0,1]$,

$$\mathcal{W}^*(A,\Sigma;J_\lambda) := \{(\lambda,u) : \lambda \in [0,1],\ u \in \mathcal{M}^*(A,\Sigma;J_\lambda)\}.$$

**Definition 7.19**

Let $(\Sigma, j)$ be a compact Riemann surface, $A \in H_2(M;\mathbb{Z})$ for $(M,\omega)$ closed symplectic, and $J_0, J_1 \in \mathcal{J}_{\text{reg}}(A,\Sigma)$. A smooth homotopy $J_\lambda$ of tame or compatible almost-complex structures between $J_0, J_1$ is *regular* if for every $(\lambda,u) \in \mathcal{W}^*(A,\Sigma;J_\lambda)$,

$$\Omega^{0,1}(\Sigma, u^*TM) = \operatorname{im} D_{J_\lambda,u} \oplus \mathbb{R}v_\lambda,$$

where $v_\lambda = (\partial_\lambda J_\lambda)\, du \circ j$, i.e. it is the image in $\Omega^{0,1}$ of the tangent vector of $\lambda \mapsto J_\lambda$. The set of all regular homotopies is denoted $\mathcal{J}_{\text{reg}}(A,\Sigma;J_0,J_1)$.

**Theorem 7.20: Regularity of the Cobordism of Moduli Space of Simple $J$-Holomorphic Curves**

Let $(\Sigma, j)$ be a compact Riemann surface, $A \in H_2(M;\mathbb{Z})$ for $(M,\omega)$ closed symplectic, and $J_0, J_1 \in \mathcal{J}_{\text{reg}}(A,\Sigma)$.

(i) If $J_\lambda \in \mathcal{J}_{\text{reg}}(A,\Sigma;J_0,J_1)$, then the moduli space $\mathcal{W}^*(A,\Sigma;J_\lambda)$ is a smooth oriented manifold with boundary,

$$\partial \mathcal{W}^*(A,\Sigma;J_\lambda) = \mathcal{M}^*(A,\Sigma;J_1) \sqcup -\mathcal{M}^*(A,\Sigma;J_0).$$

That is, it is an oriented cobordism of the boundary moduli spaces.

(ii) The set $\mathcal{J}_{\text{reg}}(A, \Sigma; J_0, J_1)$ is a countable intersection of open dense sets (second category) in the space of all smooth maps $[0,1] \to \mathcal{J}$, where $\mathcal{J}$ denotes all tame or all compatible almost-complex structures.

REMARK 7.21. The oriented cobordism between moduli spaces will only be interesting if we can establish the spaces in question are compact. This will be the work of the theory of Gromov compactness. Once we establish this, we will learn for example that when (7.1) yields a dimension of zero, that the moduli spaces $\mathcal{M}^*$ are finite set of points and their oriented count is independent of $J \in \mathcal{J}_{\text{reg}}$. △

As noted above, to apply the functional analysis theory, we need to work in some suitable Sobolev completions of our spaces of smooth maps. We introduce the terminology and background results now.

We will write $\mathcal{J}^\ell$ to denote the $\omega$-tame or $\omega$-compatible almost-complex structures of class $C^\ell$; we will be interested in the case of $\ell \geq 2$. We will write $\mathcal{B}^{k,p}$ for the space of $W^{k,p}$ maps $u : \Sigma \to M$ representing a homology class $A$. The space $W^{k,p}(\Sigma, M)$ is a smooth separable Banach manifold with tangent space,

$$T_u W^{k,p}(\Sigma, M) = W^{k,p}(\Sigma, u^*TM).$$

There are charts on this Banach manifold given by exponentiating vector fields along a curve $u$. Then $\mathcal{B}^{k,p}$ is just a connected component of $W^{k,p}(\Sigma, M)$ and hence is a smooth separable Banach manifold too.

We will assume that $kp > 2$ so that Sobolev embedding implies maps in $\mathcal{B}^{k,p}$ are continuous and Rellich's theorem implies the embedding $\mathcal{B}^{k,p} \hookrightarrow C^0(\Sigma, M)$ is compact. Using elliptic bootstrapping, we can prove the following, which shows we have not gained any new $J$-holomorphic curves by extending our function spaces.

PROPOSITION 7.22 (Elliptic Regularity). *Let $J \in \mathcal{J}^\ell$ for $\ell \geq 1$. If $u : \Sigma \to M$ is a $J$-holomorphic curve in $W^{1,p}$ for $p > 2$, then $u$ is in $W^{\ell+1,p}$. Hence $u$ is $C^\ell$. If $J$ is smooth, so is $u$.*

We can extend our vertical differential $D_u$ to an operator on the Sobolev completion. It will now define a formal adjoint $D_u^*$ by,

$$\int_\Sigma \langle \eta, D_u \xi \rangle \, \text{dvol} = \int_\Sigma \langle D_u^* \eta, \xi \rangle \, \text{dvol}\,,$$

for $\xi \in W^{k,p}(\Sigma, u^*TM)$, and $\eta \in W^{k,p}(\Sigma, \Lambda^{0,1} \otimes_J u^*TM)$.

PROPOSITION 7.23 (Fredholm Differential). *Let $J \in \mathcal{J}^\ell$ for $\ell \geq 1$ and $u \in W^{\ell,p}(\Sigma, M)$ for $p > 2$. Let $k \in \{1, \ldots, p\}$ and $q$ be the Hölder conjugate $p^{-1} + q^{-1} = 1$.*

*(i) $D_u$ and its adjoint $D_u^*$ are Fredholm with index,*

$$\operatorname{ind} D_u = -\operatorname{ind} D_u^* = n(2-2g) + 2c_1(u^*TM).$$

*(ii) Suppose $\eta \in L^q(\Sigma, \Lambda^{0,1} \otimes_J u^*TM)$ is orthogonal to the image of $D_u$ on $W^{k,p}(\Sigma, u^*TM)$. Then $\eta \in W^{\ell,p}(\Sigma, \Lambda^{0,1} \otimes_J u^*TM)$ and $\eta \in \ker D_u^*$.*

*(iii) Suppose $\xi \in L^q(\Sigma, u^*TM)$ is orthogonal to the image of $D_u^*$ acting on $W^{k,p}(\Sigma, \Lambda^{0,1} \otimes_J u^*TM)$. Then $\xi \in W^{\ell,p}(\Sigma, u^*TM)$ and $\xi \in \ker D_u$.*

This result follows from elliptic regularity, functional analysis, and using Riemann–Roch to compute the index of $D_u$. Let us expound a bit on the index computation. Since $D_u$ is a compact perturbation of a Cauchy–Riemann operator, it suffices to compute the index of the anti-holomorphic differential $\overline{\partial}$. Recall the classical Riemann–Roch theorem says (in terms of real dimensions) for a holomorphic line bundle $L$ on a genus $g$ surface,

$$\dim H^0(\Sigma, L) - \dim H^1(\Sigma, L) = \deg(L) + 1 - g.$$

By the Dolbeault theorem, the above cohomology groups come from the short chain complex,

$$0 \to \Gamma(\Sigma, L) \xrightarrow{\overline{\partial}_L} \Omega^{0,1}(\Sigma, L) \to 0.$$

That is, they are the kernel and cokernel of $\overline{\partial}$. So the left side above is half of $\operatorname{ind}(\overline{\partial}_L)$ (since it counts complex dimensions). The degree of $L$ is just $\langle c_1(L), [\Sigma] \rangle$. We can apply this to a direct sum of line bundles, and then a general bundle by the splitting principle. Considering the case of the bundle $u^*TM$, we reobtain,

$$\operatorname{ind}(\overline{\partial}) = n(2-2g) + 2\langle c_1(u^*TM), [\Sigma] \rangle.$$

We are now prepared to prove the two regularity theorems above by showing that the intersection of the Cauchy–Riemann operator section $\overline{\partial}_J(u)$ with the zero section of the Banach bundle $\mathcal{E} \to \mathcal{B}$ is transverse.

We can consider the *universal moduli space*,

$$\mathcal{M}^*(A, \Sigma; \mathcal{J}^\ell) := \{(u, J) : J \in \mathcal{J}^\ell, u \in \mathcal{M}^*(A, \Sigma; J)\} \subset \mathcal{B}^{\ell,p} \times \mathcal{J}^\ell.$$

The space $\mathcal{J}^\ell$ is a smooth separable Banach manifold with tangent space $T_J\mathcal{J}^\ell$, for $\omega$-compatible structures, given by $\Gamma(\operatorname{End}(TM, J, \omega))$, where $\operatorname{End}(TM, J, \omega)$ has fibre at $x$ given by self-adjoint endomorphisms of $T_xM$ which anticommute

with $J_x$. The space of smooth structures, which we denote $\mathcal{J}$, is modelled locally only by a Fréchet and not Banach space.

PROPOSITION 7.24. *Let $A \in H_2(M;\mathbb{Z})$, $\ell \geq 2$, $p > 2$ and $k \in \{1,\dots,\ell\}$. Then $\mathcal{M}^*(A,\Sigma;\mathcal{J}^\ell)$ is a separable $C^{\ell-k}$ Banach submanifold of $\mathcal{B}^{k,p} \times \mathcal{J}^\ell$.*

PROOF. We have a bundle $\mathcal{E}^{k-1,p} \to \mathcal{B}^{k,p} \times \mathcal{J}^\ell$ with fibre over $(u,J)$ given by $W^{k-1,p}(\Sigma, \Lambda^{0,1} \otimes_J u^*TM)$. This bundle is a $C^{\ell-k}$-Banach bundle. To see this, we can trivialize the bundle over a coordinate chart of $\mathcal{B}^{k,p}$ using a Hermitian connection (which will be of class $C^{\ell-1}$ since $J$ has class $C^\ell$). We may then differentiate $\ell - k$ times and still lie in $W^{k-1,p}$.

Now $\mathcal{F} : (u,J) \mapsto \overline{\partial}_J u$ defines a $C^{\ell-k}$ section of this bundle. We wish to show that the vertical component $D\mathcal{F}$ of the differential is surjective whenever $\mathcal{F}$ intersects the zero section at a point where $u$ is simple. Recall we have,

$$D\mathcal{F}_{(u,J)}(\xi, Y) = D_u\xi + \frac{1}{2}Y(u) \circ \mathrm{d}u \circ j.$$

We know $D_u$ is Fredholm, so that $D\mathcal{F}$ has closed image and surjectivity will follow from having dense image.

First consider $k = 1$. Suppose $\eta \in L^q(\Lambda^{0,1} \otimes_J u^*TM)$ for $q = p^*$ the Hölder conjugate is orthogonal to the image of $D\mathcal{F}$. Hence,

$$\int_\Sigma \langle \eta, D_u\xi \rangle \,\mathrm{dvol} = \int_\Sigma \langle \eta, Y(u) \circ \mathrm{d}u \circ j \rangle \,\mathrm{dvol} = 0, \tag{7.2}$$

for $\eta \in W^{1,p}(\Sigma, u^*TM)$ and $Y \in C^\ell(M, \mathrm{End}(TM, J, \omega))$. We conclude from our Fredholm differential result that $\eta$ is of class $W^{1,p}$ and in the kernel of $D_u^*$.

Let $z$ be an injective point of $u$ (since $u$ is simple). Suppose $\eta(z) \neq 0$. One can explicitly construct $Y_0 \in \mathrm{End}(T_{u(z)}M, J_{u(z)}, \omega_{u(z)})$ so that $\langle \eta(z), Y_0 \circ \mathrm{d}u\,(z) \circ j \rangle$ is positive (see [**McS2**, Lemma 3.2.2]). We can extend $Y_0$ to a $C^\ell$ section $Y$, so that $\langle \eta, Y(u) \circ \mathrm{d}u \circ j \rangle$ is positive on a neighbourhood of $z$. Multiplying $Y$ by a bump function, we would obtain a contradiction to (7.2), so we must have $\eta(z) = 0$. But the set of injective points is open and dense and hence by continuity $\eta = 0$. We conclude $D\mathcal{F}$ has dense image and so is surjective when $k = 1$.

Now we consider $k > 1$. Suppose $\eta \in W^{k-1,p}(\Lambda^{0,1} \otimes_J u^*TM)$. By the previous case, there is $(\xi, Y) \in W^{1,p}(\Sigma, u^*TM) \times C^\ell(M, \mathrm{End}(TM, J, \omega))$ with, $D\mathcal{F}_{(u,J)}(\xi, Y) = \eta$. But then,

$$D_u\xi = \eta - \frac{1}{2}Y(u) \circ \mathrm{d}u \circ j \in W^{k-1,p}.$$

By elliptic regularity, $\xi \in W^{k,p}$. And hence we obtain surjectivity for $k > 1$ too.

Because $D_u$ is Fredholm, surjectivity implies $D\mathcal{F}_{(u,J)}$ has a right inverse and so the implicit function theorem in infinite dimensions means $\mathcal{M}^*(A, \Sigma; \mathcal{J}^\ell)$ is a $C^{\ell-k}$ Banach submanifold; it inherits separability from $\mathcal{B}^{k,p} \times \mathcal{J}^\ell$. □

Now that we have regularity of the universal moduli space, we can obtain the main regularity theorems for a generic almost-complex structure and a generic homotopy.

PROOF. (of Theorem 7.18 (i)) Consider $J \in \mathcal{J}_{\text{reg}}(A, \Sigma)$ and $u \in \mathcal{M}^*(A, \Sigma; J)$, which is smooth by elliptic regularity.

Using a connection as before, we can consider the vertical part of the section $\mathcal{F}(u) = \overline{\partial}_J(u)$ as a map $W^{k,p}(\Sigma, u^*TM) \to W^{k-1,p}(\Sigma, \Lambda^{0,1} \otimes_J u^*TM)$, under exponentiating at $u$, which we denote $\mathcal{F}_u$. The differential of this at zero is precisely $D_u$. A neighbourhood of $u$ in $\mathcal{M}^*(A, \Sigma; J)$ is diffeomorphically identified with a neighbourhood of zero in $\mathcal{F}_u^{-1}(0)$. But since $J$ is regular, $\mathrm{d}\mathcal{F}_u(0) = D_u$ is surjective and so on a small neighbourhood of zero, $\mathcal{F}_u^{-1}(0)$ is a smooth submanifold of dimension $n(2-2g) + 2c_1(A)$, where the dimension follows from Proposition 7.23 (i). Thus locally at $u$, $\mathcal{M}^*(A, \Sigma; J)$ is a smooth manifold of the correct dimension.

We also need to show that the moduli space is oriented. This is a standard construction that appears often in Floer and gauge theory. The tangent space of our moduli space at $u$ is just the kernel of $D_u$. Recall $D_u$ is a zeroth order compact perturbation of a Cacuchy–Riemann operator. One can homotope $D_u$ via a family of Fredholm operators to kill off this compact perturbation. The resulting operator is complex-linear and hence carries a natural orientation on its determinant line bundle. This gives an orientation of $\det D_u = \Lambda^{\text{top}} \ker D_u$, and hence of $\ker D_u$. □

PROOF. (of Theorem 7.18 (ii)) Consider the projection $\pi : \mathcal{M}^*(A, \Sigma; \mathcal{J}^\ell) \to \mathcal{J}^\ell$. For $k = 1$, elliptic regularity means this is a $C^{\ell-1}$ maps of $C^{\ell-1}$ Banach manifolds. Note the tangent space of the domain of this map is pairs $(\xi, Y)$ so that $D_u\xi + \frac{1}{2}Y(u) \circ \mathrm{d}u \circ j = 0$ and $\mathrm{d}\pi_{(u,J)}$ is the projection $(\xi, Y) \mapsto Y$. Hence $\ker \mathrm{d}\pi_{(u,J)} = \ker D_u$. It is a standard lemma of Fredholm theory that $\mathrm{d}\pi_{(u,J)}$ will be Fredholm of the same index as $D_u$ and that $\mathrm{d}\pi_{(u,J)}$ will be surjective precisely when $D_u$ is.

We conclude the set of regular almost complex structures $\mathcal{J}^\ell_{\text{reg}}$ is just the regular values of $\pi$. By the Sard–Smale theorem, this set is second category

for $\ell$ sufficiently large (at least two bigger than the dimension of the moduli space).

Now one can consider the set $\mathcal{J}_{\text{reg},K}(A,\Sigma)$ of smooth almost-complex structures for which $D_u$ is surjective whenever $\|\mathrm{d}u\|_{L^\infty} \leq K$ and there is a point $z \in \Sigma$ with,

$$\inf_{\zeta \neq z} \frac{d(u(z),u(\zeta))}{d(u,\zeta)} \geq \frac{1}{K}.$$

Then $\mathcal{J}_{\text{reg}}$ is the intersection of the sets $\mathcal{J}_{\text{reg},K}(A,\Sigma)$ for all $K \in \mathbb{N}$. Hence it suffices to show $\mathcal{J}_{\text{reg},K}(A,\Sigma)$ is second-category. If $u_i$ is a sequence satisfying these conditions in terms of $K$ but for which $D_{u_i}$ is not surjective, then the same will be true of the limit of some convergent subsequence, hence $\mathcal{J}_{\text{reg},K}(A,\Sigma)$ is open. We can combine the facts $\mathcal{J}^\ell_{\text{reg}}$ is dense in $\mathcal{J}^\ell$ and $\mathcal{J}$ is dense in $\mathcal{J}^\ell$ with the fact the $\mathcal{J}^\ell_{\text{reg},K}(A,\Sigma)$ is open in $\mathcal{J}^\ell$ to conclude density of $\mathcal{J}_{\text{reg},K}(A,\Sigma)$ □

PROOF. (of Theorem 7.20) This proof is nearly identical to above. □

*Unparameterized Moduli Spaces.* In preparation for our analysis of bubbling and compactness, we should note that in the case $\Sigma = S^2$ we cannot expect our moduli space to be compact. This is because there is a free action of the group of biholomorphisms of $S^2$, namely $\mathrm{PSL}(2,\mathbb{C})$, on $\mathcal{M}^*(S^2,A;J)$ and this group is non-compact. To fix this, we will mostly want to talk about compactness for the moduli space of unparameterized curves.

**Definition 7.25**

We define the *unparameterized moduli space of $J$-holomorphic curves* with genus $g$ and $k$ marked points, representing class $A$ to be the space,

$$\mathcal{M}_{g,k}(A;J) := \{(\Sigma,j,u:\Sigma\to M,(z_1,\ldots,z_k))\}/\sim .$$

Here $(\Sigma,j)$ is any genus $g$ Riemann surface, $u$ is $J$-holomorphic representing the homology class $A \in H_2(M;\mathbb{Z})$ and $(z_1,\ldots,z_k)$ is a $k$-tuple of distinct points in $\Sigma$. The equivalence relation $\sim$ will identify $(\Sigma,j,u,(z_1,\ldots,z_k))$ with $(\Sigma',j',u',(w_1,\ldots,w_k))$ if and only if there exists a biholomorphism $\varphi : (\Sigma,j) \to (\Sigma',j')$ so that $u = u' \circ \varphi$ and $\varphi(z_i) = w_i$ for $i = 1,\ldots,k$.

Note that prior to taking a quotient, the moduli space of all genus $g$ curves with $k$ marked points is a fibre bundle over the space of complex structures on a genus $g$ surface with $k$ marked points with fibre over $(\Sigma,j)$ given by $\mathcal{M}_k(\Sigma,A;J)$, where the subscript $k$ denotes the marked points. If we restrict

to simple curves then generically this fibre will be smooth by our previous considerations and so the total space will be a smooth manifold.

Further observe that if we restrict to simple curves, then the group of automorphisms of a Riemann surface must act freely. This implies the above quotient will be free as well. We conclude that $\mathcal{M}^*_{g,k}(A;J)$ is a smooth manifold for generic $J$. If we don't assume our curves are simple, even if $J$ is regular and transversality holds, the moduli space will only be an orbifold with stabilizer group at a multiply covered curve $u : \Sigma \to M$ given by the group of automorphisms of $\Sigma$ preserving $u$.

Let us now find the dimension of the unparameterized moduli space $\mathcal{M}^*_{g,k}(A;J)$ by thinking of how it differs from $\mathcal{M}^*(\Sigma, A; J)$. Since we now allow for an arbitrary genus $g$ Riemann surface, we should increase the dimension according to the space of all complex structures $j_\Sigma$ on a genus $g$ surface $\Sigma$. On the other hand, we now mod out by all curve and marked point preserving biholomorphisms, so we should decrease the dimension according the space of all automorphisms of the Riemann surface $\Sigma$. Finally, we are now keeping track of $k$ points and so we should increase the dimension according to the configuration space of $k$ points in $\Sigma$.

The dimension of the space of $k$ points is clearly $2k$, since $\Sigma$ is two dimensional. Let us analyze the other dimensions in cases. For $\mathbb{C}P^1$, there is a six dimensional group of automorphisms $\mathrm{Aut}(\mathbb{C}P^1) = \mathrm{PSL}(2,\mathbb{C})$ (the Möbius transformations). There is on the other hand, only one complex structure on $S^2$ up to biholomorphism.

For a genus one surface $\mathbb{C}/\Lambda$ (for a lattice $\Lambda = \mathrm{span}_{\mathbb{Z}}(1,\tau)$ with $\tau \in \mathbb{C} \setminus \mathbb{R}$), there is a two dimensional group of automorphisms given by translation in the complex plane (possibly producted with some discrete symmetries if $\tau$ is special). There is also a two dimensional space of complex structures parameterized by the choice of lattice generator $\tau$.

For a surface of genus $g > 1$, there are only finitely many automorphisms. The dimension of the space of complex structures is $6g-6$, which follows from basic Teichmüller theory.

Thus regardless of genus, we see that the difference of the dimension of complex structures on $\Sigma_g$ and the dimension automorphisms of $\Sigma_g$ is given by $6g - 6$.

We conclude,

$$\dim \mathcal{M}^*_{g,k}(A;J) = \dim \mathcal{M}^*(\Sigma_g, A; J) + 6g - 6 + 2k.$$

Or more explicitly using (7.1),

$$\dim \mathcal{M}^*_{g,k}(A; J) = (n-3)(2-2g) + 2c_1(A) + 2k.$$

In general, there is an interesting dichotomy between when the automorphism group of a marked Riemann surface $\Sigma_{g,k}$ is finite and infinite. In the finite case, the number $6g-6+2k$ is the dimension of the moduli space of all surfaces $\Sigma_{g,k}$ (without modding out by automorphisms). In the infinite case, the dimension of the moduli space is larger than this expected formula and the dimension of the automorphism group will be the correction term.

Since the addition of $k$ marked points increases the dimension of the space of surfaces by $\Sigma_{g,k}$ by 2, we see that the automorphism group of $\Sigma_{g,k}$ is finite if and only if $2g + k \geq 3$. In this case, we say that $\Sigma_{g,k}$ is a *stable curve*. Stable curves and stable maps (i.e. $J$-holomorphic curves with finite automorphism group) will be important for Gromov compactification.

As one final note, we should explain why we want to consider marked points. One feature is that the additional structure of marked points in our moduli space allows us to define an *evaluation map*,

$$\mathrm{ev} = (\mathrm{ev}_1, \ldots, \mathrm{ev}_k) : \mathcal{M}^*_{g,k}(A; J) \to M^n \quad \text{where} \quad \mathrm{ev}_i(\Sigma, j, u, z_1, \ldots, z_k) = u(z_i).$$

This is useful for many applications in symplectic geometry, where we might want a curve with some particular intersection property. These maps are also fundamental to the definition of the Gromov–Witten invariants, which essentially count the number of pseudoholomorphic curves $u : \Sigma_{g,k} \to M$ which pass through prescribed generic homological cycles of $M$ at the marked points in $u(z_i)$.

*Multiple Covers and the Failure of Generic Transversality.* In our proof of generic transversality for simple curves the application of somewhere injectivity was slightly subtle. Essentially, recall we found a deformation $Y$ of a almost-complex structure $J$ so that $Y \circ \mathrm{d}u(z) \circ j$ pairs positively with a given deformation of the $J$-holomorphic curve $u$ at the injective point $z$. By injectivity, we can extend $Y$ to a neighbourhood of $z$ and cutoff so that this pairing has positive integral on $\Sigma$ which implies $\eta(z) = 0$. If we remove injectivity, then $\langle \eta(w), Y \circ \mathrm{d}u(w) \circ j \rangle$ may be strictly negative at a different preimage $w$ of $u(z)$. Thus the proof will not work without simplicity of $u$.

This may seem a rather minor obstacle and one might hope a generic transversality result will hold for multiply covered curves. Unfortunately, this cannot be true as we will now demonstrate via some examples. This failure of generic transversality is the source of much of the difficulty of modern Floer-theoretic studies in symplectic geometry.

To understand how generic transversality fails, suppose $u$ is a genus $g$ (unparameterized) simple $J$-holomorphic curve for regular $J$ representing some homology class $A$ in an $n$-manifold $M$. We know that $u$ belongs to the manifold $\mathcal{M}^*_{g,0}(A;J)$. Suppose $v$ is a $k$-fold branched cover of $u$ with ramification index $r$: $v = u \circ \varphi$. That is we have the following diagram,

$$\begin{array}{ccc} (\Sigma'_{g'}, j') & \xrightarrow{\quad \varphi \quad k\text{:1 cover} \quad} & (\Sigma_g, j) \\ & {\scriptstyle v} \searrow \quad \swarrow {\scriptstyle u} & \\ & (M, J). & \end{array}$$

Recall we can relate the Euler characteristics of $\Sigma$ and $\Sigma'$ using the Riemann–Hurwitz theorem,

$$(2 - 2g') = k(2 - 2g) - r.$$

Let $\mathrm{ind}(u)$ denote the index of the appropriate differential operator which cuts out the unparameterized moduli space at $u$. Analogously define $\mathrm{ind}(v)$. From our earlier Riemann–Roch formula,

$$\mathrm{ind}(v) = (n-3)(2-2g') + c_1(v^*TM).$$

From the above Riemann–Hurwitz and the fact $v_*[\Sigma'] = ku_*[\Sigma] = k[A]$,

$$\begin{aligned} &= k(n-3)(2-2g) - r(n-3) + kc_1(A) \\ &= k \cdot \mathrm{ind}(u) - (n-3)r. \end{aligned}$$

Since $u$ is simple and $J$ is regular, $\mathrm{ind}(u)$ is the actual dimension of the moduli space of curves near $u$. If $v$ is also $J$-regular, then the moduli space of (non-simple) curves near $v$ will be an orbifold of dimension $\mathrm{ind}(v)$. On the other hand, the dimension of the moduli space near $v$ must be at least $\mathrm{ind}(u) + 2r$, since each simple curve $\widetilde{u}$ near $u$ gives rise to a $2r$-dimensional family of multiple covers given by moving around the ramification points on the curve $\widetilde{v} = \widetilde{u} \circ \varphi$. Whenever we do not have $\mathrm{ind}(v) \geq \mathrm{ind}(u) + 2r$ for $\mathrm{ind}(u) \geq 0$, we cannot have generic transversality.

Here are a couple specific examples of this phenomenon. Suppose $(M, \omega)$ is a Calabi–Yau 3-fold. I.e. $M$ is a 3-dimensional Kähler manifold with holomorphically trivial canonical bundle.

Suppose $u$ is any simple holomorphic curve for any regular $J$. Then the moduli space near $u$ is smooth with,

$$\dim \mathcal{M}^*_{g,0}(M, \omega, A; J) = \mathrm{ind}(u) = (3-3)(2-2g) + 2c_1(u^*TM) = 0.$$

Where we use that $c_1(TM) = 0$ by the Calabi–Yau assumption. A degree $k$ branched cover $v$ of $u$ with ramification index $r$ then has,

$$\mathrm{ind}(v) = k \cdot \mathrm{ind}(u) - (n-3)r = 0.$$

Because of the dimension of the moduli space, the curve $u$ should persist under generic perturbation. Thus if $v$ were regular, the moduli space should be a non-empty discrete set near $v$. But, near $v$, the moduli space has dimension at least $2r$ as described above. Thus we conclude that $v$ cannot be regular whenever $r > 0$ and hence no branched covers of $J$-holomorphic curves in $M$ can be regular, no matter how we perturb $J$.

Note that it is possible the moduli space near $v$ has dimension strictly greater $2r$, since a sequence of simple covers or differently branched covers may converge to the curve $v$. There is really nothing nice we can say about the moduli space near $v$ in generality.

As another model, suppose $(M, \omega)$ is a symplectic 4-manifold and $S$ is a symplectic sphere with self-intersection number $-1$ (a good example is if $M$ is a blow-up and $S$ is its exceptional divisor). We can always find an $\omega$-compatible $J$ so that $S$ is an embedded $J$-holomorphic sphere. One has,

$$c_1([S]) = c_1(TS) + c_1(\nu_S) = \chi(S) + S \cdot S = 1.$$

Hence, we have a moduli space of dimension,

$$\dim \mathcal{M}^*_{0,0}(M, \omega, [S]; J) = (2-3)(2-0) + 2 = 0.$$

Note by automatic transversality (see below) that $S$ is cut out transversely. Consider a degree $d$ branched cover $S^2 \to S^2$ with ramification degree $2d-2$. Precomposing with $S$ this defines a multiply covered $J$-holomorphic curve $\widetilde{u}$. If this were cut out transversely then the moduli space near $\widetilde{u}$ would have dimension,

$$\mathrm{ind}(\widetilde{u}) = 2d - 2.$$

On the other hand, we know the actual dimension is at least as large as the configuration space of the ramification points, i.e. dimension $4d-4$. This yields a contradiction whenever $d \geq 2$. For the same reason, after perturbing $J$ the resulting multiple covers still cannot be regular and so generic transversality does not hold again.

**7.5. Examples from the Sphere and Automatic Transversality.** Let us think a bit about the moduli space $\mathcal{M}^*(\mathbb{C}P^1, A; J)$. Note that if $u$ is a $J$-holomorphic sphere for $J$ integrable, then $u^*TM$ is a holomorphic vector bundle and hence decomposes as a direct sum of line bundles, each of which is determined by its Chern number/degree.

LEMMA 7.26. *Let $u : \mathbb{C}P^1 \to M$ be a $J$-holomorphic sphere with $J$ integrable. Suppose every line bundle summand of $u^*TM$ has degree at least $-1$. Then $D_u$ is surjective.*

PROOF. Because $J$ is integrable, $D_u$ is just the operator $\overline{\partial}$ in holomorphic coordinates. This will respect the decomposition of $u^*TM$ into line bundles and so it will suffice to consider a single line bundle $L$. Note that $D_u$ on $L$ is the operator,

$$\overline{\partial} : \Omega^0(\mathbb{C}P^1, L) \to \Omega^{0,1}(\mathbb{C}P^1, L),$$

and its cokernel is the Dolbeault cohomology group $H^{0,1}(\mathbb{C}P^1, L)$. We just ask that this group vanish if $c_1(L) \geq -1$. Dualizing, this is equivalent to asking that $H^{1,0}(\mathbb{C}P^1, L^*)$, the space of holomorphic 1-forms on $L^*$, vanish. By Serre duality this is the group $H^0(\mathbb{C}P^1, L^* \otimes K)$, where $K$ is the canonical bundle. A line bundle on $\mathbb{C}P^1$ has global holomorphic sections if and only if it has non-negative degree. Thus $D_u$ is surjective whenever $c_1(L^* \otimes K) < 0$ for each summand $L$. Since $c_1(L^*) = -c_1(L)$ and $c_1(K) = 2$, this is the same as the condition $c_1(L) \geq -1$ for each summand $L$. □

This extends to the almost-complex case.

LEMMA 7.27. *Let $E \to S^2$ be a rank $n$ complex vector bundle and let $D : \Omega^0(S^2, E) \to \Omega^{0,1}(S^2, E)$ be a real Cauchy–Riemann operator. Suppose $E$ splits as a direct sum of complex line bundles $L_1 \oplus \cdots \oplus L_n$ so that $L_1 \oplus \cdots \oplus L_k$ is invariant under $D$ for each $k$. Then $D$ is surjective if and only if $c_1(L_k) \geq -1$ for all $k$.*

LEMMA 7.28 (Automatic Transversality). *Let $J$ be an almost-complex structure on a 4-manifold $M$ and $u : \mathbb{C}P^1 \to M$ an immersed $J$-holomorphic sphere. $D_u$ is surjective if and only if $c_1(u^*TM) \geq 1$.*

PROOF. We have $D_u(\mathrm{d}u \circ \xi) = \mathrm{d}u \circ \overline{\partial}\xi$ for $\xi \in \Gamma(T\mathbb{C}P^1)$ and so $L_0 = \operatorname{im} \mathrm{d}u$ is an invariant line bundle under $D_u$. Let $L_1$ be the orthogonal complement of $L_0$ with respect to an almost complex structure. We must have $c_1(L_0) = c_1(T\mathbb{C}P^1) = 2$. Hence $c_1(L_1) = c_1(u^*TM) - 2$ and the result then follows from the lemma above. □

COROLLARY 7.29. *Let $C$ be an embedded $J$-holomorphic sphere in a 4-manifold $M$ with self-intersection number $C \cdot C = p$. $J$ is regular for class $A = [C]$ if and only if $p \geq -1$.*

PROOF. This follows from the above lemma combined with the adjunction inequality. □

COROLLARY 7.30. *Let $M$ be symplectic and consider the product $S^2 \times M$ with homology class $A = [S^2] \times [\mathrm{pt}]$. Then for every tame almost complex structure $J$ on $M$, the product almost complex structure $i \times J$ is regular for $A$.*

PROOF. Note every $J$-holomorphic curve $u$ for this product structure has to have the form $u(z) = (\varphi(z), x_0)$ where $\varphi$ is a Möbius transformation. This is because, projecting to the factors, $u$ gives a holomorphic map $S^2 \to S^2$ of degree one (i.e. a biholomorphism) and a null-homotopic $J$-holomorphic map $S^2 \to M$ which must be constant since its energy (which is purely topological) is zero.

Now $D_u$ is complex linear and $u^*TM = u^*T_{x_0}M \oplus \varphi^*TS^2$ is a sum of $D_u$ invariant line bundles, where $\varphi^*TS^2$ has Chern number two and the other summands have Chern number zero (since $u^*T_{x_0}M$ is trivial). So by our lemma $D_u$ is surjective. □

REMARK 7.31. Results of the nature we have just proven often go by the name *automatic transversality*. This is because we are able to prove transversality/regularity without appeal to genericity of $J$. The regularity follows automatically from geometric considerations even for arbitrary $J$, or sometimes $J$ of a special form as in the last corollary. △

**7.6. Gromov Compactness and Bubbling.** We now move on to proving some compactness of our unparameterized moduli space. As remarked before, this will add significance to our earlier result that the moduli spaces of $J$-holomorphic curves for different regular $J$ are cobordant. We will say a bit about the ways in which compactness may fail, i.e. a sequence of curves might degenerate. With some energy bounds, the possible degenerations are very mild. This is the phenomena of Gromov compactness and sphere bubbling which is a cornerstone of all of pseudoholomorphic curve and Floer theory. Some results for holomorphic spheres will also be stated for holomorphic disks with Lagrangian boundary conditions, which is important for Lagrangian Floer theory. We will conclude by using compactness in tandem with earlier results to finally prove the celebrated Gromov non-squeezing theorem, which initiated much of the study of holomorphic curves in symplectic topology.

We begin with two important basic theorems which we will not prove. The first result uses elliptic bootstrapping techniques. Recall a submanifold $L$ of $(M, J)$ is *totally real* if $TL \cap J(TL) = \varnothing$ (this is an analogue of an isotropic submanifold in an almost-complex manifold).

**Theorem 7.32: Compactness**

Suppose $(M, J)$ is an almost-complex submanifold with totally real submanifold $L$ and a sequence of almost-complex structures $J_n$ converging in $C^\infty$ to $J$. Let $(\Sigma, j)$ be a Riemann surface exhausted by an increasing

sequence $\Omega_n$ of open subsets and $u_n : (\Omega_n, \Omega_n \cap \partial\Sigma) \to (M, L)$ a sequence of $J_n$-holomorphic curves. Suppose the $L^\infty$ norms of the differential $\mathrm{d}u_n$ are uniformly bounded for any compact subset of $\Sigma$. Then a subsequence of the $u_n$ converges in $C^\infty$ on compact subsets to a $J$-holomorphic curve $u : (\Sigma, \partial\Sigma) \to (M, L)$.

The proof of this is standard elliptic bootstrapping. For the next result, let $B$ be the closed unit disk in $\mathbb{C}$ and $\mathbb{H}$ be the upper half plane.

**Theorem 7.33: Removable Singularities**

Let $(M, \omega)$ be a compact symplectic manifold with compact Lagrangian $L$ and $J$ an $\omega$-tame complex structure on $M$ inducing a metric $g$.

(i) If $u : B \setminus \{0\} \to M$ is $J$-holomorphic with finite energy, then $u$ extends smoothly to a map $B \to M$.

(ii) If $u : (B \cap \mathbb{H} \setminus \{0\}, B \cap \mathbb{R} \setminus \{0\}) \to (M, L)$ is $J$-holomorphic with finite energy, them $u$ extends smoothly to a map $B \cap \mathbb{H} \to M$.

Give a proof sketch. These two results will allow us to understand how bubbling happens. A third result will let us know what bubbles can look like, and that they can only occur at finitely many points.

**Theorem 7.34: Bubble Energy**

Suppose $(M, J)$ is compact and almost-complex with a Riemannian metric and $L$ is a compact totally real submanifold. Then there is a constant $\hbar > 0$ so that, $E(u) \geq \hbar$ for every non-constant $J$-holomorphic sphere $u : S^2 \to M$ and every non-constant holomorphic disk $u : B \to M$ with $u(\partial B) \subset L$.

Proof. We just prove the result for a finite energy holomorphic map $u : \mathbb{C} \to M$. The result extends to the sphere by adding in the point at infinity. The case of a half-plane and a ball is the same.

The energy density $e = |\mathrm{d}u|^2$ satisfies a differential inequality $\Delta e \geq -Be^2$. This implies a mean value inequality,

$$\int_{B_r(z)} |\mathrm{d}u|^2 < \hbar \implies |\mathrm{d}u(z)|^2 \leq \frac{8}{\pi r^2} \int_{B_r(z)} |\mathrm{d}u|^2,$$

for some $\hbar > 0$ and every $r > 0$, $z \in \mathbb{C}$, and choice of $u$. If the energy of $u$ is less than $\hbar$, the left side holds for all $r > 0$ and hence $\mathrm{d}u$ is constantly 0. $\square$

If we have a uniform $L^p$ bound on the first derivatives of the curves $u_n$ for $p > 2$, then we may apply elliptic bootstrapping to obtain $C^\infty$ convergence on compact sets of a subsequence (this slightly generalizes the compactness result, theorem 7.32). However, it is more natural to suppose we only have control on the energy of our curves, i.e. their $L^2$ norms. Because $p = 2$ is the critical exponent in Sobolev embedding, compactness may fail in this case. Nevertheless, Gromov showed the resulting failures of compactness are very mild and only occur at finitely many points where energy localizes to form a holomorphic bubble.

Theorem 7.32 implies that compactness for a sequence of curves $\{u_n\}$ can only fail if $\sup_n \|\mathrm{d}u_n\|_{L^\infty} = \infty$. In this case, if we assume that the energy of the solutions $E(u_n)$ is uniformly bounded, then the only way that compactness can fail is through spheres bubbling off, as we now explore.

Suppose $|\mathrm{d}u_n|$ obtains its maximum at $z_n$, i.e. $|\mathrm{d}u_n(z_n)| = \|\mathrm{d}u_n\|_{L^\infty} =: c_n$. We may pass to a subsequence so that $\{z_n\}$ converges to $z_0$ and so that $\{c_n\}$ diverges to infinity. Let $\varphi : \Omega \to \Sigma$ be a holomorphic coordinate chart with $\varphi(0) = z_0$, and by passing to a subsequence $z_n \in \Omega$. Let $\varphi^* \,\mathrm{dvol} = \lambda^2 \,\mathrm{d}s \wedge \mathrm{d}t$ and we may assume $\lambda(0) = 1$ and $\lambda(z) \in [1/2, 2]$ for all $z$.

We can define sequences,

$$v_n = u_n \circ \varphi \quad \text{and} \quad w_n = \varphi^{-1}(z_n),$$

of holomorphic maps and points in our chart $\Omega$. We have,

$$c_n = \frac{|\mathrm{d}v_n(w_n)|}{\lambda(w_n)} \quad \text{and} \quad \lim_{n\to\infty} w_n = 0.$$

We may find $\varepsilon > 0$ so that the ball $B_\varepsilon(w_n)$ of radius $\varepsilon$ centered at $w_n$ is contained inside $\Omega$ for all $n$. We can then define a reparametrized sequence,

$$r_n : B_{\varepsilon c_n} \to M \quad \text{by} \quad r_n(z) = v_n\left(z_n + \frac{z}{c_n}\right).$$

By construction, $|\mathrm{d}r_n(0)| \geq 1/2$ and $\|\mathrm{d}r_n\|_{L^\infty(B_{\varepsilon c_n})} \leq 2$. Moreover,

$$E(r_n) = E(v_n|_{B_\varepsilon(w_n)}) \leq E(u_n).$$

By the compactness theorem, we may pass to a subsequence so that the $r_n$ converge uniformly in $C^\infty$ on compact sets. The limit $r$ is $J$-holomorphic and satisfies,

$$|\mathrm{d}r(0)| \geq 1/2 \quad \text{and} \quad 0 < E(r) = \int_{\mathbb{C}} r^*\omega \leq \sup_n E(u_n).$$

Using conformal invariance, we conclude that the map $\mathbb{C} \setminus \{0\} \to M$ given by $z \mapsto r(1/z)$ has finite energy and so by the removable singularities theorem

extends to a $J$-holomorphic sphere $\widetilde{r} : \mathbb{C}P^1 \to M$. This is called a *bubble*. The energy $E(\widetilde{r})$ of the bubble is positive and represents an accumulation of the energy of the sequence $u_n$ at $z$,

$$E(\widetilde{r}) \leq \liminf_{n\to\infty} E(u_n, B_\varepsilon(z_0)) \quad \text{for all} \quad \varepsilon > 0.$$

By theorem 7.34, we know $\widetilde{r}$ has energy at least $\hbar$, and hence for large $n$ an energy of at least $\hbar$ is concentrated arbitrarily close to $z_0$. We conclude there can only be finitely many points where $|\mathrm{d}u|$ can tend to infinity and bubbles will form. Away from this finite set of points, theorem 7.33 implies $u$ extends to a $J$-holomorphic sphere. At a bubbling point, more than one bubble may pop out (i.e. there could be a bubble on the bubble). One concludes that the sequence $u_n$ limits to a bubbling tree (a stable curve).

Example 7.35. Here are some basic examples of how blowups may happen.

(1) Consider a degree $d$ rational function $f : \mathbb{C}P^1 \to \mathbb{C}P^1$. The graph $\Gamma_f \subset \mathbb{C}P^1 \times \mathbb{C}P^1$ is a holomorphic curve representing $A = [\mathbb{C}P^1 \times \mathrm{pt}] + d[\mathrm{pt} \times \mathbb{C}P^1]$. Suppose $f(\infty) = 1$. $f$ is determined by its divisor $Z - P$, where $Z, P \subset \mathbb{C}$ are the unordered disjoint lists of the $d$ zeroes and $d$ poles of $f$. Compactness in the space of such holomorphic maps may fail as $Z$ and $P$ approach each other.

Consider $a \in Z$ and $b \in P$. Then $f$ contains a factor of $\dfrac{z-a}{z-b}$, which will be close to one when $a$ is close to $b$. If we consider a family $f(t)$ given by letting the zero $a(t)$ approach $b$, then the graphs $\Gamma_{f(t)}$ will Hausdorff converge to a nodal curve $\Gamma_g \cup b \times \mathbb{C}P^1$, where $g = f(z) \cdot \dfrac{z-b}{z-a}$ and $b \times \mathbb{C}P^1$ is a bubble formed at $b$.

(2) Consider the sequence $u_n : \mathbb{C}P^1 \to \mathbb{C}P^2$ of holomorphic maps,

$$u_n([x : y]) = [x^2 : y^2 : nxy].$$

Letting $[X : Y : Z]$ be coordinates on the codomain, the image of these maps define a sequence of quadrics, $XY = Z^2/n^2$.

When $n \to \infty$, this family degenerates to a pair of hyperplanes: $XY = 0$. These hyperplanes can be interpreted as bubbles.

To see this more explicitly, let us work in an affine chart given by fixing $Z = 1$. In this chart, our maps look like,

$$u_n([x : y]) = \left(\frac{x}{ny}, \frac{y}{nx}\right) \quad \text{for} \quad x, y \neq 0.$$

Hence on this chart, the curves will pointwise converge to a constant curve at $(0,0) = [0:0:1]$ (with uniform $C^\infty$ convergence on compact subsets). Outside of this chart, each curve has two points,

$$u_n([1:0]) = [1:0:0] \quad \text{and} \quad u_n([0:1]) = [0:1:0].$$

Since points close to $[1:0]$ and $[0:1]$ are being sent close to $[0:0:1]$ for large $n$, it is clear that the derivative will blow up at these two points and bubbles will form. To see the bubbles explicitly we may reparametrize,

$$v_n([x:y]) := u_n([nx:y]) \quad \text{and} \quad w_n([x:y]) := u_n([x:ny]).$$

Now on the affine chart we have,

$$v_n([x:y]) = \left(\frac{x}{y}, \frac{y}{n^2 x}\right) \quad \text{and} \quad w_n([x:y]) = \left(\frac{x}{n^2 y}, \frac{y}{x}\right).$$

Hence $v_n$ converge to the curve $v([x:y]) = (x/y, 0)$, i.e. the embedding onto the sphere at infinity given by $Y = 0$. And $w_n$ converge to the curve $w([x:y]) = (0, y/x)$, i.e. the embedding onto the sphere at infinity given by $X = 0$. Strictly speaking these curves are missing the point $[0:0:1]$, but we can fill this in to obtain our two sphere bubbles attached at $[0:0:1]$, which was the image of the constant curve given by the limit of the $u_n$.

As a sanity check, the original curves $u_n$ were degree two. In the limit we have two curves of degree one. Thus we verify a conservation of curve energy, which we prove in theorem 7.38.

(3) Consider the torus $T^2 = \mathbb{C}/\Lambda$, where $\Lambda$ is the lattic of Gaussian integers $\mathbb{Z}[i]$. One can define the Weierstrass $\wp$-function $\wp : T^2 \to S^2$ by,

$$\wp(z) = \sum_{\lambda \in \Lambda} \frac{1}{(z-\lambda)^2}.$$

This function is surjective and $\wp(z_1) = \wp(z_2)$ if and only if $z_1 \pm z_2 \in \Lambda$. Hence $\wp$ has degree two and $c_1(\wp^* TS^2) = 4$. We conclude the moduli space $\mathcal{M}_2$ of degree two meromorphic maps $T^2 \to S^2$ is eight dimensional. It turns out all $u \in \mathcal{M}_2$ have the form $u(z) = \varphi(\wp(z + z_0))$ for some $\varphi \in \mathrm{PSL}(2,\mathbb{C})$ and $z_0 \in \mathbb{C}$. For a complex sequence $a_n$ and a non-zero regular value $\wp(z_0)$, we may define a sequence of curves,

$$u_n(z) = \frac{\wp(z) - \wp(z_0)}{\wp(z) - a_n \wp(z_0)}.$$

Suppose $a_n \to 1$. Then,

$$u_n'(\pm z_0) = \frac{\wp'(\pm z_0)}{\wp(z_0)(1-a_n)},$$

will blow up as $n \to \infty$. Hence degree one holomorphic spheres will bubble off at $\pm z_0$ and $u_n$ will converge in $C^\infty$ away from $\pm z_0$ to the constant curve at 1. ▲

REMARK 7.36. In the last example, we were considering holomorphic tori instead of spheres. For surfaces with non-zero genus, a sequence of curves where the complex structure is allowed to vary may result in more complicated degenerations in its limit. For example, cubics in $\mathbb{C}P^2$ are generically embedded tori. They may however degenerate to different objects including a sphere (rational curve) with a self-intersection or singular point, or a triply covered line, or a union of three lines. △

Using the topological and analytic interaction of energy, we can deduce a compactness result sufficient for the purposes of Gromov non-squeezing. Given a finite energy holomorphic curve $u : \mathbb{C} \to M$, the removable singularities theorem means that $u(1/z) : \mathbb{C} \cong \mathbb{C}P^1 \setminus \{0\} \to M$ is also $J$-holomorphic. We say $u_n : \mathbb{C} \to M$ converge in $C^\infty$ on $\mathbb{C}P^1$ if $u_n(z)$ and $u_n(1/z)$ both converge in $C^\infty$ on compact sets.

**Theorem 7.37: Minimal Energy Spheres**

Let $(M, \omega)$ be compact symplectic and let $J$ be $\omega$-tame. Let $A \in H_2(M; \mathbb{Z})$ be spherical (i.e. in the image of the Hurewicz homomorphism) and suppose $\langle [\omega], A \rangle = \hbar$ for $\hbar$ the maximal possible constant in theorem 7.34 (i.e. $\hbar$ is the minimal energy of a $J$-holomorphic curve). The moduli space $\mathcal{M}_{0,0}(A; J)$ is compact.

PROOF. Consider a $J$-holomorphic sequence $u_n : \mathbb{C}P^1 \to M$ representing $A$. The theorem means that a subsequence should converge in $C^\infty$ after applying elements $\varphi_n \in PSL(2, \mathbb{C})$. By reparameterizing using elements of $SO(3)$, we may assume $|\mathrm{d}u_n(0)|_{\mathrm{FS}} = \sup_{z \in \mathbb{C}} |\mathrm{d}u_n(z)|_{\mathrm{FS}}$, where FS denotes the Fubini–Study metric. Explicitly,

$$|\mathrm{d}u_n(z)|_{\mathrm{FS}} = |\mathrm{d}u_n(z)|(1 + |z|^2),$$

and so $|\mathrm{d}u(z)|$ is also maximized at $z = 0$. Applying a complex rescaling around zero, we may assume $|\mathrm{d}u_n(0)| = 1$.

Applying theorem 7.32, we may restrict to a subsequence so that $u_n$ converges in $C^\infty$ on compact sets to a non-constant $J$-holomorphic curve $u : \mathbb{C} \to M$.

We can now define a sequence $v_n(z) = u_n(1/z)$ which converge in $C^\infty$ on compact subsets of $\mathbb{C}P^1 \setminus \{0\}$. In fact the $v_n$ also converge uniformly near zero. If not, we would obtain a bubble with energy at least $\hbar$ by theorem 7.34, and so the energy of $u$ would have to be zero, violating the fact it is non-constant. Thus $u_n$ converge in $C^\infty$ to $u$ on all of $\mathbb{C}P^1$. □

We now study the limit behaviour of bubbles more closely. Let $(\Sigma, j)$ be a Riemann surface, $(M, \omega)$ a compact symplectic manifold, $L \subset M$ a compact Lagrangian, $J$ an $\omega$-tame almost-complex structure, and $\hbar$ the maximal constant from theorem 7.34.

**Theorem 7.38: Convergence With Bubbling**

Let $J_n$ be a sequence of $\omega$-tame almost-cmplex structures converging to $J$ in $C^\infty$, $\Omega_n$ a sequence of open sets exhausting $\Sigma$, and $u_n : (\Omega_n, \Omega_n \cap \partial\Sigma) \to (M, L)$ a sequence of $J_n$-holomorphic curves with uniformly bounded energy.

Then, after passing to a subsequence, there is a $J$-holomorphic curve $u : (\Sigma, \partial\Sigma) \to (M, L)$ and a finite set $Z = \{z_1, \dots, z_\ell\} \subset \Sigma$ so that the following hold.

(i) $u_n$ converges to $u$ in $C^\infty$ on compact subsets of $\Sigma \setminus Z$.

(ii) For each $i$ and $\varepsilon > 0$ sufficiently small,

$$m_\varepsilon(z_i) = \lim_{n\to\infty} E(u_n|_{B_\varepsilon(z_i)}),$$

exists, depends continuously on $\varepsilon$, and satisfies,

$$m(z_i) = \lim_{\varepsilon\to 0} m_\varepsilon(z_i) \geq \hbar.$$

(iii) For any compact set of $\Sigma$ whose interior contains $Z$,

$$E(u|_K) + \sum_{i=1}^{\ell} m(z_i) = \lim_{n\to\infty} E(u_n|_K).$$

We now prove this result. We begin with a technical lemma of Hofer which shows up in all compactness arguments.

Lemma 7.39. *Let $(X, d)$ be a metric space and $f : X \to \mathbb{R}^{>0}$ be continuous. Fix $x \in X$ and $\delta > 0$ so that $B_{2\delta}(x)$ is complete. Then there is $\xi \in X$ and $\varepsilon \leq \delta$ so that,*

$$d(x, \xi) < 2\delta, \quad \sup_{B_\varepsilon(\xi)} f \leq 2f(\xi), \quad \text{and} \quad \varepsilon f(\xi) \geq \delta f(x).$$

PROOF. If this lemma were false, we could find a sequence $\{x_k\}$ with,

$$x_0 = x, \quad d(x_k, x_{k+1}) \leq \frac{\delta}{2^k}, \quad \text{and} \quad f(x_{k+1}) > 2f(x_k).$$

This would contradict completeness of the ball. To construct the sequence $x_k$, we work by induction. Take $x_0 = x$. If the lemma fails for $\xi = x_0$ and $\varepsilon = \delta$, then $\sup_{B_\delta(x_0)} f > 2f(x_0)$ and we may find $x_1$ with $d(x_0, x_1) \leq \delta$ and $f(x_0) > 2f(x_0)$. After constructing $x_1, \ldots, x_k$, we can repeat the argument with $\xi = x_k$ and $\varepsilon = \delta/2^k$ to find $x_{k+1}$ with $d(x_k, x_{k+1}) \leq \delta/2^k$ and $f(x_{k+1}) > 2f(x_k)$. □

LEMMA 7.40. *With $M, \omega, L, J_n, J, \hbar$ as in theorem 7.38, let $\Omega$ be an open set in the upper half plane $\mathbb{H}$ and $u_n : (\Omega, \Omega \cap \mathbb{R}) \to (M, L)$ be a sequence of $J_n$ holomorphic curves with uniformly bounded energy. If $|\mathrm{d}u_n(z_n)|$ diverges for some sequence $x_n \to z \in \Omega$, then for every $\varepsilon > 0$,*

$$\liminf_{n\to\infty} E(u_n|_{B_\varepsilon(z_0)\cap\mathbb{H}}) \geq \hbar.$$

*Furthermore, there is a non-constant $J$-holomorphic sphere $v : \mathbb{C}P^1 \to M$.*

PROOF. Consider $\varepsilon > 0$ so $B_\varepsilon(z_0) \cap \mathbb{H} \subset \Omega$. By the above lemma applied to the functions $z \mapsto |\mathrm{d}u_n(z)|$ on $B_\varepsilon(z_0) \cap \mathbb{H}$, the points $x_n = z_n$, and the constants $\delta_n = |\mathrm{d}u_n(z_n)|^{-1/2}$, we obtain sequences $\zeta_n \in \mathbb{C}$ and $\varepsilon_n > 0$ so that,

$$\zeta_n \to z_0, \quad \sup_{B_{\varepsilon_n}(\zeta_n)} |\mathrm{d}u_n| \leq 2c_n, \quad \varepsilon_n \to 0, \quad \text{and} \quad \varepsilon_n c_n \to 0,$$

for $c_n = |\mathrm{d}u_n(\zeta_n)|$. If $c_n \operatorname{Im} \zeta_n$ is unbounded, then a $J$-holomorphic sphere bubbles off. Otherwise, a $J$-holomorphic disk bubbles off. We will consider only the first case and the second is analogous.

After passing to a subsequence $c_n \operatorname{Im} \zeta_n \to \infty$. We can assume, shrinking $\varepsilon_n$, that $\varepsilon_n \leq \operatorname{Im} \zeta_n$ and $\varepsilon_n + |\zeta_m - z_0| \leq \varepsilon$ so $B_{\varepsilon_n}(\zeta_n) \subset \Omega$ for each $n$. Consider the sequence,

$$v_n(z) = u_n\left(\zeta_n + \frac{z}{c_n}\right).$$

These are defined on $B_{\varepsilon_n c_n}$ with $\sup_{B_{\varepsilon_n c_n}} |\mathrm{d}v_n| \leq 2$ and $|\mathrm{d}v_n(0)| = 1$. Moreover the energy of $v_n$ on $B_R$ for $R \leq \varepsilon_n c_n$ is the same as the energy of $u_n$ on $B_{R/c_n}(\zeta_n)$.

We conclude that the sequence $\{v_n\}$ has uniformly bounded derivatives on arbitrarily large domains (since $\varepsilon_n c_n \to \infty$). By theorem 7.32, $v_n$ converges in $C^\infty$ on compact subsets of $\mathbb{C}$ to a $J$-holomorphic curve $v : \mathbb{C} \to M$. The curve $v$ has finite energy and is non-constant with $|\mathrm{d}v(0)| = 1$. Thus $z \mapsto v(1/z)$

is a finite energy curve on $\mathbb{C} \setminus \{0\}$ and by theorem 7.33, it extends to a $J$-holomorphic sphere $v : \mathbb{C}P^1 \to M$. We know by theorem 7.34 that $E(v) \geq \hbar$ and so for all $\alpha \in (0,1)$ there is $R > 0$ with $E(v|_{B_R}) > \alpha\hbar$. Thus,

$$\alpha\hbar < E(v|_{B_R}) = \lim_{n\to\infty} E(v_n|_{B_R}) = \lim_{n\to\infty} E(u_n|_{B_{R/c_n}(\zeta_n)}) \leq \lim_{n\to\infty} E(u_n|_{B_\varepsilon(z_0)}).$$

Sending $\alpha \to 1$ yields our result. □

PROOF. (of Theorem 7.38) We say a point $z$ in $\Sigma$ is singular if these is a sequence $z_n \to z$ so that $|\mathrm{d}u_n(z_n)| \to \infty$. This means a bubble will form at $z$. By the previous lemma, if a sequence of $u_n$ of $J_n$-holomorphic curves has energy uniformly bounded by $c$, there can be at most $c/\hbar$ singular points.

We claim that after passing to a subsequence of the $u_n$, for every compact set $K$ contained in the complement of the finite set point set $Z$, the quantity $\|\mathrm{d}u_n\|_{L^\infty(K)}$ is uniformly bounded in $n$. We do this inductively by adding new singular points. Suppose we have a list $Z_k = \{z_1, \ldots, z_k\}$ of singular points. We may suppose there is a compact set $K \subset \Sigma \setminus Z_k$ so that the sequence $\|\mathrm{d}u_n\|_{L^\infty(K)}$ is unbounded, otherwise the theorem is proved.

We can define a sequence $w_n \in K$ so that $\|\mathrm{d}u_n\|_{L^\infty(K)} = |\mathrm{d}u_n(w_n)|$. Passing to a subsequence, the $w_n$ converge to $z_{k+1} \in \Sigma \setminus Z_k$ which is singular. Now we set $Z_{k+1} = Z_k \cup \{z_{k+1}\}$. After finitely many iterations, we must be done, since there are only finitely many singular points in $\Sigma$.

By theorem 7.32, passing to a subsequence, $u_n$ converge in $C^\infty$ on compact subsets of $\Sigma \setminus Z$ to a $J$-holomorphic curve $u : \Sigma \setminus Z \to M$ with $u(\partial\Sigma) \subset L$. By removable singularities theorem 7.33, $u$ extends to a $J$-holomorphic map $u : (\Sigma, \partial\Sigma) \to (M, L)$.

Consider disjoint balls $B_{\varepsilon_j}(z_j)$ around our singular points and pass to a subsequence so that the following limits exist,

$$m_{\varepsilon_j}(z_j) := \lim_{n\to\infty} E(u_n|_{B_{\varepsilon_j}(z_j)}).$$

Since $u_n$ converges on the lattice $B_{\varepsilon_j}(z_j) \setminus B_\varepsilon(z_j)$, the function $m_\varepsilon$ is defined for $\varepsilon \leq \varepsilon_j$ and depends continuously on $\varepsilon$. We must have $m(z_j) = \lim_{\varepsilon\to 0} m_\varepsilon(z_j) \geq \hbar$ by theorem 7.34.

To prove (iii), fix $\varepsilon$ smaller than all $\varepsilon_j$ and we compute,

$$E(u|_{K\setminus\bigcup B_\varepsilon(z_j)}) = \lim_{n\to\infty} E(u_n|_K) - \sum_{j=1}^{\ell} \lim_{n\to\infty} E(u_n|_{B_\varepsilon(z_j)})$$

$$= \lim_{n\to\infty} E(u_n|_K) - \sum_{j=1}^{\ell} m_\varepsilon(z_j).$$

Sending $\varepsilon \to 0$,

$$E(u|_K) = \lim_{n\to\infty} E(u_n|_K) - \sum_{j=1}^{\ell} m(z_j),$$

as desired. □

Our analysis of bubbling can be refined to the following pair of results. For proofs, see [**McS2**, §4.7].

PROPOSITION 7.41. *Let $J_n$ be a sequence of $\omega$-tame almost complex structures on $M$ converging in $C^\infty$ to $J$. Fix $z_0 \in \mathbb{C}$ and $r > 0$. Consider a sequence $u_n : B_r(z_0) \to M$ of $J_n$-holomorphic curves and $u : B_r(z_0) \to M$ a $J$-holomorphic curves such that:*

*(a) The sequence $u_n$ converges in $C^\infty$ to $u$ on compact subsets of $B_r(z_0) \setminus \{z_0\}$.*

*(b) $m_0 = \lim_{\varepsilon\to 0} \lim_{n\to\infty} E(u_n|_{B_\varepsilon(z_0)})$ exists and is positive.*

*Then, on a subsequence, there are Möbius transformations $\psi_n \in \mathrm{PSL}(2,\mathbb{C})$, a $J$-holomorphic sphere $v : \mathbb{C}P^1 \to M$, and finitely many distinct points $z_1, \ldots, z_\ell, z_\infty \in \mathbb{C}P^1$ so that:*

*(i) The sequence $\psi_n$ converges in $C^\infty$ to the constant function $z_0$ on compact subsets of $\mathbb{C}P^1 \setminus \{z_\infty\}$.*

*(ii) The sequence $v_n = u_n \circ \psi_n$ converges in $C^\infty$ to $v$ on compact subsets of $\mathbb{C}P^1 \setminus \{z_1, \ldots, z_\ell, z_\infty\}$. The limits $m_j = \lim_{\varepsilon\to 0} \lim_{n\to\infty} E(v_n|_{B_\varepsilon(z_j)})$ exist and are positive for $j = 1, \ldots, \ell$.*

*(iii) One has $E(v) + \sum_{j=1}^{\ell} m_j = m_0$.*

*(iv) If $v$ is constant, than $\ell \geq 2$.*

PROPOSITION 7.42. *Let $J_n$ be a sequence of $\omega$-tame almost complex structures on $M$, $z_0 \in \mathbb{C}$. Let $u, u_n, \psi_n, v, \{z_1, \ldots, z_\ell, z_\infty\}$ be as in the preceding proposition. Then $u(z_0) = v(z_\infty)$. And for every $\varepsilon > 0$ there is $\delta > 0$ and $n_0 \in \mathbb{N}$ so*

*that whenever* $n \geq n_0$ *and* $z \in \mathbb{C}P^1$,

$$d(z, z_0) + d(\psi_n^{-1} z, z_\infty) < \delta \implies d(u_n(z), u(z_0)) < \varepsilon.$$

**7.7. Gromov Nonsqueezing and Results in Symplectic Topology.** We now explore a few applications of our pseudoholomorphic curve theory to symplectic topology. Our first and primary objective is to give a complete proof of the Gromov non-squeezing theorem. We recall its statement.

**Theorem 7.43: Gromov Non-Squeezing**

If $\varphi : B^{2n}(r) \hookrightarrow B^2(R) \times \mathbb{R}^{2n-2}$ is a symplectic embedding, then $r \leq R$.

Proof. Suppose we have such an embedding $\varphi$. The image of $\varphi$ is compact and so by symplectic embedding we may consider $\varphi$ as mapping into $M := S^2 \times T^{2n-2}$ where the area of $S^2$ is $\pi R^2 + \varepsilon$ for $\varepsilon$ is small, $T^{2n-2} = \mathbb{R}^{2n-2}/L\mathbb{Z}^{2n-2}$ for $L$ large, and we use the standard product symplectic structure $\omega$.

Let $j$ be the standard complex structure on $B^{2n}(r)$. Pushing forward, this induces an almost-complex structure on $\varphi(B^{2n}(r))$ which we can extend generically to an almost-complex structure $J$ on $M$. Assume for now that $J \in \mathcal{J}_{\text{reg}}$ is regular.

Now we consider the moduli space $\mathcal{M}^*_{0,1}(A; J)$, where $A = [S^2 \times \text{pt}]$. Note all curves are automatically simple by the choice of $A$. Since $J$ is regular, this space is a smooth manifold of dimension given by,

$$2(n-3) + 2c_1(A) + 2 = 2n.$$

We know $\mathcal{M}^*_{0,0}(A; J)$ is compact modulo bubbling, but no bubbles can occur since $A$ is the spherical homology class of minimal positive energy (see theorem 7.37). Then $\mathcal{M}^*_{0,1}(A; J)$ is an $S^2$ bundle over $\mathcal{M}^*_{0,0}(A; J)$ and hence is a compact $2n$-manifold . It admits a natural evaluation map at the marked point $\text{ev} : \mathcal{M}^*_{0,1}(A; J) \to M$, which is a smooth map of $2n$-manifolds. We claim this map has degree one and in particular is surjective.

To see this, consider an almost-complex structure $J_0$ on $M$ which is the product of the standard complex structure on $S^2$ and any tame almost-complex structure on $T^{2n-2}$. By corollary 7.30, $J_0$ is regular for our homology class $A$. Moreover from the proof, the $J_0$-holomorphic curves $u$ representing class $A$ are precisely those of the form $u(z) = (\varphi(z), x_0)$ for $\varphi$ a biholomorphism of $\mathbb{C}P^1$ and $x_0 \in T^{2n-2}$ constant. If we mod out by $\text{PSL}(2, \mathbb{C})$ and consider unparameterized curves, then curves just correspond to points of $S^2 \times T^{2n-2}$, given by identifying the marked point in $S^2$ and $x_0 \in T^{2n-2}$. That is, the map $\text{ev} : \mathcal{M}^*_{0,1}(A; J_0) \to M$ is a diffeomorphism.

Now take a generic regular homotopy $\{J_t\}$ of compatible almost-complex structures between $J_0$ and $J_1 = J$. This defines a moduli space $\mathcal{M}^*_{0,1}(A; J_t)$, which will be a $(2n+1)$ manifold. It must also be compact, since it is a submersion over $[0,1]$ with compact fibres.

Thus this moduli space defines an oriented cobordism between $-\mathcal{M}^*_{0,1}(A; J_0)$ and $\mathcal{M}^*_{0,1}(A; J)$. The evaluation map extends to a map,

$$\text{ev} : \mathcal{M}^*_{0,1}(A; J_t) \to M,$$

agreeing with our other evaluation maps on the boundary, which we now call $\text{ev}_0$ and $\text{ev}_1$. The map ev provides an oriented cobordism between the maps $\text{ev}_0$ and $\text{ev}_1$. In particular, the degrees agree. Since $\text{ev}_0$ is a diffeomorphism, $(\text{ev}_0)_*[\mathcal{M}^*_{0,1}(A; J_0)] = [M]$. Hence, we also have $(\text{ev}_1)_*[\mathcal{M}^*_{0,1}(A; J_1)] = [M]$ and $\text{ev}_1$ is degree one and surjective. We conclude that for any point of $x \in M$, there is a $J$-holomorphic curve $S^2 \to M$ passing through $x$.

Recall we assumed that $J$ was regular. If it is not regular, we can find a sequence of regular almost-complex structures $J_i$ converging to $J$ and a sequence $u_i$ of $J_i$-holomorphic spheres representing class $A$ and passing through the point $x \in M$. Since there is no bubbling, Gromov compactness tells us $u_i$ converges to a $J$-holomorphic sphere $u$ which must also pass through $x$. Thus the above conclusion holds for any, even non regular, almost-complex structure.

Returning to our embedding, we know by above there is a $J$-holomorphic curve $u : S^2 \to M$ which intersects $\varphi(0)$, the image of the centre of the ball under the embedding. We can now define a map defined on some subset of $S^2$,

$$v = \varphi^{-1} \circ u : u^{-1}(\varphi(B^{2n}(r))) \to B^{2n}(r).$$

The image of this map contains $0 \in B^{2n}(r)$ and since $\varphi^* J = j$, this map is $j$-holomorphic. See figure 13 for an illustration of the maps involved.

Now we use the following theorem coming from a general result minimal surface theory (since a $j$-holomorphic curve is minimal in the $\omega$-compatible case).

**Theorem 7.44: Monotonicity Formula**

Suppose $(\Sigma, j_\Sigma)$ is a Riemann surface and $u : (\Sigma, j_\Sigma) \to (B^{2n}(r_0), j)$ is a non-constant proper holomorphic map whose image contains 0. Let $\Sigma_r = u^{-1}(B^{2n}(r))$. Then for $r < r_0$,

$$\pi r^2 \leq \int_{\Sigma_r} u^* \omega_{\text{std}}.$$

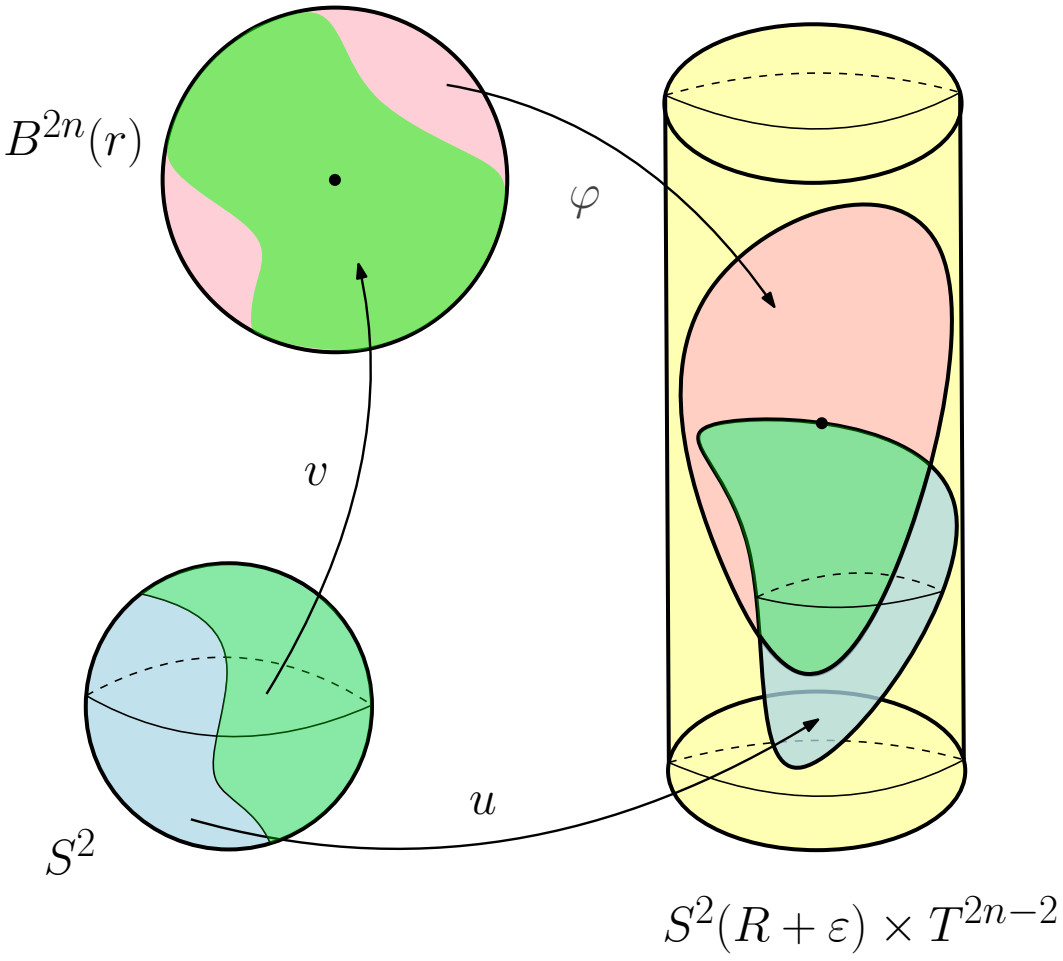



FIGURE 13. Maps in the proof of Gromov non-squeezing.

PROOF. When $r$ is some sufficiently small value $\varepsilon$, Micallef–White tells us that $u$ looks like a perturbation of $z \mapsto (z^k, 0)$ in certain local coordinates. Hence the area here will be $k\pi\varepsilon^2$. We now claim the function,

$$F(r) = \frac{1}{r^2} \int_{\Sigma_r} u^* \omega_{\text{std}},$$

is non-decreasing. This will then imply our monotonicity formula.

To do this, we just need to show $F(r) < F(R)$ whenever $r < R$ are regular in the sense that $u$ is transverse to the the boundary of the ball of radius $r$ or $R$ and hence $\Sigma_r$ and $\Sigma_R$ are smooth. This is sufficient because such regular values are dense.

Consider the symplectomorphism,

$$\Phi : (\mathbb{R} \times S^{2n-1}, \mathrm{d}(e^t \lambda_{\mathrm{std}})) \to (\mathbb{R}^{2n} \setminus \{0\}, \omega_{\mathrm{std}}) \quad \text{where} \quad \Phi(t, m) = e^{t/2} m.$$

We can equip $\mathbb{R}^{2n} \setminus \{0\}$ with the standard complex structure $j$ and pull back by $\Phi$ to obtain $\Phi^* j = \widehat{j}$ and make $\Phi$ a biholomorphism. We may remove from $\Sigma$ the points where $u$ vanishes to obtain a punctured surface $\dot{\Sigma}$. We can then define a holomorphic map,

$$(u_{\mathbb{R}}, u_S) = \Phi^{-1} \circ u : \dot{\Sigma} \to \mathbb{R} \times S^{2n-1}.$$

For $r = e^{\tau/2}$ regular,

$$\begin{aligned} F(r) &= \frac{1}{r^2} \int_{\Sigma_r} u^* \omega_{\mathrm{std}} \\ &= e^{-\tau} \int_{\dot{\Sigma}_r} (u_{\mathbb{R}}, u_s)^* \, \mathrm{d}(e^t \lambda_{\mathrm{std}}) \\ &= e^{-\tau} \int_{\partial \Sigma_r} (u_{\mathbb{R}}, u_s)^* (e^t \lambda_{\mathrm{std}}) \\ &= \int_{\partial \Sigma_r} u_S^* \lambda_{\mathrm{std}}. \end{aligned}$$

Hence,

$$\begin{aligned} F(R) - F(r) &= \int_{\partial \Sigma_R} u_S^* \lambda_{\mathrm{std}} - \int_{\partial \Sigma_r} u_S^* \lambda_{\mathrm{std}} \\ &= \int_{\Sigma_R \setminus \Sigma_r} u_S^* \mathrm{d} \lambda_{\mathrm{std}} \\ &= \int_{\Sigma_R \setminus \Sigma_r} u_S^* \mathrm{d} \lambda_{\mathrm{std}}. \end{aligned}$$

In local coordinates $(s, t)$ on the Riemann surface,

$$= \int_{\Sigma_R \setminus \Sigma_r} \mathrm{d}\lambda_{\mathrm{std}}(\mathrm{d}\Phi^{-1}(\partial_s u_S), \mathrm{d}\Phi^{-1} \partial_t (u_S)) \, \mathrm{d}s \wedge \mathrm{d}t \, .$$

Let $V = \mathrm{d}\Phi^{-1}(\partial_s u_S)$,

$$= \int_{\Sigma_R \setminus \Sigma_r} \mathrm{d}\lambda_{\mathrm{std}}(V, \widehat{j} \, V) \, \mathrm{d}s \wedge \mathrm{d}t \, .$$

Since $\widehat{j}$ is $\mathrm{d}\lambda_{\mathrm{std}}$-compatible,

$$> 0.$$

Thus we obtain our result. □

Now we may finish the proof of Gromov non-squeezing. Using the monotonicity formula we have,

$$\pi(r-\varepsilon)^2 \leq \int_{u^{-1}(\varphi(B^{2n}(r')))} v^*\omega_{\text{std}}.$$

Recalling $\omega$ was the (standard) symplectic form on $S^2 \times T^{2n-2}$,

$$\begin{aligned} &= \int_{u^{-1}(\varphi(B^{2n}(r')))} u^*\omega \\ &\leq \int_{S^2} u^*\omega \\ &= \langle [\omega], [S^2 \times \text{pt}] \rangle. \end{aligned}$$

This just gives the symplectic area of $S^2$.

$$= \pi R^2 + \varepsilon.$$

Taking $\varepsilon \to 0$ we deduce $r \leq R$ as desired. □

We now discuss some work of Gromov concerning symplectomorphism groups and symplectic rigidity.

**Theorem 7.45: Gromov**

The symplectomorphism group of $S^2 \times S^2$, with the product of the Fubini–Study forms on each, has two connected components. The identity component is homotopy equivalent to $SO(3) \times SO(3)$.

Proof. (Sketch) Let $\text{Symp}_0$ denote the component of the symplectomorphism group inducing the identity on $H_2(M;\mathbb{Z})$. Note $\text{Symp} \,/\, \text{Symp}_0$ has elements since we can symplectically interchange the spheres. Hence we need to show $\text{Symp}_0$ is homotopy equivalent to $S^2 \times S^2$.

Let $\mathcal{J}$ be the set of compatible almost-complex structures. There is an action $\text{Symp}_0 \times \mathcal{J} \to \mathcal{J}$ by $(\psi, J) \mapsto \psi_* J$. The stabilizer of the standard complex structure $J_0$ under this action is $SO(3) \times SO(3)$, i.e. these are the Kähler isometries. Hence $\text{Symp}_0 / SO(3) \times SO(3)$ is the orbit of $J_0$ under our action.

We wish to show the map $\text{Symp}_0 / SO(3) \times SO(3) \to \mathcal{J}$ given by $\psi \mapsto \psi_* J_0$ has a left inverse. This implies $\text{Symp}_0 / SO(3) \times SO(3)$ is contractible, and we will be done.

Let $A = [S^2 \times \text{pt}]$ and $B = [\text{pt} \times S^2]$. Consider the universal moduli space,

$$\mathcal{M}_0 = \{(u, v, J) : J \in \mathcal{J}, u \in \mathcal{M}(A; J), v \in \mathcal{M}(B; J), (0,0) \in \operatorname{im} u \cap \operatorname{im} v\}.$$

Let $G = \mathrm{PSL}(2,\mathbb{C})$. There is a free right action on $\mathcal{M}_0$ by $G \times G$. An extension of our automatic transversality theory [**McS2**, Prop. 9.4.4] yields that the $G \times G$-invariant projection $\mathcal{M}_0 \to \mathcal{J}$ is a submersion and the fibres are $G \times G$ orbits. Hence, $\mathcal{M}_0$ is a principal $G \times G$-bundle over $\mathcal{J}$. Since $\mathcal{J}$ is contractible, $\mathcal{M}_0$ is homotopy equivalent to $G \times G$ and hence $SO(3) \times SO(3)$ (since $SO(3)$ is the maximal compact of $G$).

Now we have a commutative diagram,

$$\begin{array}{ccccc} SO(3)\times SO(3) & \longrightarrow & \mathrm{Symp}_0 & \longrightarrow & \mathrm{Symp}_0/SO(3)\times SO(3) \\ \downarrow & & \tau\downarrow & & \downarrow \\ G\times G & \longrightarrow & \mathcal{M}_0 & \longrightarrow & \mathcal{J} \end{array}$$

Here the embedding $\mathcal{T} : \mathrm{Symp}_0 \to \mathcal{M}_0$ is given by,

$$\mathcal{T}(\psi) = (\psi(\cdot, w_*), \psi(z_*, \cdot), \psi_* J_0) \quad \text{where} \quad (z_*, w_*) = \psi^{-1}(0,0).$$

Note $\mathcal{T}$ is $SO(3)\times SO(3)$-equivariant and hence induces a map $\mathrm{Symp}_0 / SO(3)\times SO(3) \to \mathcal{M}_0/SO(3) \times SO(3)$, where the codomain is contractible. Thus $\mathcal{T}$ gives a homotopy equivalence $SO(3) \times SO(3) \to \mathcal{M}_0$.

But $\mathcal{T}$ is injective on $SO(3) \times SO(3)$. It has image,

$$\mathcal{M}_{00} = \{(u, v, J_0) : u(z) = (\varphi(z), 0), v(w) = (0, \phi(w)), \varphi, \phi \in SO(3)\}.$$

Thus $\mathcal{M}_0$ deformation retracts onto $\mathcal{M}_{00}$ via some homotopy $\Psi_t$. If we can construct a left inverse $\mathcal{F} : \mathcal{M}_0 \to \mathrm{Symp}_0$ to $\mathcal{T}$, then $\mathcal{F} \circ \Psi_t \circ \mathcal{T}$ will give a deformation retract of $\mathrm{Symp}_0$ onto $SO(3) \times SO(3)$.

Given $(u, v, J) \in \mathcal{M}_0$, we obtain two maps,

$$S^2 \to \mathcal{M}(A; J)/G : w \mapsto [u_w] \quad \text{and} \quad S^2 \to \mathcal{M}(B; J)/G : z \mapsto [v_z].$$

And we will have $u(z) \in \mathrm{im}[v_z]$ and $v(w) \in \mathrm{im}[u_w]$. Let $\psi_{u,v,J} = \psi : M \to M$ be given by $\psi(z, w) = \mathrm{im}[u_w] \cap \mathrm{im}[v_z]$. This will be a diffeomorphism satisfying $\psi(z, w_*) = u(z)$ and $\psi(z_*, w) = v(w)$ where $z_*, w_* \in S^2$ satisfy $u(z_*) = v(w_*) = (0, 0)$.

We can define a one parameter family of symplectic forms $\omega_t = \omega + t(\psi^*\omega - \omega)$. These are cohomologous since $\psi$ acts trivially on $H^2(M; \mathbb{R})$. By Moser isotopy, we can deform $\psi$ to a symplectomorphism. We thus obtain a path $f_t$ of diffeomorphisms with $f_0 = \mathrm{id}$ and $f_1^* \psi^* \omega = \omega$. Now let $\mathcal{F}(u, v, J) = \psi_{u,v,J} \circ f_1$. If $\psi_{u,v,J}$ is a symplectomorphism, $\mathcal{F}(u, v, J) = \psi_{u,v,J}$ and so $\mathcal{F}$ is a left inverse to $\mathcal{T}$. $\square$

REMARK 7.46. If we deform the product symplectic form on $S^2 \times S^2$ even simply by making one sphere larger than the other, then the homotopy type will change. △

A similar result which we restrain from proving is the following.

**Theorem 7.47: Gromov** [McS2, Thm. 9.5.3]

The symplectomorphism group of $\mathbb{C}P^2$ is homotopy equivalent to $PU(3)$.

From our result on $S^2 \times S^2$, one can also deduce results about symplectomorphisms for $\mathbb{C}^2$. The following is a kind of unique continuation result for symplectomorphisms. It expresses the *symplectic rigidity* of $\mathbb{C}^2$.

**Theorem 7.48: Gromov**

Suppose $X$ is a symplectic 4-manifold with $H_2(X) = 0$. Suppose the complement of a compact set in $X$ is symplectomorphic to the complement of a compact set of $\mathbb{C}^2$. Then $X$ is symplectomorphic to $\mathbb{C}^2$.

To prove this, one uses the compactness to compactify to $S^2 \times S^2$. And then one can apply the above result. The assumption $H_2(X) = 0$ is needed to disallow bubbling. If we drop the assumption, then one could also have blowups of $\mathbb{C}^2$ symplectomorphic to $\mathbb{C}^2$ off a compact set.

We conclude with a result showing symplectic rigidity for $\mathbb{C}P^2$, modulo a deep result of Taubes.

**Theorem 7.49**

Let $(X, \omega)$ be a symplectic 4-manifold with the homology of $\mathbb{C}P^2$ which contains an embedded symplectic 2-sphere with self intersection number one. Then $(X, \omega)$ is symplectomorphic to $\mathbb{C}P^2$ with a multiple of the Fubini–Study form.

Every symplectic structure on $\mathbb{C}P^2$ is symplectomorphic to $(\mathbb{C}P^2, \lambda\omega_{FS})$ for some $\lambda > 0$.

PROOF. Let $C$ be our embedded symplectic 2-sphere and suppose it is in homology class $A$. We know $A$ is primitive and so the complement of a tubular neighbourhood of $X$, call it $W$ is homologically a 4-ball with boundary the total space of a circle bundle over $C$ with Euler class one. Using the symplectic neighbourhood theorem, $\omega$ on this tubular neighbourhood is symplectomorphic to the Fubini–Study form on $\mathbb{C}P^2$ in a neighbourhood of $\mathbb{C}P^1$, which in turn

is isomorphic to the standard symplectic form on $\mathbb{C}^2$ in an annulus around 0. Using the previous result on the rigidity of $\mathbb{C}^2$, there is a symplectomorphism from $W$ to a ball in $\mathbb{C}^2$. Gluing back in the neighbourhood of $C$ must thus give $\mathbb{C}P^2$ with some multiple of the Fubini–Study form (depending on the size of the ball in $\mathbb{C}^2$).

Taubes has shown that every symplectic structure on $\mathbb{C}P^2$ has such a symplectically embedded 2-sphere $C$ and hence the second claim follows. $\square$

CHAPTER 2

# Algebraic Topology

## 1. Basic Theory

We begin with some basic fundamental concepts. In particular we build familiarity with CW complexes, important types of spaces and maps between them, and the basic notion of equivalence in algebraic topology.

**1.1. CW Complexes, Elementary Spaces, and Operations.** Most of the spaces we deal with have the following description which allows for easier calculations and implies nice properties.

**Definition 1.1**

A *CW complex* $X$ is a space formed as the inductive limit of spaces $X_0 \subset X_1 \subset \cdots$, where $X_n$ is obtained from $X_{n-1}$by,

$$X_n = X_{n-1} \bigsqcup_{\alpha \in A, \varphi_\alpha} e_\alpha,$$

where $A$ is an indexing set, $e_\alpha$ are homeomorphic to $D^n$ and we glue $e_\alpha$ to $X_{n-1}$ by continuous maps $\varphi_\alpha : \partial e_\alpha \to X_{n-1}$, so that $X$ satisfies:

**Weak Topology:** $C \subset X$ is closed iff $C \cap X_i$ is closed for all $i$.

We say $X_n$ is the *$n$-skeleton* of $X$ and a choice of decomposition $\{X_n\}$ of $X$ is a *cell structure*. A Hausdorff space is a CW complex iff it can be partitioned into cells $e_\alpha$ homeomorphic to disks so that it satisfies the weak topology axiom and:

**Closure Finiteness:** The boundary of any cell is in the closure of a finite collection of lower dimensional cells.

Now we recall a number of fundamental spaces and describe their CW structure where applicable.

Important Spaces

**Euclidean space:** We have the basic spaces $\mathbb{R}, \mathbb{C}, \mathbb{H}$ and their Cartesian products $\mathbb{R}^n, \mathbb{C}^n, \mathbb{H}^n$. In the direct limit under inclusion, we obtain $\mathbb{R}^\infty, \mathbb{C}^\infty, \mathbb{H}^\infty$. These are homeomorphic but are each interesting for their distinct algebraic structures.

**Spheres:** One has the sphere $S^n$ which can be formed as a CW complex from an $n$-cell glued by its whole boundary to a 0-cell. Alternatively, $S^n$ can be formed inductively by gluing two $n$-cells by the identity to $S^{n-1}$. This second structure allows us to define $S^\infty$ as a direct limit, and a simple lemma proves $S^\infty$ is contractible.

**Projective Spaces:** The space of real (resp. complex or quaternionic lines) in $\mathbb{R}^{n+1}$ (resp. $\mathbb{C}^{n+1}$ or $\mathbb{H}^{n+1}$) defines the projective space $\mathbb{R}P^n$ (resp. $\mathbb{C}P^n$ or $\mathbb{H}P^n$). These spaces can also be formed from $S^n$ (resp. $S^{2n+1}$ or $S^{4n+3}$) by quotienting by $\mathbb{Z}_2$ (resp. $S^1$ or $S^3$) acting by unit scalar multiplication. We can form projective space as a CW structure inductively forming $P^n$ from $P^{n-1}$ by gluing on an $n$-cell (resp. $2n$ or $4n$-cell) via the quotient map just described. We obtain $\mathbb{R}P^\infty, \mathbb{C}P^\infty, \mathbb{H}P^\infty$ as a direct limit.

**Grassmannians:** One has the Grassmannians $G(n,k)$ defined as the space of $k$-dimensional subspaces of $\mathbb{R}^n$ (note $G(n+1,1) = \mathbb{R}P^n$). Analogously we may define complex and quaternionic Grassmannians $\mathbb{C}G(n,k)$, $\mathbb{H}G(n,k)$ and oriented Grassmannians $G_+(n,k)$.

There is a CW structure on $G(n,k)$, the Schubert decomposition, whose cells are in bijection with Young diagrams that fit in a $k \times (n-k)$ rectangle (we won't describe the gluing). The dimension of each cell is the number of squares in the diagram. We can do the same for $\mathbb{C}G(n,k)$ and $\mathbb{H}G(n,k)$ doubling (resp. quadrupling) the cells' dimensions, and for $G_+(n,k)$ doubling the number of cells. In the direct limit one obtains $G(\infty,k)$ and taking another limit, $G(\infty,\infty)$ (analogously for the other versions).

**Stiefel Manifolds:** The Stiefel manifolds $V(n,k)$ (resp. $\mathbb{C}V(n,k)$, $\mathbb{H}V(n,k)$ are the spaces of orthonormal (resp. unitary, symplectic) $k$-frames in $\mathbb{R}^n$ (resp. $\mathbb{C}^n$, $\mathbb{H}^n$). These have Bruhat cell partitions which we won't describe.

**Lie Groups:** Groups which have the structure of a smooth manifold with smooth group multiplication and inversion are called Lie groups. Important examples include $\mathrm{O}(n)$ and $\mathrm{U}(n)$ the orthogonal and unitary groups of $n \times n$ orthogonal and unitary matrices. These are compact analogues of the general linear groups $\mathrm{GL}(n,\mathbb{R})$ and $\mathrm{GL}(n,\mathbb{C})$, the space of invertible real and complex matrices. These groups have "special" subgroups where we restrict the determinant to be one, namely $\mathrm{SO}(n), \mathrm{SU}(n), \mathrm{SL}(n,\mathbb{R})$, and $\mathrm{SL}(n,\mathbb{C})$ respectively. There is also the group $\mathrm{GL}^+(n,\mathbb{R})$ of positive determinant real matrices. One can also define the symplectic groups $\mathrm{Sp}(2n,\mathbb{R}), \mathrm{Sp}(2n;\mathbb{C})$ of real and complex matrices preserving the standard symplectic form. The quaternionic unitary group can be described as,

$$\mathrm{Sp}(n) := \mathrm{Sp}(2n;\mathbb{C}) \cap \mathrm{U}(2n).$$

It's useful to remember the dimension of some of these as manifolds:

$$\begin{aligned} \dim(\mathrm{O}(n)) = \dim(\mathrm{SO}(n)) &= \frac{n(n-1)}{2} \\ \dim(\mathrm{U}(n)) &= n^2 \\ \dim(\mathrm{SU}(n)) &= n^2 - 1 \\ \dim(\mathrm{Sp}(2n;\mathbb{R})) = \dim(\mathrm{Sp}(n)) &= n(2n+1) \\ \dim(\mathrm{Sp}(2n;\mathbb{C})) &= 2n(2n+1). \end{aligned}$$

The tangent space to the identity of a Lie group defines its Lie algebra, there are several of these worth remembering:

$$\begin{aligned} \mathfrak{gl}(n) &= \{n \times n \text{ matrices}\} \\ \mathfrak{sl}(n) &= \{\text{traceless matrices}\} \\ \mathfrak{o}(n) = \mathfrak{so}(n) &= \{\text{skew-symmetric matrices}\} \\ \mathfrak{u}(n) &= \{\text{skew-adjoint matrices}\} \\ \mathfrak{su}(n) &= \{\text{traceless skew-adjoint matrices}\} \\ \mathfrak{sp}(n) &= \{A : JA + A^*J = 0\}. \end{aligned}$$

There are several "accidental isomorphisms" that occur in low dimensions, for example:

$$S^1 \cong \mathrm{SO}(2) \cong \mathrm{U}(1), \quad S^3 \cong \mathrm{SU}(2) \cong \mathrm{Sp}(1) \quad \text{and} \quad \mathrm{Sp}(1,\mathbb{C}) \cong \mathrm{SL}(2,\mathbb{C}).$$

One also has $\mathrm{SO}(4) \cong \mathrm{SO}(3) \times \mathrm{SU}(2)$ but only as manifolds, not as groups. There are also some more general topological descriptions one can obtain. For example, $\mathrm{SU}(3)$ is an $S^3$ fibre bundle over $S^5$.

**Homogeneous Spaces:** Many of the spaces we have just seen can be described as homogeneous spaces, i.e. quotients of compact Lie groups by Lie subgroups. We have,

$$\begin{aligned} S^{n-1} &= \mathrm{O}(n)/\mathrm{O}(n-1) = \mathrm{SO}(n)/\mathrm{SO}(n-1) \\ S^{2n-1} &= \mathrm{U}(n)/\mathrm{U}(n-1) = \mathrm{SU}(n)/\mathrm{SU}(n-1) \\ S^{4n-1} &= \mathrm{Sp}(n)/\mathrm{Sp}(n-1) \\ G(n,k) &= \mathrm{O}(n)/\mathrm{O}(k) \times \mathrm{O}(n-k) \\ V(n,k) &= \mathrm{O}(n)/\mathrm{O}(n-k) \overset{\text{if } n>k}{=} \mathrm{SO}(n)/\mathrm{SO}(n-k). \end{aligned}$$

And similar can be inferred for oriented, complex, or quaternionic analogues.

With many spaces at our disposal, we now consider some basic operations to be performed on spaces.

## Operations

**Products:** We can take the Cartesian product of two spaces $X \times Y$. For infinite products, we use the weak topology.

**Cylinders, Cones and Suspensions:** Given $X$, the cylinder over $X$ is $X \times I$. The cone over $X$ is $CX = X \times I / X \times \{1\}$. The suspension over $X$ is $\Sigma X$ formed as two copies of $CX$ glued along $X \times \{0\}$. Note $\Sigma S^n = S^{n+1}$. If $X$ is a pointed space, we instead consider the reduced suspension also denoted $\Sigma X$, which is the suspension quotient the line $\{x_0\} \times I$. For nice spaces this is homotopy equivalent to the usual suspension.

**Attachings:** Given $f : X \to Y$, the mapping cylinder of $f$ is,

$$M_f = X \times I \bigsqcup_{(x,1)\sim f(x)} Y.$$

The mapping cone of $f$ is,

$$C_f = CX \bigsqcup_{(x,0)\sim f(x)} Y.$$

If $X = Y$, the mapping torus of $f$ is,

$$T_f = X \times S^1/(x,1) \sim (f(x),0).$$

**Joins:** The join of $X$ and $Y$ is given by,

$$X * Y = X \times Y \times I/(x,y_1,0) \sim (x,y_2,0) \quad \text{and} \quad (x_1,y,1) \sim (x_2,y,1).$$

**Mapping Spaces:** Given $X, Y$ the space $C(X,Y)$ of continuous maps $X \to Y$ can be equipped with the compact-open topology, i.e. it has a basis $\mathcal{B} = \{B(K,U)\}$ indexed by compacts sets $K \subset X$ and open sets $U \subset Y$ so that $B(K,U)$ consists of all continuous maps mapping $K$ inside $U$. In particular, we have the path space $C(I,X) = \mathcal{P}X$ and the loop space $C(S^1,X) = \Lambda X$. We often restrict to paths or loops with fixed endpoints to obtain $EX$ and $\Omega X$ respectively.

**Wedge Sum:** Given pointed spaces $X, Y$ we can glue their base points together to get a wedge sum $X \vee Y$.

**Smash Product:** Given pointed spaces, we have the smash product,

$$X \wedge Y = X \times Y/X \vee Y.$$

One has $X \wedge S^k = \Sigma^k X$.

**1.2. Homotopy.** We discuss now the central equivalence notion for the study of algebraic topology.

**Definition 1.2**

Two maps $f, g : X \to Y$ are *homotopic* if there is a map $F : X \times I \to Y$ so that if $F_t := F(\cdot, t)$ then $F_0 = f, F_1 = g$. This is an equivalence relation on $C(X,Y)$. We denote the set of homotopy equivalence class as $[X,Y]$. We can similarly define homotopy relative to a subset $A \subset X$ by demanding $F_t|_A$ is independent of $A$. For the particular case of a point, we have the set of pointed homotopy classes $[(X,x_0),(Y,y_0)]$. When $X$ and $Y$ are pointed we will write $[X,Y]$ to mean the homotopy classes relative to the basepoints.

Two spaces $X$ and $Y$ are *homotopy equivalent* if there are maps $f : X \to Y$, $g : Y \to X$ called *homotopy equivalences* so that $g \circ f$ and $f \circ g$ are homotopic to $\mathrm{id}_X$ and $\mathrm{id}_Y$ respectively.

A space is *contractible* if it is homotopy equivalent to a point.

PROPOSITION 1.3. *The following are equivalent:*

*(i)* $X$ *and* $Y$ *are homotopy equivalent.*

*(ii) For any* $Z$*, there is a natural bijection* $[X, Z] \to [Y, Z]$.

*(iii) For any* $Z$*, there is a natural bijection* $[Z, X] \to [Z, Y]$.

PROOF. $(i) \implies (iii)$ : Suppose $f : X \to Y$ and $g : Y \to X$ provide a homotopy equivalence. Then $f_*$ gives a natural map $[Z, X] \to [Z, Y]$ with inverse $g_*$.

$(iii) \implies (i)$ : Take $Z = X$, giving a bijection $[X, X] \to [X, Y]$. Let $f$ be the image of the identity. Taking $Z = Y$, gives bijection $[Y, X] \to [Y, Y]$. Let $g$ be the inverse image of the identity. Naturality implies $f \circ g$ and $g \circ f$ are homotopic to the identities.

A similar argument gives $(i) \iff (ii)$. □

**Definition 1.4**

We say $X$ and $Y$ are *weak homotopy equivalent* if the conditions $(ii)$ or $(iii)$ in the above proposition hold just for any CW complex $Z$.

REMARK 1.5. In the special case where a map $f : X \to Y$ induces the natural bijection, we will later see this is the same, and more commonly defined, as saying $f$ induces an isomorphism in homotopy or homology. △

**1.3. Cofibrations and Fibrations.** There are two special types of maps in topology that generalize the ideas of embeddings and submersions/fibre bundles from topology.

**Definition 1.6**

A map $i : A \to X$ is a *cofibration* if for every homotopy $F : A \times I \to Y$ so that $F_0$ extends to a map $X \to Y$, the entire homotopy extends. Equivalently any commutative square of the following form admits the following lift

$$\begin{array}{ccc} A & \xrightarrow{F} & \mathcal{P}Y \\ \downarrow{\scriptstyle i} & \nearrow & \downarrow{\scriptstyle \pi_0} \\ X & \longrightarrow & Y. \end{array}$$

If $A \to X$ is a fibration, $X/i(A)$ is called the *homotopy cofibre* of $i$.

A map $\pi : E \to B$ is a *Hurewicz* or a *strong Serre fibration* if for any homotopy $F : X \times I \to B$ so that $F_0$ lifts to a map $X \to E$, the entire

homotopy lifts. Equivalently any commutative square of the following form admits the following lift

$$\begin{array}{ccc} X & \longrightarrow & E \\ \downarrow{\scriptstyle i_0} & \nearrow & \downarrow{\scriptstyle \pi} \\ X \times I & \xrightarrow[F]{} & B. \end{array}$$

A map $\pi : E \to B$ is a *Serre fibration* if the above is known to hold only for $X$ a CW complex (or equivalently a disk). If $\pi : E \to B$ is a fibration with path connected base, and $b \in B$, then $\pi^{-1}(b)$ is called the *homotopy fibre* of $\pi$.

To show these generalize geometric notions, we have the following pair of results.

**Theorem 1.7: Borsuk's Theorem**

For every CW pair $A \subset X$, i.e. a CW complex $X$ and subcomplex $A$, the inclusion $i : A \hookrightarrow X$ is a cofibration.

Proof. Let $F : A \times I \to Y$ so that $F_0$ extends to $f : X \to Y$. We wish to extend $F$ to $X \times I$. We can work inductively on the dimension of cells in $X \setminus A$. On the 0 skeleton, we just define $F$ to constantly agree with $f$. Now suppose $F$ has been extended to the $n$ skeleton of $X$. On a given $n+1$ cell $e^{n+1}$, we wish to define $F$ on the cylinder $e^{n+1} \times I$ given that we know it on the "vase" $\partial e^{n+1} \times I \cup e^{n+1} \times \{0\}$. Note that there is a deformation retraction map from the cylinder to the vase and we can extend $F$ on the cylinder by composing this retraction with $F$ defined on the vase. Doing this for all $n+1$ cells gives a continuous extension of $F$ to $X_{n+1} \times I$. We conclude by induction. □

**Theorem 1.8**

Every fibre bundle $\pi : E \to B$ is a Serre fibration. If $B$ is paracompact, $\pi$ is a Hurewicz fibration.

Proof. We only deal with the Serre case. If the fibre bundle is trivial, this follows from Borsuk's theorem. For a non-trivial bundle $E$ consider the case where we have a relative homotopy from the disk $(D^n, \partial D^n)$ we wish to extend. Then we can pullback the bundle and it will be trivial by Feldbau's lemma below. Hence we may apply the above result. For homotopies from general CW complexes, we may just glue relative homotopies on each cell. □

A couple more important results about fibrations.

**Theorem 1.9: Feldbau's Lemma**

Every Hurewicz fibration $\pi : E \to B$ with contractible CW base $B$ and homotopy fibre $F$ is homotopy equivalent to the projection $B \times F \to B$. If $\pi$ is only Serre, there is still a weak homotopy equivalence.

PROOF. We prove only for a fibre bundle. By naturality of bundles, if $f : B \to *$ and $g : * \to B$ so that $g \circ f \sim \mathrm{id}_B$ we have that $E = (g \circ f)^* E = f^* g^* E$. But $g^* E$ is a bundle over a point and hence trivial, so its pullback by $f$ is too. □

The following justifies the definition of homotopy fibre.

**Theorem 1.10**

The fibres of a Hurewicz fibration over a path connected base are homotopy equivalent. The fibres of Serre fibration over a path connected base are weak homotopy equivalent.

PROOF. Consider a path $\gamma$ between two points $x_0, x_1$ in the base with fibres $F_0, F_1$. Given a map $Z \to F_0$, we can use $\gamma$ and the definition of a fibration to obtain a map $G : Z \times I \to E$ so that $G_0$ lands in $F_0$ and $G_1$ lands in $F_1$. This gives a map from $[Z, F_0]$ to $[Z, F_1]$. We may compose paths to compose maps and homotopic paths induce homotopic maps. Taking the inverse path we conclude our map is a bijection and it is easily seen to be natural in $Z$. Thus by Proposition 1.3, $F_0, F_1$ are homotopy equivalent. If we only have a Serre fibration, we are restricted to $Z$ a CW complex from which we conclude we have a weak homotopy equivalence. □

A final interesting result is as follows.

**Theorem 1.11**

Every map $f : X \to Y$ is homotopic (by which we mean equivalent in the homotopy category) to a fibration and to a cofibration.

PROOF. $Y$ is homotopic to the mapping cylinder $M_f$ and so our map is equivalent in the homotopy category to the embedding $X \hookrightarrow M_f$, which is a cofibration. The homotopy cofibre of this map is the mapping cone $C_f$

For the fibration case, first by above, we may assume $f : X \to Y$ is an embedding. Now consider the space $\mathcal{P}_X Y$ of continuous paths in $Y$ that begin in $f(X)$. Truncating the paths gives a homotopy equivalence with $X$. Hence our map is equivalent in the homotopy category to the map $\mathcal{P}_X Y \to Y$ sending a path to its endpoint. This is naturally a fibration with homotopy fibre $E_X Y$, the space of paths in $Y$ beginning in $X$ and ending at $y_0$. □

REMARK 1.12. Let $f : X \to Y$ be a map which we homotope by above to be a cofibration and inclusion. We can extend to a coexact sequence of cofibrations,

$$X \xrightarrow{f} Y \to C_f \to \Sigma X \to \Sigma Y \to \Sigma C_f \to \Sigma^2 X \to \cdots .$$

We can see what's going on here with a simple picture seen in Figure 1 (noting if we take $X \subset Y$, then we can identify $C_f$ with gluing to $Y$ a cone over $X \subset Y$).

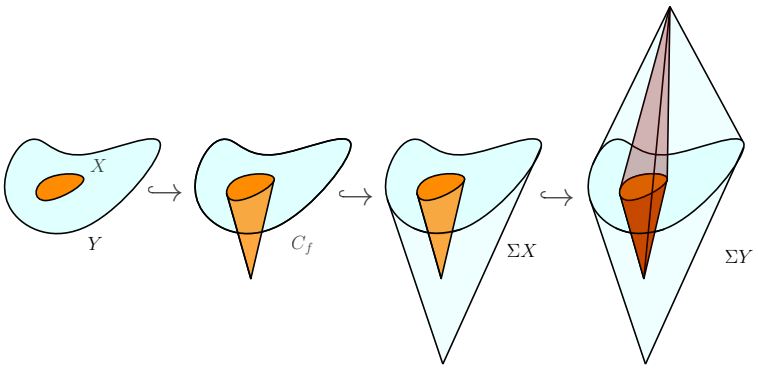



FIGURE 1. Puppe sequence

Similarly, there is an exact sequence associated to a fibration,

$$\cdots \to \Omega^2 Y \to \Omega(E_X Y) \to \Omega X \to \Omega Y \to E_X Y \to X \to Y.$$

These are both called the *Puppe sequence*. △

**1.4. Cellular Approximation.** We say a map $X \to Y$ of CW complexes is *cellular* if $X_n$ maps into $Y_n$ for each $n$.

**Theorem 1.13: Cellular Approximation Theorem [FF, pp. 52]**

A continuous map $f : X \to Y$ of CW complexes is homotopic to a cellular map. More generally, if $A \subset X$ is a subcomplex and $f|_A$ is cellular, than $f$ is homotopic relative to $A$ to a cellular map.

PROOF. (Sketch) We work by induction. Suppose $f : X \to Y$ is cellular on $X_{n-1} \cup A$ and let $e^n$ be an $n$-cell of $X \setminus A$. $f$ maps the closure of $e^n$ to a compact set of $Y$. Thus it can only meet finitely many cells of $Y$. Let $\varepsilon$ be one of these cells of $Y$ of maximal dimension $m$.

There is a technical lemma that we skip which implies, for $m > n$, $f|_{A \cup X^n \cup e^n}$ is homotopic to a map $f'$, cellular on $A \cup X^n$, so $f'(e^n)$ intersects the same cells as $f(e^n)$ and is not surjective onto $\varepsilon$. By Borsuk's theorem $f'$ extends to $X$. On $e^n$, we may define a homotopy of $f'$ pushing its image off of $\varepsilon$ by radial projection away from a point not in $f'(e^n)$. By Borsuk's theorem, this homotopy can be extended to $X$. Inductively, we can ensure $f(e^n)$ lands in no cells of dimension greater than $n$. Doing this for every cell, we obtain our result inductively. □

## 2. The Fundamental Group and Coverings

We keep this section short and sweet. Results about fundamental groups can usually be reproved as exercises with just a little creativity.

### 2.1. Fundamental Group.

We give a very abbreviated treatment.

**Definition 2.1**

The *fundamental group* $\pi_1(X, x_0)$ of a space is the set $[(S^1, *), (X, x_0)]$ endowed with the operation of loop composition. Equivalently, it is the path components of $\Omega X$.
Up to isomorphism, the fundamental group only depends on the path component of $x_0$ and so we often abbreviate to $\pi_1(X)$.
If $X$ is connected and $\pi_1(X)$ is trivial, we say $X$ is *simply connected*.

A continuous map $f : X \to Y$ induces a group homomorphism $f_* : \pi_1(X) \to \pi_1(Y)$ by post-composition with a loop. This is natural with respect to composition of maps and depends only on the homotopy class of our maps.

COROLLARY 2.2. *Homotopy equivalent spaces have the same fundamental group.*

### 2.2. Covering Spaces.

We again only sketch the basics.

**Definition 2.3**

A *covering map* $p : T \to X$ is a fibre bundle with discrete fibre; we call $T$ a *covering space* of $X$. Equivalently, $X$ is covered by neighbourhoods

$V_i$ whose preimage under $p$ is a disjoint collection of open sets $U_{ij}$ so that $p : U_{ij} \to V_i$ is a homeomorphism for each $i$ and $j$.

**Theorem 2.4: Map Lifting**

Let $p : T \to X$ be a covering and $Z$ path connected and locally path connected. A map $f : Z \to X$ lifts to $F : Z \to T$ if and only if $f_*\pi_1(Z, z_0) \subset p_*\pi_1(T, t_0)$. If we specify, $F(z_0) = t_0$ with $p(t_0) = f(z_0)$, the map is unique.

In particular, a path $\gamma : I \to X$ always lifts to $\widetilde{\gamma} : I \to T$ and the lift is unique if we specify $\widetilde{\gamma}(0)$. By a relative version of the map lifting, if $F_t : Z \to X$ is a homotopy and $F_0$ lifts to $\widetilde{F_0} : Z \to T$, then there is a unique lift $\widetilde{F_t} : Z \to T$ of $F_t$ agreeing with $\widetilde{F_0}$ for $t = 0$.

**Theorem 2.5: Classification of Covering Spaces**

If $X$ is path connected, locally path connected, and semi-locally simply connected, then for every subgroup $H \subset \pi_1(X)$, there is a covering space $p : X_H \to X$ with $p_*(\pi_1(X_H)) = H$ (and hence $\pi_1(X_H) \cong H$). Moreover there is a bijection between isomorphism classes of coverings and conjugacy classes of subgroups of $\pi_1(X)$.

PROOF. We only construct the *universal cover* $\widetilde{X}$, which is the simply-connected cover. Other covering spaces are realized as quotients of $\widetilde{X}$ in an intuitive way. Fix a basepoint $x_0 \in X$. We define $\widetilde{X}$ as the homotopy classes of paths in $X$ with fixed endpoints that begin at $x_0$. There is a map $p : \widetilde{X} \to X$ by projecting to the end of a path which is clearly a covering. Moreover, given a loop $\gamma$ in $X$ that lifts to a loop in $\widetilde{X}$, we see that the lifted loop $[\gamma_t]$ represents the homotopy class of the loop $\gamma$ truncated at time $t$. But the fact it lifts to a loop means $[\gamma] = [\gamma_0]$ is null-homotopic. Hence, $p_*(\pi_1(\widetilde{X})) = 0$. But uniqueness in Theorem 2.4 implies that $p_*$ is injective, and so $\pi_1(X) = 0$.

To see the correspondence, note that by changing the base point we obtain conjugate subgroups of $\pi_1(X)$ and isomorphic coverings. □

**Definition 2.6**

Given a covering space $p : T \to X$, an isomorphism $f : T \to T$ is called a *deck transformation* if it commutes with $p$. Note a deck transformation is fully determined by the image of one point.

**Definition 2.7**

A covering $p : T \to X$ is *normal* (or sometimes called *regular*) if $p_*\pi_1(T)$ is a normal subgroup of $\pi_1(X)$.

PROPOSITION 2.8. *A covering $p : T \to X$ is normal if and only if the group of deck transformations acts transitively on the set $\{p^{-1}(x_0)\}$, where $x_0$ is a point of $X$.*

PROPOSITION 2.9. *The group of deck transformations $D$ of a covering $p : T \to X$ is isomorphic to the quotient of the normalizer,*

$$D = \pi_1(X)/N_{\pi_1(X)}(p_*\pi_1(T)).$$

We can finally make some computations.

COROLLARY 2.10. $\pi_1(S^1) = \mathbb{Z}$.

PROOF. Note $\mathbb{R} \to S^1$ given by $z \mapsto e^{2\pi iz}$ is the universal covering of $S^1$. We know that a subgroup of the group of deck transformations isomorphic to $\mathbb{Z}$ is given by translations by $2\pi in$ for $n \in \mathbb{Z}$. But since these translations act transitively on the preimages of the basepoint, this is in fact all of the deck transformations. By Proposition 2.9, we conclude $\pi_1(S^1) = D = \mathbb{Z}$. □

COROLLARY 2.11. $\pi_1(\mathbb{R}P^n) = \mathbb{Z}_2$ *for* $n > 1$.

PROOF. We will prove momentarily that $\pi_1(S^n) = 0$ for $n > 1$. Thus the universal cover of $\mathbb{R}P^n$ is $S^n$ under the double cover quotient map identifying antipodal points. The map $p : S^n \to S^n$ sending a point to its antipode is a deck transformation which swaps the two preimages of the basepoint. Hence the covering is normal, $p$ is the only non-trivial deck transformation, and $\pi_1(S^n) = D = \mathbb{Z}_2$. □

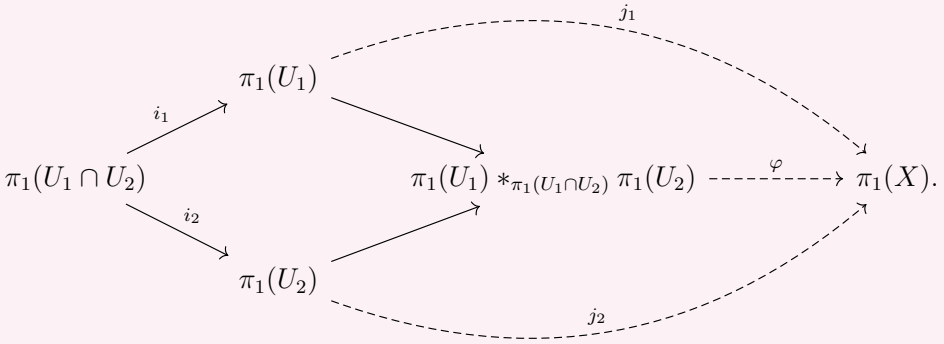
COROLLARY 2.12. $\pi_1(T^2) = \mathbb{Z}^2$.

PROOF. The universal cover is $\mathbb{R}^2 \to T^2$ given by the quotient map by the integer lattice. There are deck transformations $\mathbb{R}^2 \to \mathbb{R}^2$ given by translations by $(n, m)$ for $(n, m) \in \mathbb{Z}^2$. These act transitively on the preimages of the basepoint and so the covering is normal. We have $\pi_1(T^2) = D = \mathbb{Z}^2$. Alternatively, we could use that $\pi_1(T^2) = \pi_1(S^1 \times S^1) = \pi_1(S^1) \times \pi_1(S^1) = \mathbb{Z}^2$. □

**2.3. Seifert–Van Kampen Theorem and $\pi_1$ of CW Complexes.** The following is very useful in more advanced computations where there is no clear candidate for universal cover (or we want to show the space itself is simply connected).

**Theorem 2.13: Seifert–Van Kampen**

Let $X = U_1 \cup U_2$ with $U_1, U_2$ open, path connected and $U_1 \cap U_2$ path connected. Then the maps on $\pi_1$ induced by inclusions of $U_1 \cap U_2, U_1, U_2$ and $X$ define a pushout square, i.e. $\varphi$ is an isomorphism in the following colimit diagram:




Equivalently, if $\pi_1(U_2 \cap U_2)$ has generator set $G_0$, $\pi_1(U_1)$ has presentation $\langle G_1 | R_2 \rangle$, and $\pi_1(U_2)$ has presentation $\langle G_2 | R_2 \rangle$, then $\pi_1(X)$ has presentation,

$$\left\langle G_1, G_2 \middle| R_1, R_2, i_1 G_0 = i_2 G_0 \right\rangle.$$

COROLLARY 2.14. *For $n > 1$, the sphere $S^n$ is simply connected.*

PROOF. Let $U_1, U_2$ be open neighbourhoods of the two hemispheres of the sphere. They are each simply connected and have connected intersection a thickened $n - 1$ sphere. The pushout of the maps $\mathbb{Z} \to 0, \mathbb{Z} \to 0$ is clearly 0 and so by Theorem 2.13, $\pi_1(S^n) = 0$. □

Note by Seifert–Van Kampen, if we have a nice space $X$ and we attach a 2-cell $D$ to it via $\varphi : \partial D \to X$, the effect on $\pi_1$ will be to kill the element of $\pi_1(X)$ represented by $\varphi(\partial D)$. Furthermore, if we attach an $n$-cell to $X$ for $n > 2$, it will not change its fundamental group.

Now consider a one-dimensional path connected CW complex $X$. This is the same thing as a graph. By contracting edges joining distinct vertices of $X$, one sees that $X$ is homotopy equivalent to a (possibly infinite) bouquet of circles. We now find the fundamental group of $X$.

PROPOSITION 2.15. *If $X, Y$ are pointed spaces with base point contractible neighbourhoods, then $\pi_1(X \vee Y) = \pi_1(X) * \pi_1(Y)$.*

PROOF. This directly follows from Theorem 2.13 applied to the complements of these contractible neighbourhoods. □

COROLLARY 2.16. *The fundamental group of a bouquet of circles is a free group with generators corresponding to the circles.*

PROOF. For finite bouquets this follows inductively from the above proposition. For infinite bouquets, note any loop lies in only finitely many circles by compactness, and hence belongs in $\pi_1$ to the free group generated by those circles. □

Combining our findings, we can deduce the following theorem.

**Theorem 2.17: The Fundamental Group of a CW Complex**

Suppose $X$ is a connected CW complex with a single zero cell. Then $\pi_1(X)$ has a presentation whose generators correspond to the one cells of $X$ and whose relations impose that the loop class $[\partial e^2 \to X_1]$ as a product of generators is trivial for any two cell $e^2$ of $X$.

PROOF. The fundamental group of the one skeleton $X_1$ is a free group generated by the one-cells, since $X_1$ is a bouquet of circles, one circle per one-cell. Attaching a two-cell has the effect of killing its parameterized boundary and hence imposing a relation of the form described above. Doing this for all two cells inductively, we conclude $\pi_1(X_2)$ is as described in the theorem statement. Attaching a cell of dimension three or greater has no effect on $\pi_1$ and hence $\pi_1(X) = \pi_1(X_2)$ is as claimed. □

EXERCISE 2.18. Our assumption $X$ has a single zero cell is no restriction, since a connected CW complex $X$ is always homotopy equivalent to one with a single zero cell. ▲

We can now easily find the fundamental groups of all our favourite spaces.

**Some Fundamental Groups**

- $\pi_1(\Sigma_g) = \langle a_1, \ldots, a_g, b_1, \ldots, b_g | [a_1, b_1] \cdots [a_g, b_g] = e \rangle$.
- $\pi_1((\mathbb{R}P^2)^{\#g}) = \langle c_1 \ldots, c_g | c_1^2 \cdots c_g^2 \rangle$.
- $\pi_1(\mathbb{R}P^n) = \mathbb{Z}_2$ for $2 \leq n \leq \infty$ with universal cover $S^n$.
- $\pi_1(G(n,k)) = \mathbb{Z}_2$ for $1 \leq k \leq n-1$ and $n > 2$ with universal cover $G_+(n,k)$.
- The complex and quaternionic projective spaces, Grassmannians, and Stiefel manifolds are all simply connected.
- $\pi_1(\mathrm{SO}(n)) = \mathbb{Z}_2$ for $n \geq 3$. Its double cover is called $\mathrm{Spin}(n)$, the *Spin group*. There are exceptional isomorphisms,

  $\mathrm{Spin}(2) \cong S^1$, $\mathrm{Spin}(3) \cong S^3 \cong \mathrm{SU}(2)$,

  $\mathrm{Spin}(4) \cong \mathrm{SU}(2) \times \mathrm{SU}(2)$, $\mathrm{Spin}(5) \cong \mathrm{Sp}(2)$, $\mathrm{Spin}(6) \cong \mathrm{SU}(4)$.
- $\pi_1(\mathrm{SU}(n)) = \pi_1(\mathrm{Sp}(n)) = 0$ and $\pi_1(\mathrm{U}(n)) = \mathbb{Z}$. The universal cover of $\mathrm{U}(n)$ is $\mathbb{R} \times \mathrm{SU}(n)$ given by $(\theta, A) \mapsto e^{2\pi i\theta} A$.
- $\pi_1(\mathrm{Sp}(2n, \mathbb{R})) = \mathbb{Z}$ and $\pi_1(\mathrm{Sp}(2n, \mathbb{C})) = 0$. The double cover of $\mathrm{Sp}(2n, \mathbb{R})$ is the *metaplectic group* $\mathrm{Mp}(2n, \mathbb{R})$. It is *not* a matrix group.

## 3. Homology and Cohomology

We try to stick to the fundamentals, emphasizing a range of versions of homology, and focusing on algebraic tools and geometric intuition with minimal actual examples.

**3.1. Singular Homology and Its Many Imitators.** We begin in total generality. Take $R$ to be a commutative ring (actually any abelian group would be fine); we almost exclusively use $\mathbb{Z}, \mathbb{Z}_2, \mathbb{Q}$ or $\mathbb{R}$. If unspecified, $R$ should be taken to be $\mathbb{Z}$.

**Definition 3.1: Homology**

A *chain complex* $(A_i, \partial)$ is a collection of $R$-modules $A_i$ indexed by $\mathbb{Z}$ and boundary maps $\partial : A_i \to A_{i-1}$ so that $\partial^2 = 0$. The failure of the

complex,

$$\cdots \xrightarrow{\partial} A_{i+1} \xrightarrow{\partial} A_i \xrightarrow{\partial} A_{i-1} \xrightarrow{\partial} \cdots$$

to be an exact sequence is measured by its *homology groups*,

$$H_i(A) := \ker(\partial : A_i \to A_{i-1}) / \operatorname{im}(\partial : A_{i+1} \to A_i).$$

We call elements in this kernel *cycles* and elements in this image *boundaries*, so homology is the group of cycles modulo the boundaries.

A morphism of chain complexes $f : (A_*, \partial_A) \to (B_*, \partial_B)$, or a *chain map*, is a collection of module homomorphisms $f_i : A_i \to B_i$ so that $f_{i-1} \circ \partial_A = \partial_B \circ f_i$. Equivalently, the following commutes,

$$\begin{array}{ccccccccc}
\cdots & \longrightarrow & A_{i+1} & \xrightarrow{\partial_A} & A_i & \longrightarrow & A_{i-1} & \longrightarrow & \cdots \\
 & & \downarrow{\scriptstyle f_{i+1}} & & \downarrow{\scriptstyle f_i} & & \downarrow{\scriptstyle f_{i-1}} & & \\
\cdots & \longrightarrow & B_{i+1} & \xrightarrow{\partial_B} & B_i & \longrightarrow & B_{i-1} & \longrightarrow & \cdots .
\end{array}$$

A chain map $f : A_i \to B_i$ can be easily seen to induce a map on homology $f_* : H_i(A) \to H_i(B)$.

A *chain homotopy* between chain maps $f, g : (A_*, \partial_A) \to (B_*, \partial_B)$ is a collection of maps $h_i : A_i \to B_{i+1}$ so that $h \circ \partial_A + \partial_B \circ h = f - g$.

If there is a chain map $f : (A_*, \partial_A) \to (B_*, \partial_B)$ and a chain map $g$ in the other direction so that $f \circ g$ and $g \circ f$ are chain homotopic to the identities, then we say $f$ is a *quasi-isomorphism*.

PROPOSITION 3.2. *If $f, g : (A_*, \partial_A) \to (B_*, \partial_B)$ are chain homotopic, then they induce the same map on homology. If $f : (A_*, \partial_A) \to (B_*, \partial_B)$ is a quasi-isomorphism, then $(A_*, \partial_A)$ and $(B_*, \partial_B)$ have isomorphic homology.*

PROOF. Let $[x] \in H_i(A)$ be represented by $x \in \ker(\partial_A) \subset A_i$. Then, as an element of $H_i(B)$,

$$[f(x) - g(x)] = [h \circ \partial_A x] + [\partial_B \circ h(x)].$$

Since $x \in \ker \partial_A$,

$$= [\partial_B \circ h(x)].$$

Since we quotient by the image of $\partial_B$,

$$= 0.$$

So $[f(x)] = [g(x)]$ for any $[x] \in H_i(A)$ and the induced maps on homology thus agree. □

**Definition 3.3**

A *cochain complex* $(A_i, \mathrm{d})$ is like a chain complex but with coboundary maps $\delta : A_i \to A_{i+1}$ so that $\delta^2 = 0$. We extract from the complex its *cohomology groups*,

$$H^i(A) := \ker(\delta : A_i \to A_{i+1}) / \operatorname{im}(\mathrm{d} : A_{i-1} \to A_i),$$

which is the group of *cocycles* modulo *coboundaries*. We define *cochain maps* and *cochain homotopies* in the analogous way.

*Singular Homology.* Let $X$ be a topology space and let $C_n(X; R)$ be the (gigantic) free $R$-module generated by continuous maps $\Delta^n \to X$, for $\Delta^n$ the standard $n$-dimensional simplex. Elements of $C_n(X; R)$ are called *singular $n$-chains*.

We will orient our simplices in the sense that we keep track of an ordering of their vertices. We will give the boundary strata of a simplex compatible orientations induced from the ordering of their subcollection of vertices. Let $\Delta^n_i$ be the face of $\Delta^n$ excluding the $i$th vertex of $\Delta^n$. Given a map $\sigma : \Delta^n \to X$, we define its boundary,

$$\partial\sigma = \sum_{i=0}^{n} (-1)^i \sigma|_{\Delta^n_i}.$$

This extends by linearity to a map $C_n(X; R) \to C_{n-1}(X; R)$. Clearly $\partial^2 = 0$ and so we have a chain complex. We define the *singular homology of $X$* as the homology of the singular chain complex,

$$H_n(X; R) := H_n(C_*(X; R)).$$

Given a continuous map $f : X \to Y$, we obtain a pushforward map $f_* : C_n(X; R) \to C_n(Y; R)$, given by post-composing a singular simplex $\Delta^n \to X$ with $f$ and extending linearly. Explicitly written out, it is immediate that $f_*$ is a chain map.

**Theorem 3.4**

If $f, g : X \to Y$ are homotopic, they induce the same same maps on homology $f_*, g_* : H_n(X; R) \to H_n(Y; R)$. Hence homology is an invariant of the homotopy type of a topological space.

Proof. By our previous proposition, it suffices to exhibit a chain homotopy between $f_*, g_* : C_n(X; R) \to C_n(Y; R)$. Let $F : X \times I \to Y$ be a homotopy between $f$ and $g$. Given a singular simplex $\sigma : \Delta^n \to X$, we can

define a map,

$$H(\sigma) : \Delta^n \times I \xrightarrow{\sigma \times \mathrm{id}} X \times I \xrightarrow{F} Y.$$

We may split $\Delta^n \times I$ along diagonals into $n+1$ copies of $\Delta^{n+1}$, and interpret $H(\sigma)$ as a sum of $n+1$ singular $n+1$ simplices in $Y$. Extending by linearity, we thus obtain a map $H : C_n(X;R) \to C_{n+1}(X;R)$.

Note that $\partial H(\sigma)$ consists of several faces. There are the bottom and top faces coming from restricting $H(\sigma)$ to $\Delta^n \times \partial I$ which correspond to $f_*(\sigma)$ and $g_*(\sigma)$ respectively. And there are the side faces coming from restricting $H(\sigma)$ to $(\partial \Delta^n) \times I$, which correspond to $H(\partial \sigma)$. Analyzing orientations we conclude,

$$\partial H(\sigma) + H(\partial \sigma) = f_*(\sigma) - g_*(\sigma).$$

Hence, $H$ is a chain homotopy of $f_*$ and $g_*$. □

*Singular Cohomology.* Define a *singular cochain* as an $R$-valued functional on the space of singular chains. The $R$-module of singular cochains is the dual group to the singular chains: $C^n(X;R) = \mathrm{Hom}(C_n(X), R)$. We define $\delta : C^n(X;R) \to C^{n+1}(X;R)$ dual to $\partial$, so that given $\varphi \in C^n(X;R)$ and $\sigma \in C_{n+1}(X)$, $\delta(\varphi)(\sigma) = \varphi(\partial\sigma)$. Clearly $\delta^2 = 0$ and so $C^n(X;R)$ becomes a cochain complex. We define the *singular cohomology* of $X$ as the cohomology of the singular cochain complex,

$$H^n(X;R) := H^n(C^*(X;R)).$$

*Simplicial Homology.* We say a space $X$ has a *simplicial structure*, if it can be decomposed into simplices in the following sense:

(1) There is a collection of maps $\sigma : \Delta^n \to X$, where $n$ may vary, so that each map is injective on the interior of $\Delta^n$.

(2) Every point of $X$ is in the image of one of the maps, and no point is in the image of two of the maps restricted to the interior of $\Delta^n$ for the same $n$.

(3) Each map $\sigma : \Delta^n \to X$ when restricted to a face of $\Delta^n$ agrees with one of the maps $\sigma' : \Delta^{n-1} \to X$.

Remark 3.5. This definition is not quite correct (every map should be uniquely determined by its face maps) but any structure of the above form can be refined (by twice applying barycentric subdivision) to obtain a true simplicial structure. △

Let $X$ have a simplicial structure, and let $C_n^\Delta(X;R)$ be the free $R$-module generated by the structure maps $\Delta^n \to X$. This is a much smaller subcomplex of $C_n(X;R)$. By property (3), the boundary restricts to a map on this

subcomplex. And we may thus define the *simplicial homology* of $X$ as,

$$H_n^\Delta(X;R) := H_n(C_*^\Delta(X;R)).$$

The advantage of this theory is we can actually compute this homology, and moreover we may do it purely combinatorially. It turns out this homology is independent of our triangulation and in fact $H_n^\Delta(X;R) = H_n(X;R)$ for any simplicial complex; we will not prove this directly, but instead observe simplicial homology as a special case of the following construction.

*Cellular Homology.* Let $X$ be a CW complex with a specified cell structure. We define $\mathcal{C}_n(X;R)$ as the free $R$-module generated by the $n$-cells of $X$. We define a boundary map $\partial : \mathcal{C}_n(X;R) \to \mathcal{C}_{n-1}(X;R)$ as follows. If $e$ is an $n$-cell,

$$\partial e = \sum_{f \subset X_{n-1}} d_{ef} \cdot f,$$

where the sum is over $(n-1)$-cells $f$ in $X$, and then we extend by linearity. Here the coefficient $d_{ef}$ is the degree of the map $S^{n-1} \to S^{n-1}$ given by composing the attaching map $\partial D^n \to X_{n-1}$ of $e$ with the quotient map $X_{n-1} \to S^{n-1}$ given by identifying the complement of $f$ to a point. Recall the *degree* of a map $S^{n-1} \to S^{n-1}$ can be computed as the image of the unit under the induced map on $H^{n-1}(S^{n-1};\mathbb{Z})$. Alternatively, one can homotope the map to be smooth and make an oriented count of a generic point's preimage.

It's not too hard to see $\partial^2 = 0$ and hence conclude $\mathcal{C}_*(X;R)$ form a chain complex. We define the *cellular homology* of $X$ as the homology of the cellular chain complex,

$$H_n(X;R) := H_n(\mathcal{C}_*(X;R)).$$

It's easy to check this coincides with simplicial homology for a simplicial complex. We will give an alternate definition of cellular homology later that makes it manifestly equal to singular homology.

*De Rham Cohomology.* Let $X$ be a smooth manifold. It has a tangent bundle $TX$ and its dual the cotangent bundle $T^*X$. Taking exterior powers of this bundle gives the bundles $\bigwedge^n T^*X$, whose fibres are wedge products of covectors.

Sections of the bundle $\bigwedge^n T^*X$ are called *differential $n$-forms*; the space of $n$-forms is denoted $\Omega^n(X)$. If $x^1, \ldots, x^m$ are local coordinates on $X$, then $TX$ has a local basis of sections $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^m}$ and we denote the dual sections $\mathrm{d}x^1, \ldots, \mathrm{d}x^m$. We define the *exterior derivative* $\mathrm{d} : \Omega^n(X) \to \Omega^{n+1}(X)$ to be the unique $\mathbb{R}$-linear map satisfying the following properties:

(1) $\mathrm{d}f\,(V) = V(f)$ for any $f \in C^\infty(X)$ and $V \in \Gamma(TX)$.

(2) d satisfies a super-Leibniz rule,

$$\mathrm{d}(\alpha \wedge \beta) = \mathrm{d}\alpha \wedge \beta + (-1)^{|\alpha|}\alpha \wedge \mathrm{d}\beta.$$

(3) $\mathrm{d}^2 = 0$.

Locally, d is specified on an $n$-form by,

$$\mathrm{d}\left(\sum_{|I|=n} f_I \, \mathrm{d}x^I\right) = \sum_{i=1}^{m} \sum_{|I|=n} \frac{\partial f_I}{\partial x^i} \, \mathrm{d}x^i \wedge \mathrm{d}x^I \, .$$

The groups $\Omega^n(X)$ thus define a cochain complex over $\mathbb{R}$ with differential the exterior derivative d, called the de Rham complex. The *deRham cohomology* of $X$ is the cohomology of the de Rham complex,

$$H^n_{dR}(X) := H^n(\Omega^*(X); \mathbb{R}).$$

A consequence of the Whitney approximation theorem is that we if consider not all singular chains, but only those that are regular enough to be integrated over, we still obtain the same singular cohomology theory. One then has a map,

$$F : \Omega^n(X; R) \to \mathrm{Hom}(C_n(X), \mathbb{R}) = C^n(X; \mathbb{R}) \quad \text{by} \quad \omega \mapsto F_\omega(\sigma) = \int_{\Delta^n} \sigma^* \omega.$$

By computing locally and patching together with the Mayer-Vietoris theorem (discussed later) one concludes that $F$ is an isomorphism. Hence, for any smooth manifold, the de Rham cohomology coincides with the singular cohomology; this is de Rham's theorem.

*Morse Homology.* One final (more esoteric but dear to my heart and research) cousin of singular homology for compact manifolds is defined as follows. Let $f$ be a *Morse function* on a smooth closed manifold $X$. That is, the critical points of $f$ are non-degenerate, or equivalently the Hessian of $f$ at its critical points is non-singular. The *index* of a critical point $x$ of $f$ is the dimension of the negative eigenspace of the Hessian of $f$ at $x$.

If we pick a Riemannian metric $g$ on $X$, it defines a gradient vector field $\nabla f$ by the condition $g(\nabla f, V) = V(f)$ for any $V \in \Gamma(TX)$. This defines a complete gradient flow $\varphi_t$. We define the *unstable and stable manifolds* of a critical point $x$ respectively as,

$$W^u(x; f, g) = \{p \in X : \lim_{t \to -\infty} \varphi_t(p) = x\}$$
$$W^s(x; f, g) = \{p \in X : \lim_{t \to \infty} \varphi_t(p) = x\}.$$

These are open disks of dimension $\mathrm{ind}(x)$ and $\dim(X) - \mathrm{ind}(x)$ respectively by the Hartman–Grobman theorem using the flow to provide an explicit diffeomorphism.

We say that the pair $(f, g)$ is *Morse–Smale* if $f$ is Morse and for any pair of critical points, $x, y$ the intersection of $W^u(x; f, g)$ and $W^s(y, f; g)$ is transverse (i.e. the tangent spaces of the submanifolds span $TX$ at any point of intersection); it is a consequence of the Sard–Smale theorem that generically pairs $(f, g)$ are Morse–Smale. In this case, the intersection is a manifold of dimension $\mathrm{ind}(x) - \mathrm{ind}(y)$. Note the intersection also has a free $\mathbb{R}$-action from the flow. Quotienting by this action, we obtain a manifold of dimension $\mathrm{ind}(x) - \mathrm{ind}(y) - 1$, which we denote, $\mathcal{M}(x, y)$. Moreover, $\mathcal{M}(x, y)$ has a natural compactification as a manifold with corners, whose $k$th boundary stratum consists of $k$-times "broken flow lines" connecting $x$ to $y$ and stopping at $k$ intermediate critical points.

We now define the Morse complex for $(f, g)$ a Morse–Smale pair, with chain group $CM_n(X; f, g)$ the free $\mathbb{Z}_2$-module generated by the critical points of $f$ of index $n$. The set of such critical points is finite by the Morse condition. We define a boundary operator $\partial : CM_n(X) \to CM_{n-1}(X)$ by the linear extension of,

$$\partial x = \sum_{\mathrm{ind}(y)=n-1} \#\mathcal{M}(x, y) \cdot y.$$

Counting dimensions, $\mathcal{M}(x, y)$ is a zero dimensional compact manifold, i.e. a finite set of points, and so we can count the set modulo 2. We can also define this complex over $\mathbb{Z}$ using the observation that $W^u$ and $W^s$ are orientable, and the $\mathbb{R}$-action direction carries a natural orientation; we would then count $\mathcal{M}(x, y)$ as an oriented set of points.

To see that $\partial^2 = 0$, we note that the $\langle \partial^2 x, y \rangle$ consists of broken flow lines connecting $x$ to $y$ and stopping at one intermediate critical point. But by our description of the compactification, this is precisely the boundary of the 1-manifold $\mathcal{M}(x, y)$. A 1-manifold has no boundary points modulo 2 or counting orientations, and so $\langle \partial^2 x, y \rangle = 0$ for any $x$ and $y$.

Thus we obtain the *Morse homology* of $X$ as the homology of the Morse complex,

$$HM_n(X; f, g) := H_n(CM_*(X; f, g); \mathbb{Z}_2).$$

This again agrees with singular homology, which we now show. We may define a map $F : CM_n(X) \to C_n(X; \mathbb{Z}_2)$ by sending a critical point $x$ to the singular map extending the inclusion of the descending manifold of $X$ by compactifying (strictly we should be dealing with some kind of current instead of singular chains). This is a chain map because the boundary of the descending manifold

consists of the descending manifolds of critical points of index one less (counted mod 2 or with appropriate signs). We define $G : C_n(X; \mathbb{Z}_2) \to CM_n(X)$ that sends a generic singular $n$-simplex $\sigma$ to a sum of index $n$ critical points $x$ with coefficients counting mod 2 or with orientations the flow lines passing through the image of $\sigma$ at $t = 0$ and terminating at $x$. This is a chain map because the space of flow lines from $\sigma$ terminating at $q$ of index $n-1$ is a one-manifold with boundary given by flow lines of $\partial\sigma$ terminating at $q$ (i.e. $\langle G(\partial\sigma), q\rangle$) and flow lines of $\sigma$ terminating at some $p$ of index $n$ followed by flow lines from $p$ to $q$ (i.e. $\langle \partial(G\sigma), q\rangle$).

To a singular $n$-simplex $\sigma$, we associated $H(\sigma)$ which is the $n+1$-simplex given by compactifying the full forward gradient flow of $\sigma$. This extends to a map $H : C_n(X; \mathbb{Z}_2) \to C_{n+1}(X; \mathbb{Z}_2)$. Note that $\partial H(\sigma)$ consists of $\sigma$ itself, the forward gradient flow of $\partial\sigma$, and the descending manifolds of the critical points $\sigma$ flows to at infinity. That is, keeping track of signs,

$$\partial H(\sigma) + H(\partial\sigma) = F \circ G(\sigma) - \sigma,$$

hence $F \circ G$ is chain homotopic to the identity. On the other hand, $G \circ F$ is verbatim the identity on chain groups. Hence $F$ is a quasi-isomorphism and so our chain complexes have the same homology.

It is possible to extend to the case of compact manifolds $X$ with boundary. We can set up a Morse–Smale pair as usual under the additional assumption that $\nabla f$ points inward along all of $\partial X$. We can define the Morse complex like usual and obtain homology $HM_*(X)$ that again agrees with singular homology. If we instead assume that $\nabla f$ points outward along $\partial X$, we will define again Morse homology in the same way. However in this case the Morse homology computes something called the relative homology $H_*(X, \partial X)$ that we will shortly study more.

For more details on our construction of Morse homology and many more aspects of the theory, one can read the lecture notes of Hutchings [**Hutchings**].

Remark 3.6. Note, by studying the unstable manifolds which are homeomorphic to open balls (i.e. cells) and how they glue (i.e. to lower dimensional cells), we see that Morse homology endows every compact $n$-manifold with a CW structure. We will use Morse theory for convenient proofs of many fundamental results down the road. △

**3.2. Geometric Constructions: Reduced Homology, Homology of a Pair, Excision, and Mayer–Vietoris.** We will proceed to consider a number of extensions of the basic homology theory and a couple relevant

theorems, all based on some geometric ideas. Throughout we deal with singular homology over $\mathbb{Z}$ acknowledging that other theories and coefficients work identically.

*Reduced Homology.* What is $H_0(X)$? We know a singular 0-chain is a linear combination of maps $* \to X$. This is always a cycle since $\partial C_0(X) = 0$. A 0-chain is a boundary if there is a sum of maps $I \to X$ whose boundary is our 0-chain. Equivalently, two maps $* \to X$, representing points $x_1, x_2 \in X$ (their image) are the same in homology if and only if $x_1$ and $x_2$ are joined by a path.

We conclude $H_0(X)$ is a free $\mathbb{Z}$-module generated by the path components of $X$. This can be a little awkward since we might hope that $H_n(X) = 0$ for $X$ contractible; indeed this is true for $n > 1$ since $H_n(X) = H_n(*) = 0$. But, $H_0(*) = \mathbb{Z}$. To deal with this, we will sometimes want to reduce the dimension of the zeroth homology by one. Thus we define the *reduced homology* $\widetilde{H}_n(X)$ of $X$ as,

$$\widetilde{H}_n(X) = \begin{cases} H_n(X) & n \neq 0 \\ H_0(X)/[\sum x_i] & n = 0, \end{cases}$$

where the $x_i$ are maps $* \to X$ with image $x_i$, having picked one point $x_i$ in each path component of $X$. More formally, we could modify the chain complex $C_n(X)$ by setting $C_{-1}(X) = \mathbb{Z}$. And we could set $\partial : C_0(X) \to C_{-1}(X)$ to send any singular 0-simplex $* \to X$ to 1 and extend linearly. The homology of this extended complex will be exactly $\widetilde{H}_n(X)$.

*Homology of a Pair.* One has the following extension of the usual definition of homology. Given a subspace $A \subset X$, let $C_n(X, A)$ denote the quotient $C_n(X)/C_n(A)$. These are *relative singular n-chains*. This is a free $\mathbb{Z}$-module generated by simplices that don't entirely land within $A$. The boundary map clearly restricts to this subcomplex and the resulting homology $H_n(X, A)$ is called the *relative homology* of $(X, A)$.

EXERCISE 3.7. Show $H_n(X, *) = \widetilde{H}_n(X)$. ▲

**Theorem 3.8: Long Exact Sequence of a Pair and a Triple**

Given a topological pair $(X, A)$, its homology forms a long exact sequence,

$$\cdots \to H_n(A) \xrightarrow{i_*} H_n(X) \xrightarrow{j_*} H_n(X, A) \xrightarrow{\partial_*} H_{n-1}(A) \to \cdots .$$

The maps $i_*, j_*$ are induced by the inclusions $A \to X$ and $(X, \varnothing) \to (X, A)$ respectively. The map $\partial_*$ sends a relative $n$-cycle to its boundary, which is an $(n-1)$-cycle in $A$.

More generally, for a triple $B \subset A \subset X$, there is a long exact sequence,

$$\cdots \to H_n(A, B) \xrightarrow{i_*} H_n(X, B) \xrightarrow{j_*} H_n(X, A) \xrightarrow{\partial_*} H_{n-1}(A, B) \to \cdots .$$

PROOF. By definition of relative chains, there is a short exact sequence,

$$0 \to C_n(A) \xrightarrow{i_*} C_n(X) \xrightarrow{j_*} C_n(X, A) \to 0.$$

By the snake lemma, this becomes a long exact sequence in homology, with the maps as described. The triple case is the same. □

To simplify the computation of homology of a pair, one can apply the following result.

**Theorem 3.9**

Given a topological pair $(X, A)$ so that the inclusion $A \to X$ has homotopy cofibre $B = X \cup_A CA$, then $H_n(X, A) \cong \widetilde{H}_n(B)$. In particular, if $A \to X$ is a cofibration, for example a CW pair, then $H_n(X, A) = \widetilde{H}_n(X/A)$.

REMARK 3.10. The general proof is a little involved, using the concept of transformators. Fortunately for the case of CW pairs, which is what we care most about, this equality follows quite readily from the construction of cellular homology. △

We use the following important lemma.

LEMMA 3.11 (Refinement Lemma). *Let $X$ be a topological space with a collection $\mathcal{U}$ of sets whose interiors cover $X$. Let $C_n^{\mathcal{U}}(X)$ be the subcomplex of $C_n(X)$ generated by singular simplices which land inside sets of $\mathcal{U}$. Then the inclusion $C_n^{\mathcal{U}}(X) \to C_n(X)$ is a quasi-isomorphism. I.e. cycles in $X$ are homologous to ones subordinate to $\mathcal{U}$ and homologous cycles subordinate $\mathcal{U}$ differ by the boundary of a chain subordinate to $\mathcal{U}$.*

PROOF. Given a singular simplex $\sigma : \Delta^n \to X$, it can be decomposed via multiple applications of barycentric subdivision into a sum of simplices subordinate to $\mathcal{U}$. It's geometrically clear this subdivided chain is homologous

to $\sigma$. Moreover, its geometrically clear that if two cycles subordinate to $\mathcal{U}$ are cohomologous, we can subdivide their difference to be the boundary of a chain subordinate to $\mathcal{U}$. □

Proof. (of Thm. 3.9) Cover $B$ by $CA$ and $X \cup C'A$, the complement of a neighbourhood of the cone point; call this two set open cover $\mathcal{U}$. By the refinement lemma, the homology can be computed subordinate to $\mathcal{U}$. We have,

$$C_n^{\mathcal{U}}(B, CA) = C_n^{\mathcal{U}}(B)/C_n^{\mathcal{U}}(CA) = C_n(X \cup C'A)/C_n(C'A) = C_n(X \cup C'A, C'A).$$

We conclude the homology of $(B, CA)$ is the same as that of $(X \cup C'A, C'A)$, which by homotopy invariance is also that of $(X, A)$. Hence,

$$\widetilde{H}_n(B) = H_n(B, \mathrm{pt}) = H_n(B, CA) = H_n(X, A).$$

If $(X, A)$ is a cofibrant pair, $B \cong X/A$ giving the second statement. □

*Excision.* This is a crucial result for many geometric arguments.

**Theorem 3.12: The Excision Theorem**

Let $(X, A)$ be a topological pair. Suppose $B \subset A$ is such that the closure of $B$ is contained in the interior of $A$. Then the inclusion map induces an isomorphism in homology,

$$H_n(X \setminus B, A \setminus B) \cong H_n(X, A).$$

Proof. Consider the collection $\mathcal{U}$ consisting of $A$ and $B^c$. By our assumption, the interiors of this collection cover $X$. We have,

$$C_n^{\mathcal{U}}(X, A) = C_n^{\mathcal{U}}(X)/C_n^{\mathcal{U}}(A) = C_n(X \setminus B)/C_n(A \setminus B).$$

By the refinement lemma, we obtain our result. □

*Mayer–Vietoris.* Homology possesses an analogue of the Seifert–Van Kampen theorem that allows for ease of computations, by splitting a space into simpler pieces.

**Theorem 3.13: Mayer–Vietoris Theorem**

Suppose $X$ is the union of the interiors of subspaces $A$ and $B$. Then there is a long exact sequence,

$$\cdots \to H_n(A \cap B) \xrightarrow{i_*} H_n(A) \oplus H_n(B) \xrightarrow{j_*} H_n(X) \xrightarrow{\partial_*} H_{n-1}(A \cap B) \to \cdots.$$

Here $i_*$ denotes the difference of the maps induced by the two inclusions, and $j_*$ is the sum of the maps induced by the two inclusions. The map $\partial_*$ can described as follows. If $\sigma \in H_n(X)$, it is homologous to a sum

$x + y$ of chains in $A$ and $B$. We have $\partial x = -\partial y \in H_{n-1}(A \cap B)$. Then, we have that $\partial_* \sigma = \partial x$.

PROOF. Take $\mathcal{U}$ to be the collection $A$ and $B$. Then we have the short exact sequence,

$$0 \to C_n(A \cap B) \to C_n(A) \oplus C_n(B) \to C_n^{\mathcal{U}}(X) \to 0,$$

where the first map sends $\sigma$ to $(\sigma, -\sigma)$ and the second sends $(x, y)$ to $x + y$. By the refinement lemma, the homology of $C_n^{\mathcal{U}}(X)$ is $H_n(X)$. Thus by the snake lemma, we obtain the desired sequence. We easily deduce the maps are as described. □

*A Rigourous Construction of Cellular Homology.* First we need to actually know some homology.

PROPOSITION 3.14. *The homology of the $n$-sphere is,*

$$H_m(S^n) = \begin{cases} \mathbb{Z} & m = 0 \text{ or } n \\ 0 & \text{else.} \end{cases}$$

PROOF. We may cover the sphere by open neighbourhoods of the two hemispheres. These are contractible and their intersection is homeomorphic to the $n - 1$ sphere. We then proceed by induction. Since we know the homology of a pair of points, we may apply Mayer–Vietoris to find $H_m(S^1)$. Inductively, the computation of $H_m(S^n)$ determines from Mayer–Vietoris the homology of $H_m(S^{n+1})$. □

PROPOSITION 3.15.

$$H_m(D^n, S^{n-1}) = \begin{cases} \mathbb{Z} & m = n \\ 0 & \text{else.} \end{cases}$$

PROOF. By Theorem 3.9,

$$H_m(D^n, S^{n-1}) = \widetilde{H}_m(D^n / S^{n-1}) = \widetilde{H}_m(S^n).$$

□

PROPOSITION 3.16. *The reduced homology of a bouquet of $n$-sphere is,*

$$\widetilde{H}_m\left(\bigvee_{\alpha \in A} S^n_\alpha\right) = \begin{cases} \bigoplus_{\alpha \in A} \mathbb{Z} & m = n \\ 0 & \text{else.} \end{cases}$$

PROOF. In fact, for any space $X_\alpha$,

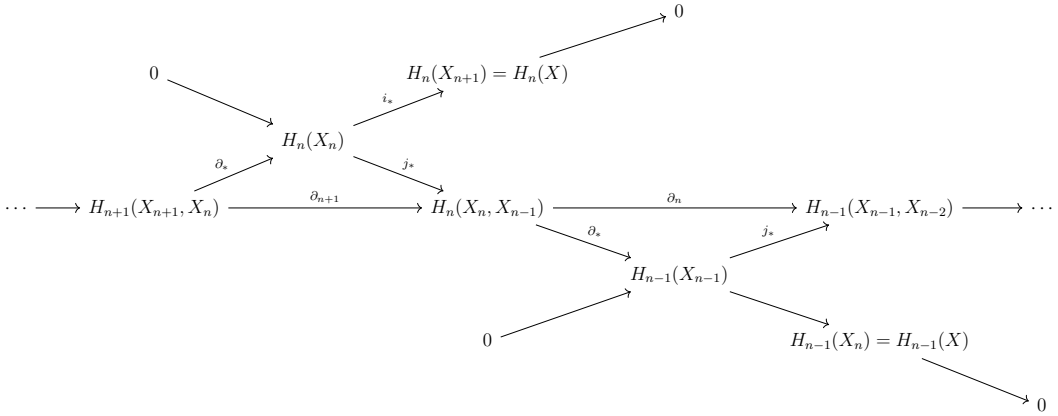
$$\widetilde{H}_m\left(\bigvee_{\alpha\in A} X_\alpha\right) = \bigoplus_{\alpha\in A} \widetilde{H}_m(X_\alpha).$$

To see this, the bouquet of spaces is the disjoint union quotiented by identifying base points. Thus by Theorem 3.9, our homology is the homology of $\bigsqcup_{\alpha\in A} X_\alpha$ relative to the union of base points. It is easy to see, the homology of the disjoint union of the $X_\alpha$ is the direct sum of the homologies of the $X_\alpha$. Thus our result follows from the long exact sequence of a pair. □

Now given a CW complex $X$, note that $X_n/X_{n-1}$ is a bouquet of $n$ spheres indexed by the cells of $X_n$. Hence we know the relative homology $H_*(X_n, X_{n-1})$.

Note also that $H_k(X_n) = 0$ for $k > n$. This follows from the fact $H_k(X_0) = 0$ for $k > 0$ and applying the long exact sequence of a pair inductively. Lastly, note that the map $H_k(X_n) \to H_k(X)$ is an isomorphism for $k < n$ and an epimorphism for $k = n$. If $X$ is finite dimensional, this follows directly from the long exact sequence of a pair. For infinite dimensional spaces, note each singular simplex must lie in a finite subcomplex of $X$, from which the conclusion again follows.

Now for the CW complex $X$, we can construct the following commutative diagram, where the diagonal sequences are truncations of the long exact sequence of a pair, and the horizontal maps are defined to make the diagram commute.




Note the composition $\partial_n \circ \partial_{n+1}$ includes the segment $\partial_* \circ j_*$ of the short exact sequence of a pair. And hence $\partial^2 = 0$. Thus, the horizontal sequence forms a chain complex. Note from our description of the relative homology groups that $H_n(X_n, X_{n-1}) \cong \mathcal{C}_n(X; \mathbb{Z})$. By following the arrows in the above diagram, it can be seen that $\partial$ is given by the same description as our boundary in cellular homology. Hence we have recovered the cellular chain complex and the cellular

homology. We can also show this description of cellular homology coincides with singular homology.

**Theorem 3.17**

If $X$ is a CW complex, its cellular and singular homologies coincide.

Proof. We know from the diagram, $H_n(X)$ is the quotient of $H_n(X_n)$ by the image of the boundary map $\partial_*$. Since $j_*$ is injective, $\operatorname{im}(\partial_*) = \operatorname{im}(\partial_{n+1})$. On the other hand,

$$H_n(X_n) = \operatorname{im}(j_*) = \ker(\partial_*) = \ker(j_* \circ \partial_*) = \ker(\partial_n).$$

Thus,

$$H_n(X) = H_n(X_n)/\operatorname{im}(\partial_*) = \ker(\partial_n)/\operatorname{im}(\partial_{n+1}) = H_n(\mathcal{C}_*(X); \mathbb{Z}).$$

□

*Eilenberg–Steenrod Axioms.* There is an axiomatic approach to homology which we now describe.

**Definition 3.18: Eilenberg–Steenrod Axioms**

A sequence of functors,

$$H_q(\cdot, \cdot; \pi) : \mathbf{hTopPair} \to \mathbf{AbGp},$$

from the homotopy category of topological pairs to the category of abelian groups together with natural transformations,

$$\partial : H_q(X, A; \pi) \to H_{q-1}(A; \pi) := H_{q-1}(A, \varnothing; \pi),$$

is called a *homology theory* if it satisfies the following axioms.

**Dimension:** If $X$ is a point, $H_0(X; \pi) = \pi$ and $H_q(X; \pi) = 0$ for all $q \neq 0$.

**Exactness:** Inclusion maps $A \to X$ and $(X, \varnothing) \to (X, A)$ induce a long exact sequence,

$$\cdots \to H_q(A; \pi) \to H_q(X; \pi) \to H_q(X, A; \pi) \xrightarrow{\partial} H_{q-1}(A; \pi) \to \cdots.$$

**Excision:** If $X$ is the union of the interiors of $A$ and $B$ then the inclusion $(A, A \cap B) \to (X, B)$ induces an isomorphism,

$$H_*(A, A \cap B; \pi) \to H_*(X, B; \pi).$$

**Additivity:** If $(X, A)$ is the disjoint union of pairs $(X_i, A_i)$, then the inclusions $(X_i, A_i) \to (X, A)$ induce an isomorphism,

$$\bigoplus_i H_*(X_i, A_i; \pi) \to H_*(X, A; \pi).$$

**Weak Equivalence:** If $f : (X, A) \to (Y, B)$ is a weak equivalence (see Theorem 4.25) then it induces an isomorphism,

$$f_* : H_*(X, A; \pi) \to H_*(Y, B; \pi).$$

**Theorem 3.19**

The Eilenberg–Steenrod axioms characterize a unique homology theory. Hence any sequence of functors satisfying these axioms agrees with the singular homology.

Proof. We have showed that singular homology satisfies all of these properties, with the exception of Weak Equivalence which we prove in Proposition 4.33. We will prove later in Theorem 4.31 that every space $X$ has a weak equivalence $f : X^{CW} \to X$ from a CW complex and so it suffices to prove that any homology theory for CW complexes agrees with cellular and hence singular homology.

One can deduce that all the geometric constructions of this section, like reduced homology and Mayer–Vietoris, can be obtained just from our axioms. In particular, the homology of the sphere and the relative homology of the disk to its boundary follow from the axioms. This allows us to construct the cellular chain complex as a sequence of relative homology groups of successive skeleta $H_*(X_n, X_{n-1})$ for any homology theory. The same arguments we applied before show the resulting cellular homology agrees with the underlying homology theory and has the usual geometric interpretation of its differential. Hence the cellular homology (and consequently the singular homology) of any CW complex $X$ is equal to its homology in our arbitrary homology theory. □

Of course totally analogous constructions hold for cohomology.

This argument gives us an alternate route to proving a certain homology theory agrees with singular homology: we just need to verify these axioms. This is a reasonable way to show that simplicial and Morse homology, and deRham cohomology all reproduce singular (co)homology.

If one wants to obtain something like homology that gives new algebraic homotopy invariants of space with similar geometric properties, we need to weaken these axioms. What turns out to be the best way to do this may seem surprising because it involves removing the most trivial of the axioms.

**Definition 3.20**

A *generalized* or *extraordinary homology theory* is a sequence of functors $E_q(\cdot,\cdot) : \mathbf{hTopPair} \to \mathbf{AbGp}$ that satisfy all of Eilenberg–Steenrod axioms, except possibly the dimension axiom.

Much of modern algebraic topology is focused on studying these theories and we will meet a few of the most important ones later.

**3.3. Algebraic Constructions: The Universal Coefficient Theorem, Künneth Formula, and Products on Cohomology.**

*Algebraic Preliminaries: Tor and Ext.* We will consider coefficients in the category of $\mathbb{Z}$-modules, i.e. abelian groups. Tensoring with a group $B$ is a *right exact functor*, meaning that if,

$$M \to N \to A \to 0 \quad \text{is exact, then so is} \quad M \otimes B \to N \otimes B \to A \otimes B \to 0.$$

However, this need not be true on the left. Conversely, the functor $\mathrm{Hom}(\cdot, B)$ is *left exact*, so that given an exact sequence, $M \to N \to A \to 0$, the sequence,

$$\mathrm{Hom}(M,B) \leftarrow \mathrm{Hom}(N,B) \leftarrow \mathrm{Hom}(A,B) \leftarrow 0$$

is also exact. The reverse statement need not be true. The Tor and Ext functors measure this failure to be exact on the other sides.

**Definition 3.21: Tor and Ext**

Given an $\mathbb{Z}$-module $A$, we may always find a *free resolution* of $A$, which is a short exact sequence,

$$0 \to F_1 \to F_0 \to A \to 0,$$

for $F_1, F_0$ free abelian groups (just use a group presentation).

Given $\mathbb{Z}$-modules $A$ and $B$, the *Tor functor* $\mathrm{Tor}(A,B)$ is the unique $\mathbb{Z}$-module so that, for a given free resolution $F_1 \to F_0 \to A$ of $A$,

$$0 \to \mathrm{Tor}(A,B) \to F_1 \otimes B \to F_0 \otimes B \to A \otimes B \to 0,$$

is exact.

The *Ext functor* $\mathrm{Ext}(A,B)$ is the unique $\mathbb{Z}$-module so that,

$$0 \to \mathrm{Hom}(A,B) \to \mathrm{Hom}(F_0,B) \to \mathrm{Hom}(F_1,B) \to \mathrm{Ext}(A,B) \to 0$$

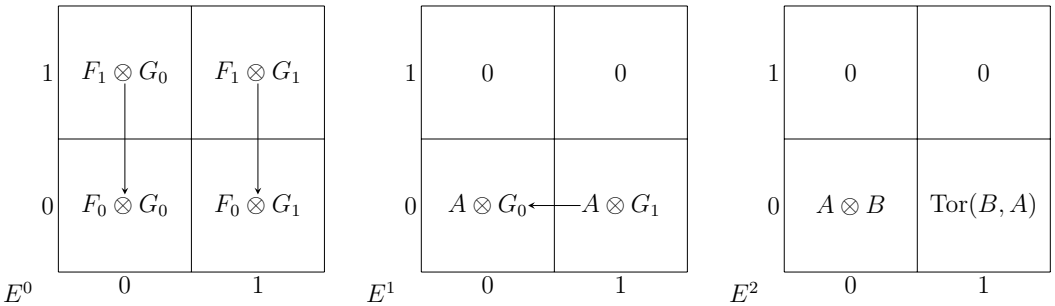
is exact. Both definitions are independent of the choice of free resolution.

PROPOSITION 3.22. *Tor is symmetric:* $\mathrm{Tor}(A,B) = \mathrm{Tor}(B,A)$.

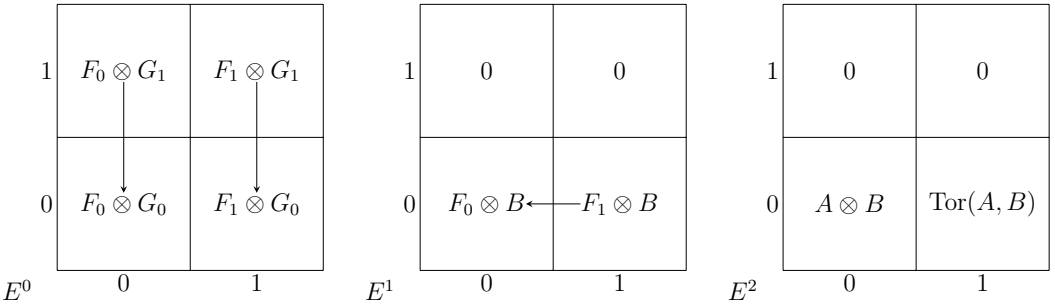
PROOF. We prove this with spectral sequences which we will define in a few sections. Suppose we have free resolutions,

$$0 \to F_1 \to F_0 \to A \to 0 \quad \text{and} \quad 0 \to G_1 \to G_0 \to B \to 0.$$

Tensoring together the $F_i$ and $G_j$ gives a bicomplex. We may extract a spectral sequence from each of the two gradings. In one direction, we have the following first three pages (we use $\mathrm{Tor}(A,F) = 0$ when $F$ is free, which we prove shortly).




At which point the spectral sequence terminates. If we apply the spectral sequence in the other direction, we have the following first three pages.




At which point the spectral sequence terminates. In either case, the spectral sequence computes the homology of the total bigraded complex. We conclude that $\mathrm{Tor}(A,B) = \mathrm{Tor}(B,A)$. □

PROPOSITION 3.23. *The following hold.*

*(1)* $\mathrm{Tor}(A,F) = 0$ *for* $F$ *free.*

*(2)* $\mathrm{Tor}(\mathbb{Z}_m, \mathbb{Z}_n) \cong \mathbb{Z}_{(m,n)}$.

*(3) For* $A, B$ *finitely generated,* $\mathrm{Tor}(A,B) \cong \mathrm{Tor}(A) \otimes \mathrm{Tor}(B)$.

*(4)* $\mathrm{Tor}(A, \mathbb{Q}) = 0$ *for any* $A$.

*(5)* $\mathrm{Ext}(F, B) = 0$ *for* $F$ *free.*

*(6)* $\mathrm{Ext}(\mathbb{Z}_m, \mathbb{Z}_n) = \mathbb{Z}_{(m,n)}$.

*(7)* $\mathrm{Ext}(\mathbb{Z}_m, \mathbb{Z}) = \mathbb{Z}_m$.

*(8)* $\mathrm{Ext}(A, \mathbb{Q}) = 0$ *for any* $A$.

PROOF. (1) Given a free resolution $0 \to F_1 \to F_0 \to A \to 0$, we get an injective morphism of free abelian groups $F_1 \to F_0$. But then the induced map $F_1 \otimes F \to F_0 \otimes F$ is clearly injective too if $F$ is free. Hence,

$$\mathrm{Tor}(A, B) = \ker(F_1 \otimes F \to F_0 \otimes F) = 0.$$

(2) $\mathbb{Z}_m$ has the free resolution,

$$0 \to \mathbb{Z} \xrightarrow{\times m} \mathbb{Z} \to \mathbb{Z}_m \to 0.$$

Tensoring with $\mathbb{Z}_n$ gives an exact sequence,

$$0 \to \mathrm{Tor}(\mathbb{Z}_m, \mathbb{Z}_n) \to \mathbb{Z}_n \xrightarrow{\times m} \mathbb{Z}_n \to \mathbb{Z}_m \otimes \mathbb{Z}_n \to 0.$$

Thus,

$$\mathrm{Tor}(\mathbb{Z}_m, \mathbb{Z}_n) = \ker(\mathbb{Z}_n \xrightarrow{\times m} \mathbb{Z}_n) = \mathbb{Z}_{(m,n)}.$$

(3) This follows from (1) and (2), plus the fundamental theorem of finitely generated abelian groups and the fact Tor distributes over direct sums.

(4) Given a free resolution $0 \to F_1 \to F_0 \to A$, $F_1 \to F_0$ is injective. Tensoring a free group $F$, i.e. no torsion, with $\mathbb{Q}$ has the consequence that $f \otimes (p/q) = f' \otimes (p'/q')$ if and only if $pq'f = p'qf'$. In particular, if $f \otimes (1/q) = e \otimes 0$ then $f = 0$ and hence $F_1 \otimes \mathbb{Q} \to F_0 \otimes \mathbb{Q}$ is also injective.

(5) We may take the free resolution $0 \to 0 \to F \to F \to 0$, from which the conclusion is clear.

(6) Again take the free resolution $0 \to \mathbb{Z} \xrightarrow{\times m} \mathbb{Z} \to \mathbb{Z}_m \to 0$. Applying $\mathrm{Hom}(\cdot, \mathbb{Z}_n)$ gives,

$$0 \to \mathrm{Hom}(\mathbb{Z}_m, \mathbb{Z}_n) \to \mathrm{Hom}(\mathbb{Z}, \mathbb{Z}_n) \xrightarrow{\circ \times m} \mathrm{Hom}(\mathbb{Z}, \mathbb{Z}_n) \to \mathrm{Ext}(\mathbb{Z}_m, \mathbb{Z}_n) \to 0.$$

Recall that $\mathrm{Hom}(\mathbb{Z}_m, \mathbb{Z}_n) \cong \mathbb{Z}_{(n,m)}$ and $\mathrm{Hom}(\mathbb{Z}, \mathbb{Z}_n) \cong \mathbb{Z}_n$, both being determined by where we send the unit. So the above becomes,

$$0 \to \mathbb{Z}_{(n,m)} \to \mathbb{Z}_n \xrightarrow{\times m} \mathbb{Z}_n \to \mathrm{Ext}(\mathbb{Z}_m, \mathbb{Z}_n) \to 0.$$

Thus,

$$\mathrm{Ext}(\mathbb{Z}_m, \mathbb{Z}_n) = \mathrm{coker}(\mathbb{Z}_n \xrightarrow{\times m} \mathbb{Z}_n) = \mathbb{Z}_n / m\mathbb{Z}_n \cong \mathbb{Z}_{(m,n)}.$$

(7) Using the same free resolution and Applying $\mathrm{Hom}(\cdot, \mathbb{Z})$ gives,

$$0 \to \mathrm{Hom}(\mathbb{Z}_m, \mathbb{Z}) \to \mathrm{Hom}(\mathbb{Z}, \mathbb{Z}) \xrightarrow{\circ \times m} \mathrm{Hom}(\mathbb{Z}, \mathbb{Z}) \to \mathrm{Ext}(\mathbb{Z}_m, \mathbb{Z}) \to 0.$$

We have $\mathrm{Hom}(\mathbb{Z}, \mathbb{Z}) = \mathbb{Z}$ while $\mathrm{Hom}(\mathbb{Z}_m, \mathbb{Z}) = 0$. So this becomes,

$$0 \to \mathbb{Z} \xrightarrow{\times m} \mathbb{Z} \to \mathrm{Ext}(\mathbb{Z}_m, \mathbb{Z}) \to 0.$$

Hence $\mathrm{Ext}(\mathbb{Z}_m, \mathbb{Z}) \cong \mathbb{Z}_m$ as claimed.

(8) It is a classic fact that $\mathbb{Q}$, and any divisible group, is an injective $\mathbb{Z}$-module, so that given an injective map $F_1 \to F_0$, any element of $\mathrm{Hom}(F_1, \mathbb{Q})$ lifts to an element of $\mathrm{Hom}(F_0, \mathbb{Q})$. Hence, $\mathrm{Hom}(\cdot, \mathbb{Q})$ is right exact and $\mathrm{Ext}(A, \mathbb{Q}) = 0$ for any $A$. □

*Universal Coefficient Theorem.* Recall that for a group $G$, $C_*(X; G) = C_*(X; \mathbb{Z}) \otimes G$ and $C^*(X; G) = \mathrm{Hom}(C_*(X; \mathbb{Z}); G)$. This need not be true on the level of homology, since tensoring and Hom fail to be exact. However, it's not so surprising that we can relate the homology between different coefficients as well as relate the homology and the cohomology using the Ext and Tor.

First, we have a simple way to relate homology with different coefficients.

Proposition 3.24 (Coefficient Long Exact Sequence). *Suppose $C_*(X)$ is a chain complex of free abelian groups. Suppose there is an exact sequence of abelian groups $0 \to A \to B \to C \to 0$. Then there is a long exact sequence in homology,*

$$\cdots \to H_n(X; A) \to H_n(X; B) \to H_n(X; C) \xrightarrow{\beta} H_{n-1}(X; A) \to \cdots .$$

*And in cohomology,*

$$\cdots \to H^n(X; A) \to H^n(X; B) \to H^n(X; C) \xrightarrow{\beta} H^{n+1}(X; A) \to \cdots .$$

*The maps $\beta$ are called the* Bockstein homomorphisms.

Proof. Since the chain complex is free, we have that $\mathrm{Tor}(G, C_n(X)) = \mathrm{Ext}(C_n(X), G) = 0$. Hence $\cdot \otimes C_n(X)$ and $\mathrm{Hom}(C_n(X), \cdot)$ are exact functors and we obtain short exact sequences,

$$0 \to C_n(X; A) \to C_n(X; B) \to C_n(X; C) \to 0,$$

and,

$$0 \to C^n(X; A) \to C^n(X; B) \to C^n(X; C) \to 0.$$

The result follows from the snake lemma. □

**Theorem 3.25: Universal Coefficient Theorem**

Suppose $C_*(X)$ is a chain complex of free abelian groups. There are short exact sequences for any $n \in \mathbb{Z}$ and any abelian group $G$,

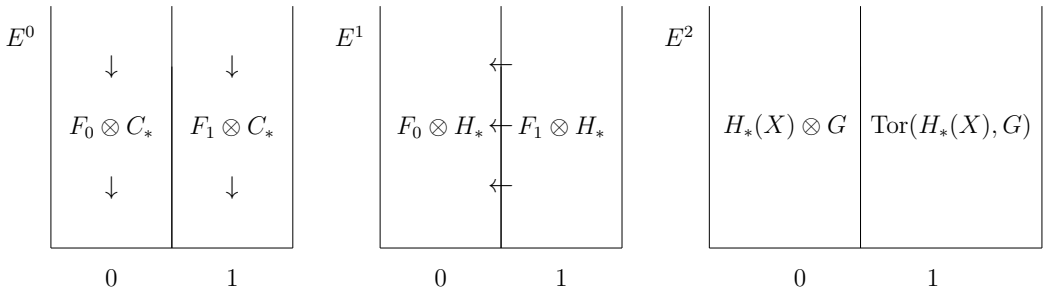
(i) $$0 \to H_n(X) \otimes G \to H_n(X;G) \to \operatorname{Tor}(H_{n-1}(X), G) \to 0,$$

(ii) $$0 \to H^n(X) \otimes G \to H^n(X;G) \to \operatorname{Tor}(H^{n+1}(X), G) \to 0,$$

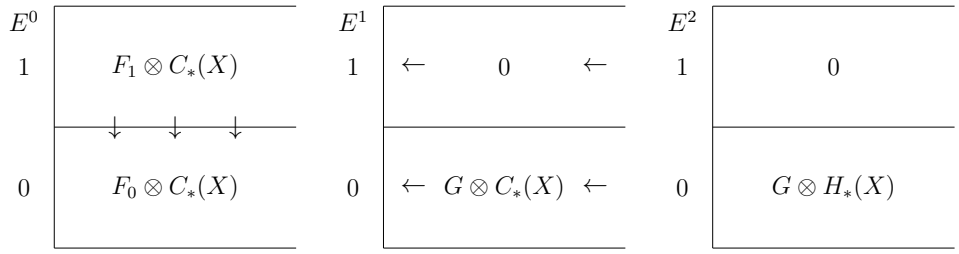
(iii) $$0 \to \operatorname{Ext}(H_{n-1}(X), G) \to H^n(X;G) \to \operatorname{Hom}(H_n(X), G) \to 0.$$

These sequences are canonical. They split non-canonically.

PROOF. (i) Let $0 \to F_1 \to F_0 \to G \to 0$ be a free resolution of $G$. Then $0 \to F_1 \to F_0 \to 0$ defines a complex and tensoring with $C_*(X)$, we obtain a bicomplex. We can look at the spectral sequence associated to each grading. In one direction, we degenerate after the following three pages,




While in the direction, we degenerate after the following three pages,




In both spectral sequences, the $E^\infty$ page should give the homology of the total bicomplex. Thus we have,

$$\bigoplus_{p+q=n} E^\infty_{p,q} = (H_n(X) \otimes G) \oplus \operatorname{Tor}(H_{n-1}(X); G) = H_n(X;G).$$

(ii) This follows from the same argument applied to the bicomplex given by tensoring with $C^*(X)$.

(iii) The full chain complex $C_*(X)$ decomposes as a direct sum over $n$ of short complexes,

$$0 \to B_n \xrightarrow{\partial_{n+1}} Z_n \to 0,$$

of boundaries and cycles. This short sequence is a resolution for $H_n(X)$. Applying the $\mathrm{Hom}(\cdot, G)$ functor gives,

$$0 \to \mathrm{Hom}(H_n(X), G) \to \mathrm{Hom}(Z_n, G) \to \mathrm{Hom}(B_n, G) \to \mathrm{Ext}(H_n(X), G) \to 0.$$

Thus the short co-complex,

$$0 \to \mathrm{Hom}(Z_n, G) \xrightarrow{\delta_{n+1}} \mathrm{Hom}(B_n, G) \to 0,$$

has $(n+1)$st homology $\mathrm{Ext}(H_n(X), G)$ and $n$th homology $\mathrm{Hom}(H_n(X), G)$.

The direct sum of the homologies of all these short co-complexes must be the homology of $H^*(X; G)$. Taking the relevant homology from the $Z_n, B_n$ and $Z_{n-1}, B_{n-1}$) complexes, we have,

$$H^n(X; G) = \mathrm{Hom}(H_n(X), G) \oplus \mathrm{Ext}(H_{n-1}, G).$$

□

The universal coefficient theorem has the following corollary which is simpler to remember and is used all time.

**Theorem 3.26**

Suppose the homology groups $H_n(X)$ of a space are all finitely generated. Then,

$$\boxed{H^n(X) = H_n^{\text{free}}(X) \oplus H_{n-1}^{\text{tor}}(X),}$$

where free and tor denote the free and torsion parts of the homology.

Proof. This is immediate from the universal coefficient theorem if we remember the simple description of torsion for finitely generated abelian groups. □

*Künneth Formula.* What can we say about the homology of a product of spaces? If we have two CW complexes $X, Y$, we can obtain a cell structure on $X \times Y$ whose cells are products of pairs of cells from $X$ and $Y$ with attaching maps induced from the attaching maps for $X$ and $Y$.

This means that $X \times Y$ has a cellular complex which is the tensor product of the complexes of $X$ and $Y$ with boundary operator the sum of the boundary operators on $X$ and $Y$. While this tensor product description doesn't descend to homology, it does after some correction from Tor groups. This relationship

extends from CW complexes to all spaces by the technique of CW approximation.

**Theorem 3.27: The Künneth Formula**

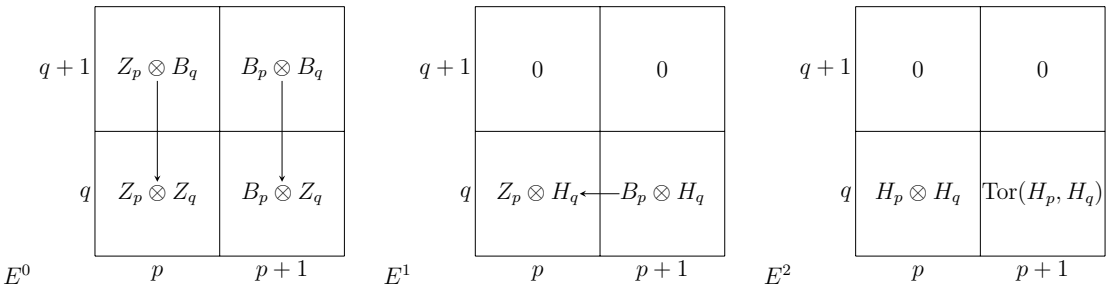
Given two spaces $X$ and $Y$, the homology of their product is,

$$H_n(X \times Y) \cong \bigoplus_{p+q=n} H_p(X) \otimes H_q(X) \bigoplus \bigoplus_{p+q=n-1} \operatorname{Tor}(H_p(X), H_q(X)).$$

Proof. As we show later, every space $X$ is weak homotopy equivalent to a CW complex. As we show later, weak homotopy equivalent spaces have isomorphic homology. Hence, without loss of generality $X$ and $Y$ are CW complexes.

By our comments above, the cellular complexes satisfy $\mathcal{C}_*(X \times Y) = \mathcal{C}_*(X) \otimes \mathcal{C}_*(Y)$.

Consider the short complexes $0 \to B_p \to Z_p \to 0$ and $0 \to B_q \to Z_q \to 0$. For $B, Z$ boundaries and cycles. The tensor product of these is a graded filtered complex. The associated spectral sequence has the following three pages after which it degenerates.




Taking the direct sum over all pairs of complexes $B_p \to Z_p, B_q \to Z_q$ gives $\mathcal{C}_*(X \times Y)$. And the direct sum of these spectral sequences should compute the cohomology of the product $H_*(X \times Y)$. Checking the indexing of our entries in the spectral sequence, we obtain the claimed formula. □

*Multiplications.* While the construction of cohomology may seem superfluous given the universal coefficient theorem, cohomology has the benefit over homology of having the structure of a ring. Throughout, suppose $R$ is a commutative ring.

**Definition 3.28: Cup Product**

Let $\varphi \in C^n(X;R)$ and $\psi \in C^m(X;R)$. The *cup product* $\varphi \smile \psi \in C^{n+m}(X;R)$ is defined so that for a singular simplex $\sigma : \Delta^{n+m} \to X$,

$$(\varphi \smile \psi)(\sigma) = \varphi(\sigma|_{[0,\dots,n]}) \cdot \psi(\sigma|_{[n,\dots,m]}).$$

Here, $[0,\dots,n]$ denotes the face of $\Delta^{n+m}$ containing the first $n+1$ vertices and $[n,\dots,m]$ denotes the face of $\Delta^{n+m}$ containing the last $m+1$ vertices. Extending by linearity, we obtain a module homomorphism,

$$\smile : C^n(X;R) \times C^m(X;R) \to C^{n+m}(X;R).$$

Exercise 3.29. Show the cup product has the following properties:

(1) $(\alpha \smile \beta) \smile \gamma = \alpha \smile (\beta \smile \gamma)$.

(2) $\delta(\alpha \smile \beta) = \delta\alpha \smile \beta + (-1)^{|\alpha|}\alpha \smile \delta\beta$.

(3) Given $f : X \to Y$, $f^*(\alpha \smile \beta) = f^*\alpha \smile f^*\beta$.

(4) The cup product is super-commutative (i.e. commutative with a sign $(-1)^{|\alpha|\cdot|\beta|}$) up to some quasi-isomorphism permuting the vertices of simplices. ▲

This implies the cup product descends to a multiplication on cohomology,

$$\smile : H^n(X;R) \otimes_R H^m(X;R) \to H^{n+m}(X;R),$$

which is associative, super-commutative, and natural with respect to pullback by continuous maps. This makes the direct sum of the cohomology groups, which we denote $H^*(X;R)$, a graded ring.

There are two more operations on (co)homology which we sometimes use.

**Definition 3.30**

Suppose $X_1, X_2$ are two spaces, and let $\pi_i : X_1 \times X_2 \to X_i$ be the two projections. Given $\alpha \in H^n(X_1;R)$ and $\beta \in H^m(X_2;R)$, we can define the *cross product*,

$$\times : H^n(X_1;R) \times H^m(X_2;R) \to H^{n+m}(X_1 \times X_2;R),$$

by $\alpha \times \beta = \pi_1^*\alpha \smile \pi_2^*\beta$.

Given $\alpha \in H^n(X;R)$ and $\sigma : \Delta^{n+m} \to X$, we can define the *cap product* $a \frown \alpha \in H_m(X;R)$ by,

$$\sigma \frown \alpha = \alpha(\sigma|_{[0,\dots,n]})\sigma|_{[n,\dots,n+m]}.$$

Extending by linearity, we obtain a module homomorphism,

$$\frown : C_{n+m}(X;R) \times H^n(X;R) \to C_m(X;R).$$

EXERCISE 3.31. Show the cap product has the following properties:

(1) $(\partial\sigma) \frown \alpha = \sigma \frown \delta\alpha + (-1)^{|\alpha|}\partial(\sigma \frown \alpha)$.

(2) $\sigma \frown (\alpha \smile \beta) = (a \frown \alpha) \frown \beta$. ▲

Thus the cap product descends to an operation on homology,

$$\frown : H_{n+m}(X;R) \otimes H^n(X;R) \to H_m(X;R).$$

We may also perform a relative version of these operations. Suppose $A, B \subset X$ are sufficiently nice, for example they have an open neighbourhood deformation retracting to them. Then we may define relative versions of the cup and cap product:

$$\smile : H^*(X,A) \otimes H^*(X,B) \to H^*(X, A \cup B)$$
$$\frown : H_*(X,A) \otimes H^*(X,B) \to H_*(X, A \setminus B).$$

*Homology with Twisted Coefficients.* For a couple of our constructions later, it will be useful to study homology where the coefficients depend on some local data.

**Definition 3.32**

A *local coefficient system* $\{G_x\}$ on a space $X$ is a choice of abelian group $G_x$ for each $x \in X$ and a homomorphism $\Phi_\gamma : G_{\gamma(0)} \to G_{\gamma(1)}$ for each $\gamma : [0,1] \to X$ so that:

(i) $\Phi_\gamma$ depends only on the homotopy class of $\gamma$ with fixed endpoints.

(ii) If $\gamma_1, \gamma_2$ are composable, $\Phi_{\gamma_1 * \gamma_2} = \Phi_{\gamma_1} \circ \Phi_{\gamma_2}$.

(iii) For $\gamma$ a constant path, $\Phi_\gamma = \mathrm{id}$.

EXAMPLE 3.33. An important example for later is that if $E \to B$ is a Serre fibration, then the homology of its fibres $\{H_*(E_x)\}$ defines a local coefficient system on the base $B$. ▲

Consider a local coefficient system $\mathscr{G} = \{G_x\}$ on $X$. Let $\sigma : \Delta^k \to X$ be a singular simplex. Note since $\Delta^k$ is contractible that $G_\sigma(t)$ are canonically isomorphic for all $t \in \Delta^k$ and we may denote the local coefficients on the

simplex by $G_\sigma$. Let $C_*(X;\mathscr{G})$ be the free $\mathbb{Z}$-module generated by pairs $(\sigma, g)$ where $\sigma$ is a singular simplex and $g \in G_\sigma$ modulo the relation,

$$(\sigma, g_1) + (\sigma, g_2) = (\sigma, g_1 + g_2).$$

We can extend the differential $\partial$ from the usual singular complex to this complex by letting it act linearly on the $\mathscr{G}$ coefficients. This still satisfies $\partial^2$ and so $C_*(X;\mathscr{G})$ is a chain complex.

**Definition 3.34**

The homology of $X$ with *local* or *twisted coefficients* in a local coefficient system $\mathscr{G}$ is the homology of the above chain complex,

$$H_*(X;\mathscr{G}) = H(C_*(X;\mathscr{G})).$$

The cohomology with local coefficients is defined analogously.

**3.4. Homology on Manifolds.** The well behaved structure of smooth manifolds allows for important additional properties and a better geometric understanding of the homology. We will assume throughout that our manifolds are compact (possibly with boundary) and oriented. The same conclusions hold for unoriented manifolds over $\mathbb{Z}_2$ instead of $\mathbb{Z}$.

*Fundamental Class.* For manifolds, we should think of elements of $H_*(X)$ as representing something like submanifolds of $X$. Our intuition is that if $\dim(X) = n$, then $H_n(X)$ should be generated by a single element representing $X$ itself. This is essentially true, and there are a few paths to a proof.

**Definition 3.35**

A simplical structure for $X$ is an $n$-dimensional *pseudomanifold* if,

(1) The maximum dimension of a simplex in the simplicial structure is $n$.
(2) Any two $n$-simplices of $X$ is connected by a finite chain of $n$-simplices which pairwise share an $n-1$ dimensional face.
(3) Every $n-1$ dimensional simplex is the face of precisely two $n$-simplices of $X$ (if we allow for boundaries, they may be the face of just one $n$-simplex as well).

It is a classical theorem of topology that every smooth $n$-manifold is homeomorphic to a $n$-dimensional triangulated subspace of Euclidean space, so that the homeomorphism is smooth on the $n$-simplices. And moreover, this triangulation is a $n$-dimensional pseudomanifold.

**Theorem 3.36: Fundamental Class**

Given a smooth compact oriented $n$-manifold $X$, we have $H_n(X, \partial X) = \mathbb{Z}$ if $X$ is orientable, and equals zero otherwise. For any smooth compact $n$-manifold $H_n(X, \partial X; \mathbb{Z}_2) = \mathbb{Z}_2$.

PROOF. We can use our pseudomanifold structure to study simplicial homology. Let $a \in H_n(X, \partial X)$ be non-zero and represented by a linear combination of $n$-simplices with no boundary. Since every $(n-1)$-simplex in $X/\partial X$ is a face of two $n$-simplices, and the $n$-simplices are all connected, we immediately deduce that every $n$-simplex of $X$ appears in $a$, and furthermore all have the same coefficient up to a sign. If we work modulo 2 then we are done, giving a proof of the second part.

Working over $\mathbb{Z}$, we see that $H_n(X, \partial X)$ is generated by at most one element. Such a generator $a$ exists if and only if we can compatibly orient all the $n$-simplices of $X$ to match signs in our linear combination. If we have such compatible orientations, then the orientations of each simplex patch together to give an orientation of $X$. Conversely, if we cannot compatibly orient all the simplices, then there is a path through a sequence of $n$-simplices connecting an $(n-1)$ simplex to itself which reverses orientation, and so $X$ is not orientable. □

PROOF. (Take Two) I have come up with another admittedly quite long and arduous proof, which avoids the need for pseudomanifolds (which I dislike) and instead uses Morse homology (which I like very much). The first part borrows an argument from here to show we can find a Morse function on $X$ with a unique local maximum. We'll instead find a function with a unique local minimum, and by flipping a sign we get what we want. Pick any Morse function $f$. Let $p_1, \ldots, p_a$ be the local minima, i.e. index zero critical points and $q_1, \ldots, q_b$ the index one critical points ($b$ may equal zero but $a \geq 1$ by the extreme value theorem). By decreasing $f$ adding a negative bump function near the $p_i$ we can ensure $f(p_i) < f(q_j)$ for any pair $i, j$.

Pick a compatible metric $g$. The network of gradient flow lines between the $p_i$ and the $q_j$ will form an finite embedded graph in $X$. Note that this embedded graph is the one-skeleton of a cell structure for $X$. Since $X$ is connected, the graph must be connected (by cell approximation any path connecting 0-cells in $X$ lands in the 1-skeleton). Now consider a spanning tree $T$ for the subcollection of vertices $p_1, \ldots, p_a$, which exists by basic graph theory. Let $U$ be a tubular neighbourhood of $T$ inside $X$. By adding to $f$ a large negative bump function subordinate to $U$ we may ensure that $f(q_i) < C < f(q_j)$ for any

pair of index 1 critical points with $q_i$ in our graph and $q_j$ outside the graph and some constant $C$. Now consider $N = f^{-1}((-\infty, C])$. This is an open submanifold with boundary of $X$ containing $T$ and no other critical points of $f$. Applying the gradient flow of $f$ on $N \setminus T$ determines a smooth deformation retract of $N$ to $T$, which is itself contractible. Hence $N$ is diffeomorphic to a closed $n$-ball in $X$. We may alter $f$ on $N$, replacing it by some radial distance function smoothed at the boundary. This altered $f$ has a unique local minimum on $N$ and hence on all of $X$.

Considering the Morse complex for a Morse function $f$ with a unique local maximum $p$, we have $CM_n(X) \cong \mathbb{Z}$ generated by $p$. It just remains to show $\partial p = 0$ iff $X$ is oriented and always $\partial p = 0 \pmod 2$. Fix any critical point $q$ of index $n-1$. The Hessian of $f$ is symmetric so its eigenvalues are real and non-zero by our Morse assumption. Hence we can apply the Hartman–Grobman theorem which allows us to locally linearize the gradient flow of $f$ near $q$. We conclude from the local linear picture and the fact $\mathrm{Hess}_q(f)$ has only one positive eigenvalue, that there are exactly two gradient flow lines that stably converge to $q$. This immediately gives us that $\langle \partial p, q \rangle = 0 \pmod 2$.

The final (difficult) ingredient is to show these two flow lines converging to $q$ have opposite orientation precisely when $X$ is oriented. While orientations in Morse theory are hard, they are not so bad for the case of codimension one. We endow all the unstable manifolds of critical points with orientations of our choosing (this is possible because they are diffeomorphic to contractible disks). If $\gamma$ is a flow line from $p$ of index $n$ to $q$ of index $n-1$, then the tangent space to $X$ at $q$ has an orientation from the orientation of the unstable manifold at $q$ direct summed with the asymptotic direction of the gradient field along $\gamma$ (i.e. using the vector pointing out of the gradient flow line at $q$). We give a sign $\pm 1$ to our flow line according to if this orientation agrees or disagrees with the orientation of $T_qX$ given by the unstable manifold of $p$; this is the sign in the oriented count of flow lines defining the Morse boundary operator.

We can change our Morse function by a sign, so that $p$ is the unique critical point of index zero and the $q_i$ are critical points of index one. From Morse theory as we've described it, there is thus a cell structure for $X$ with a single zero cell given by $p$ and one cells given by the gradient flow lines. In particular, each index one critical point $q$ has two flow lines from $q$ to $p$ that together determine a loop in $M$, which is a one cell with both ends glued to $p$. Hereafter, we change our function back by a sign, so again $p$ is the global maximum and the points $q_i$ have index $n-1$, with the new knowledge that these flow lines from $p$ to $q$ determine a 1-skeleton for a CW structure of $X$.

Note that by choice of a metric, $TX$ has structure group $\mathrm{O}(n)$ and orienting $TX$ is equivalent to a reduction of structure group to $\mathrm{SO}(n)$. Such reduction is the same as a section of the quotient bundle $Q$ with fibre $\mathrm{O}(n)/\mathrm{SO}(n) \cong \mathbb{Z}_2$. Clearly $\mathbb{Z}_2$ has $\pi_0$ group $\mathbb{Z}_2$ and all its higher homotopy groups are trivial. We will conclude from obstruction theory, but one can probably prove directly, that there is an obstruction to extending a section of $Q$ from the 0-skeleton of $X$ (which can always be constructed trivially) to the 1-skeleton given by an element of $H^1(X;\mathbb{Z}_2)$ (this is the first Stiefel-Whitney class $w_1(TX)$). And once we have a section of $Q$ on the one-skeleton of $X$, there is no further obstruction to extending it to all of $X$.

We conclude that $X$ will be orientable if and only if we can define an orientation of $TX$ over the gradient flow lines from $p$ to the index $n-1$ critical points $q_i$. If we remove the points $q_i$, we are left with a "spider shaped" graph which is contractible and hence we can define an orientation of $TX$ over it. Thus $X$ is orientable if and only if the orientation extends to each index one critical point $q$. This will be the case if and only if the orientations as defined on the two flow lines converging to $q$ meet compatibly at $q$.

Now pick an orientation on the one-skeleton of gradient flow lines punctured at the index $n-1$ critical points. At one such point $q$, use the two gradient flow lines converging to $q$ to define two candidate orientations for $T_qX$. We can represent these orientations by an ordered basis for $T_qX$ whose first $n-1$ vectors span and induce a given orientation on the tangent space to the unstable manifold of $q$ and whose last vector spans the tangent to the gradient flow line, pointing in whichever direction necessary to agree with the overall orientation of the punctured one-skeleton. Since the first $n-1$ vectors in each oriented basis agree, we see that the orientation extends to $q$ if and only if the last vector in the pair of oriented bases point in the same direction. Because, by Hartman–Grobman, the two gradient flow lines meeting at $q$ lie on opposite sides of the tangent space to the unstable manifold of $q$, the orientations extend if and only if the orientation on exactly one of two gradient lines includes the direction agreeing with the direction of gradient flow. Referring back to our description of signs in Morse homology, the orientation extends to $q$ if and only if exactly one of the two gradient flow lines carries a positive sign in the boundary operator of the Morse complex. That is, the orientation extends to $q$ if and only if $\langle \partial p, q\rangle = 0$ (and not $\pm 2$). Repeating this analysis for every index $n-1$ critical point, we see that the orientation extends to the one skeleton of $X$, and hence to all of $X$ by our discussion above, if and only $\partial p = 0$.

So if $X$ is orientable, $p$ is a non-zero generator for $H^n(M;\mathbb{Z}) \cong \mathbb{Z}$ While if $X$ is not orientable, $\partial p \neq 0$ and $H^n(M;\mathbb{Z}) = 0$. Nothing we said should use

in a fundamental way that the manifold be closed, so we can with little work extend to the case of compact manifolds with boundary. Phew. □

EXERCISE 3.37. Show by excision that for any $x$ in the interior of $X$ a smooth manifold of dimension $n$, $H_i(X, X \setminus \{x\}) \cong H_i(S^n)$. This is called the *local homology* of $X$ at $x$. Furthermore, show that the map,

$$i_* : H_n(X, \partial X) \to H_n(X, X \setminus \{x\}),$$

induced by inclusion of the pair is an in isomorphism. Show that if $X$ is oriented, the local homology $H_n(X, X \setminus x)$ has a canonical generator induced by the orientation. ▲

If $X$ is oriented with a given orientation, then there is a canonical generator of $H_n(X, \partial X)$, described in the language of the above proof as follows. We consider a signed sum of the $n$-simplices of $X$ which assigns a coefficient $\pm 1$ to a simplex based off if the smooth embedding $\Delta^n \to X$ is orientation preserving or reversing. Equivalently, the isomorphism described in the above exercise should send the generator of $H_n(X, \partial X)$ to the canonical generator of the top local homology.

**Definition 3.38: Fundamental Class**

If $X$ is a compact oriented $n$-manifold, the canonical generator of $H^n(X, \partial X)$ is called the *fundamental class* and denoted $[X, \partial X]$. If $X$ is closed, we write $[X] \in H_n(X)$ for the fundamental class. If $X$ is non-oriented we use the same name and notation to refer to the unique non-zero element of $H_n(X, \partial X; \mathbb{Z}_2)$.

**Definition 3.39: Degree**

The *degree* of a continuous map $f : M \to N$ between compact $n$-dimensional oriented manifolds is the integer by which $f_* : H_n(M, \partial M) \to H_n(N, \partial N)$ multiplies elements by under the isomorphism of top homology with $\mathbb{Z}$. Equivalently, the degree $d$ satisfies $f_*([M, \partial M]) = d[N, \partial N]$.

We can always homotope our map to be smooth, for which we have the following equivalent definitions of degree.

(1) By Sard's theorem, $f$ has an open dense subset of regular values. The degree of $f$ is an oriented count of the preimage of any regular value.

(2) If $\omega$ is a volume form on $N$ normalized to have volume one, then the degree is given by,

$$d = \int_M f^*\omega.$$

To see these two are equivalent, note that the critical values are closed and measure zero, so do not affect the integral. On the regular value set, we can define a partition of unity and apply the change of variables theorem to deduce the equivalence. To see (1) is the same as the topological version, pick a regular value $y$ and note that the map,

$$H_n(M) \to H_n(M, M \setminus f^{-1}(y)) \to H_n(N, N \setminus y)$$

should be multiplication by the degree. But by excision, $H_n(M, M \setminus f^{-1}(y))$ is a direct sum of the local homology at each point of $f^{-1}(y)$, and so the above map factors as a multiplication by the sum of local degrees of the map near each point of $f^{-1}(y)$. For our smooth map, each local degree is just plus or minus one dependent on orientation from which we recover definition (1).

*Poincaré Duality.* We are now ready for the most fundamental result about the topology of manifolds.

**Theorem 3.40: Poincaré Isomorphism Theorem**

If $X$ is a smooth compact oriented $n$-manifold, one has for each $i$,

$$H_i(X, \partial X) \cong H^{n-i}(X) \quad \text{and} \quad H_i(X) \cong H^{n-i}(X, \partial X).$$

In particular, in the case of a closed oriented manifold, one has,

$$H_i(X) \cong H^{n-i}(X).$$

If $X$ is not oriented, the same isomorphisms hold over $\mathbb{Z}_2$. Moreover, these isomorphisms are induced by capping with the fundamental class $[X, \partial X]$ or its dual in cohomology.

Remark 3.41. The isomorphism in the case of a closed manifold is called *Poincaré duality*, while its generalization to manifolds with boundary is called *Lefschetz duality*. For non-compact manifolds, there is a more general statement provided we switch one of the groups in each isomorphism to (co)homology with compact support. This can be defined as $\operatorname{colim} H^*(M, M \setminus K)$ taken over compact sets $K$ directed by inclusion. Equivalently in de Rham theory, it is the cohomology of the complex of differential forms with compact support. $\triangle$

PROOF. We follow a somewhat unorthodox argument through Morse theory. Consider a Morse function $f$; we can perturb $f$ so that all of its critical points have distinct critical values. Near some critical value $c$ denote $M_- = f^{-1}((-\infty, c-\varepsilon))$ and $M_+ = f^{-1}((-\infty, c+\varepsilon))$ for $\varepsilon$ sufficiently small.

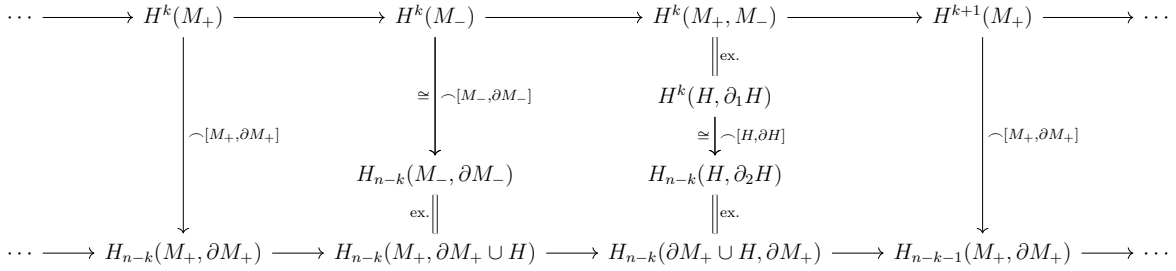
Note that if there are no critical points on the interval $[a, b]$, then gradient flow determines a diffeomorphism between $f^{-1}((-\infty, a))$ and $f^{-1}((-\infty, b))$. In particular, the topology of $f^{-1}((-\infty, a))$ only changes when we pass through the level set of a critical point. Thus, we may inductively prove Poincaré duality by assuming it holds for $M_-$ and showing it then holds for $M_+$ for an arbitrary critical value $c$.

Recall, the Morse lemma says that there are local coordinates $(x_1, \ldots, x_n)$ near a critical point so that our Morse function has the form.

$$f(x) = f(c) - x_1^2 - \cdots - x_i^2 + x_{i+1}^2 + \cdots + x_n^2,$$

where $i = \text{ind}(c)$. Thus, $M_+$ is obtained from $M_-$ by gluing on a handle $H = D^i \times D^{n-i}$. The boundary of $H$ is the union of two pieces $\partial_1 H = \partial D^i \times D^{n-i}$ and $\partial_2 H = D^i \times \partial D^{n-i}$. We see that the handle $H$ is glued to $M_-$ along $\partial_1 H$.

Now assume inductively that Poincaré duality holds for $M_-$. The base case is Poincaré duality for a disk, which is clear. Consider the triple $(M_+, \partial M_+ \cup H, \partial M_+)$ and the pair $(M_+, M_-)$ We may consider the homology sequence of the triple and the cohomology sequence of the pair, which are connected by the cap product.




Here, ex. denotes excision isomorphisms. We know the cap products in the middle two columns are isomorphisms, since we know Poincaré duality holds for the disk $H$ and we assume it holds for $M_-$. Thus, we have a morphism of long exact sequences, which is an isomorphism on two of every three columns. The five lemma implies the other columns are also isomorphisms, and so the map,

$$\frown [M_+, \partial M_+] : H^k(M_+) \to H_{n-k}(M_+, \partial M_+),$$

is an isomorphism. By induction, this extends to hold for $(M, \partial M)$. Dualizing our sequences and using the universal coefficient theorem gives the second

isomorphism. The mod 2 case follows from the same argument using the mod 2 fundamental class. □

PROOF. (Take Two) We can provide a simpler Morse theoretic proof of the isomorphism (without the fact it is induced by capping with $[M]$).

Take $(f, g)$ to be a Morse–Smale pair on $M$. We can study its negative gradient flow to obtain the usual Morse complex and then the Morse homology. Now study the Morse homology of the pair $(-f, g)$. Note that the critical points of $f$ of index $i$ are in correspondence with the critical points of $-f$ of index $n - i$:

$$CM_i(M; f, g) \cong CM_{n-i}(M; -f, g). \tag{3.1}$$

Under this isomorphism, we can write the Morse chain complex of $-f$ as a cochain complex,

$$\cdots \to CM_{i-1}(M; f, g) \xrightarrow{\delta} CM_i(M; f, g) \xrightarrow{\delta} CM_{i+1}(M; f, g) \to \cdots . \tag{3.2}$$

Given critical points $p, q$ of index $i$ and $i + 1$ with respect to $f$ (so $n - i$ and $n - i - 1$ with respect to $-f$), the coefficient $\langle \delta p, q \rangle$ is an oriented count of negative gradient flow lines of $-f$ from $p$ to $q$. But a negative gradient flow line of $-f$ is the same as a positive gradient flow line of $f$. And so changing the direction, $\langle \delta p, q \rangle$ is an oriented count of negative gradient flow lines of $f$ from $q$ to $p$. As a warning, these orientations may be different from the ones in the Morse complex of $f$; we can only guarantee they agree in the case where $M$ is orientable and thus admits coherent orientations for all its unstable manifolds (think back to our fundamental class result). We conclude, modulo two, $\langle \delta p, q \rangle = \langle \partial q, p \rangle$ where $\partial : CM_{i+1}(M; f, g) \to CM_i(M; f, g)$ is the usual Morse boundary, and the equality holds over $\mathbb{Z}$ if $M$ is oriented.

Hence under 3.1, $\delta$ is the adjoint of $\partial$ and the complex 3.2 is dual to the Morse chain complex of $f$. Since the homology of $CM(M; f, g)$ computes the singular homology of $M$, our dual complex 3.2 computes the singular cohomology of $M$.

But on the other hand, $(-f, g)$ is a Morse–Smale pair and we know its Morse complex has to compute the singular homology of $M$. Since the Morse complex of $-f$ is the same as 3.2 after re-indexing, we conclude,

$$H_i(M; \mathbb{Z}_2) \cong H^{n-i}(M; \mathbb{Z}_2),$$

and for $M$ oriented this extends to integers coefficients.

For the case where $M$ has boundary, if we start with a Morse function $f$ pointing inward on the boundary it will compute $H_*(M; \mathbb{Z}_2)$. From the above analysis, the Morse complex of $-f$ will compute $H^{n-*}(M; \mathbb{Z}_2)$. On the other

hand $-f$ is a Morse function pointing outward on the boundary and so its Morse complex will compute $H_*(M, \partial M; \mathbb{Z}_2)$. Conversely, if $\widetilde{f}$ points outward on the boundary, the Morse complex of $-\widetilde{f}$ simultaneously computes $H^{n-*}(M, \partial M; \mathbb{Z}_2)$ and $H_*(M; \mathbb{Z}_2)$. These results again extend over $\mathbb{Z}$ if $M$ is oriented. □

Poincaré duality together with the universal coefficient theorem give relatively strict restrictions on the potential homology of a smooth manifold. In particular we have for a closed oriented $n$-manifold $X$,

$$H_i^{\text{free}}(X) \cong H_{n-i}^{\text{free}}(X) \quad \text{and} \quad H_i^{\text{tor}}(X) \cong H_{n-i-1}^{\text{tor}}(X).$$

As an example of the applications of this, consider the following exercise.

Exercise 3.42. Suppose $M$ is a smooth closed 3-manifold with $H_1(M; \mathbb{Z}) = 0$. Apply the UCT and the (mod 2) Poincaré isomorphism to conclude that $H_3(M) = \mathbb{Z}$ and so by Theorem 3.36, $M$ is oriented (this also follows from obstruction theory). Then apply Poincaré duality and the UCT to show $M$ is a *homology sphere*, that is $H_*(M) = H_*(S^3)$. ▲

Remark 3.43. Let $M$ be a closed $n$-manifold which is not oriented or maybe not even orientable. A version of Poincaré duality still holds provided we use local coefficients. Define the local coefficient system $\mathscr{O}_x = H_n(M, M \setminus \{x\})$. The condition that $M$ is orientable is equivalent to finding a section of this system, i.e. a trivialization of it. Any closed manifold has a local coefficient fundamental class $[M] \in H_n(M; \mathscr{O})$ defined so that its image under $M \to (M, M \setminus \{x\})$ is the unit in $\mathscr{O}_x$. Capping with this fundamental class will then give an isomorphism,

$$H^k(M; \mathscr{G}) \cong H_{n-k}(M; \mathscr{G} \otimes \mathscr{O}),$$

for any local coefficient system $\mathscr{G}$. This can be proved just by inserting local coefficients into our proof above. △

*Intersection Theory.* We are suddenly going to be quite fast and loose with our arguments, favouring some basic geometric intuition. This is well in the spirit of intersection theory, which personally I feel is one of the crown jewels of algebraic topology.

First we briefly discuss a concept that will appear in greater detail in later chapters when we study vector bundles. Suppose $V$ is a rank $r$ oriented vector bundle over a closed base $n$-manifold. Let $DV$ be the corresponding unit disk bundle of $V$ with respect to some metric (or just some tubular neighbourhood of the zero-section of $V$) and $SV = \partial DV$ the unit sphere bundle.

**Definition 3.44: Thom Class**

PROPOSITION 3.45. *There is a unique element $t_V \in H^r(DV, SV)$ so that for any fibre $D_xV$ of the bundle, the map*

$$H^r(DV, SV) \to H^r(D_xV, S_xV) \cong H^r(D^r, S^r),$$

*sends $t_V$ to the fundamental cohomology class $[D^r, S^r]$ (with respect to the orientation of the bundle).*

This element $t_V$ is called the *Thom class* of the bundle $V$. One can equivalently view it as living in $H^r(TV)$, where $TV$ is the *Thom space* of $V$ defined as $V$'s one point compactification.

Thom classes will allow us to answer the following basic question. Suppose $Z$ is an $n-m$ dimensional oriented submanifold of a closed oriented $n-$manifold $M$. We can define a fundamental submanifold class $[Z] \in H_{n-m}(M)$ as the image of the usual fundamental class $[Z] \in H_{n-m}(Z)$ under the map induced by inclusion. Our natural question is: what is the Poincaré dual of $[Z]$?

We are looking for $\alpha \in H^m(M)$ so that $[M] \frown \alpha = [Z]$. First, let $U$ be a tubular neighbourhood around $Z$; by a classical theorem, $U$ is diffeomorphic to a neighbourhood of the zero section of the normal bundle $\nu$ of the embedding $Z \to M$. We can instead include $Z \to U$ and consider $[Z] \in H^n(U)$. By the refinement lemma, we can thus restrict to look for $\alpha \in H^m(M, M \setminus U)$, which by excision is the same as $\alpha \in H^m(U, \partial U)$.

Again by refinement and some locality arguments for manifolds, it suffices to work locally. Pick a small neighbourhood of $Z$ homeomorphic to a disk $D^{n-m}$ over which the normal bundle $\nu$ trivializes. Taking a product with the fibre, we get a neighbourhood of $U$ of the form $D^{n-m} \times D^m$. On this neighbourhood, by excision $[M]$ becomes $[D^{n-m} \times D^m, \partial(D^{n-m} \times D^m)] = [D^{n-m}, \partial D^{n-m}] \times [D^m, \partial D^m]$ and $[Z]$ becomes $[D^m, \partial D^m]$. If we let $\widetilde{[X]}$ denote the fundamental cohomology class of $X$. It is a simple lemma that in $D^n = D^m \times D^{n-m}$,

$$[D^{n-m}, \partial D^{n-m}] \times [D^m, \partial D^m] \frown \widetilde{[D^{n-m}, \partial D^{n-m}]} = [D^n, \partial D^n].$$

We deduce that the restriction of $\alpha \in H^m(U, \partial U)$ should coincide with the fundamental cohomology class of the fibre $[D^{n-m}, S^{n-m}]$. By Proposition 3.45, this is exactly the Thom class of the normal bundle $t_\nu \in H^{n-m}(U, \partial U)$. By excision and pullback by inclusion, this becomes a class in $H^{n-m}(M)$, which we denote $t_Z$.

PROPOSITION 3.46. *Given an oriented $(n-m)$ dimensional submanifold $Z$ of an oriented $n$-manifold $M$, the Thom class $t_Z \in H^{n-m}(M)$ of the normal*

*bundle of $Z$, realized as a tubular neighbourhood of $Z \subset M$ satisfies,*

$$[M] \frown t_Z = [Z],$$

*and hence $t_Z$ is Poincaré dual to the fundamental class of $Z$. The same holds with boundary, or mod 2 in the unoriented case.*

From this we deduce the following very important theorem.

**Theorem 3.47**

Let $Z$ and $W$ be closed oriented submanifolds of a closed oriented manifold $M$. Perturb $Z$ and $W$ so that they intersect transversally. By convention, we orient $Z \cap W$ using the orientation of $Z$ and the coorientation of $W$. Then,

$$t_Z \smile t_W = t_{Z \cap W},$$

or equivalently,

$$PD([Z]) \smile PD([W]) = PD([Z \cap W]).$$

The same holds modulo 2 in the unoriented case.

Proof. Let $U$ and $V$ be tubular neighbourhoods of $Z$ and $W$ respectively. Suppose $Z$ has codimension $r$ and $W$ has codimension $s$. We have $t_Z \in H^r(M, M \setminus U)$ and $t_W \in H^s(M, M \setminus V)$. Thus,

$$t_Z \smile t_W \in H^{r+s}(M, M \setminus (U \cap V)) \cong H^{r+s}(U \cap V, \partial(U \cap V)),$$

where we use excision. As before, by refinement and locality arguments, it suffices to analyze this cup product locally. Let $D^{n-r-s}$ be a small open disk in $Z \cap W$, over which all the normal bundles trivialize. The restrictions of the normal bundles $U, V$ over this disk have the form,

$$D^{n-r-s} \times D^r \quad \text{and} \quad D^{n-r-s} \times D^s$$

respectively. By transversality, the normal bundle to the embedding $Z \cap W \to M$ restricted to this disk has the form $D^{n-r-s} \times D^r \times D^s$.

By a similar computation to above for fundamental cohomology classes of disks inside $D^r \times D^s$,

$$[\widetilde{D^r, \partial D^r}] \smile [\widetilde{D^s, \partial D^s}] = [\widetilde{D^r \times D^s, \partial(D^r \times D^s)}].$$

Noting that the Thom classes of $Z$, $W$, and $Z \cap W$ restricted to this neighbourhood correspond to the fundamental classes of the fibres of the respective embeddings by Proposition 3.45, we conclude $t_Z \smile t_W = t_{Z \cap W}$. □

This result is powerful for computing ring structures.

EXAMPLE 3.48. The cellular complex of $\mathbb{C}P^n$ has a single generator in even gradings and is zero in odd gradings. Thus the differential is trivial and we have,

$$H_k(\mathbb{C}P^n) = \begin{cases} \mathbb{Z} & 0 \leq k \leq 2n, \text{ even} \\ 0 & \text{otherwise.} \end{cases}$$

By Poincaré duality, the cohomology is the same. Note from the cellular complex that a hyperplane $\mathbb{C}P^k$ for $k < n$ is a closed submanifold of dimension $2k$ which represents the generator of $H_{2k}(\mathbb{C}P^n)$. If we take two generic hyperplanes of codimensions $2n-2k$ and $2n-2j$, their intersection will be a hyperplane of codimension $4n-2k-2j$. We conclude from the above theorem that,

$$PD([\mathbb{C}P^k]) \cap PD([\mathbb{C}P^j]) = PD([\mathbb{C}P^{j+k-n}]).$$

We easily deduce that the cohomology ring is,

$$H^*(\mathbb{C}P^n) = \mathbb{Z}[x]/(x^{n+1}), \quad |x| = 2.$$

Here $x$ is Poincaré dual to the fundamental class of a hyperplane $[\mathbb{C}P^{n-1}]$.

The same computation holds in the unoriented case for real projective spaces,

$$H^*(\mathbb{R}P^n; \mathbb{Z}_2) = \mathbb{Z}_2[x]/(x^{n+1}), \quad |x| = 1.$$

Here $x$ is Poincaré dual to the fundamental class of a hyperplane $[\mathbb{R}P^{n-1}]$. ▲

*Some Geometric Applications.* The simplest application of homology is as a homotopy invariant to distinguish spaces which are not homeomorphic. But there is much more we can do with it. We outline a few of the additional invariants and results that can be proved from basic homology theory.

**Definition 3.49**

Let $C_*(X)$ be a chain complex with finitely many free and finite dimensional chain groups. The *Euler characteristic* of $C_*(X)$ is given as,

$$\chi(C_*(X)) := \sum_i (-1)^i \dim C_*(X).$$

It is a simple exercise that this equals,

$$\chi(C_*(X)) := \sum_i (-1)^i \dim H_i^{\text{free}}(X).$$

If the latter but not the former is well defined, we still call it the Euler characteristic. For the case where our chain complex is $\mathcal{C}_*(X)$ the cellular complex of a space with finite dimensional homology, we abbreviate to $\chi(X)$, which we call the Euler characteristic of $X$.

Exercise 3.50. Prove the following properties of the Euler characteristic.

(1) $\chi(M \times N) = \chi(M) \cdot \chi(N)$.

(2) If $\widetilde{M}$ is a $k$-sheeted cover of $M$, $\chi(\widetilde{M}) = k\chi(M)$.

(3) If $M, N$ are nice subsets of $X$ (so that Mayer–Vietoris applies), then,
$$\chi(M \cup N) = \chi(M) + \chi(N) - \chi(M \cap N).$$

(4) A genus $g$ surface $\Sigma_g$ has $\chi(\Sigma_g) = 2 - 2g$. Deduce Euler's formula for a convex polyhedron: $V - E + F = 2$. What is the corresponding formula for a polyhedron homeomorphic to a three-holed torus? What about for an $n$-dimensional convex polytope?

We give one important result about the Euler characteristic.

**Definition 3.51**

Let $V$ be a vector field on a smooth $n$-manifold $M$. If $x$ is an isolated zero of $V$, then in a small neighbourhood of $x$, the normalized vector field $V(x)/\|V(x)\|$ (with respect to some metric) restricted to a small sphere around $x$ defines a map $S^{n-1} \to S^{n-1}$. This map's degree is invariant under continuous deformations, and so is independent of the metric (since the set of metrics is contractible) or the choice of sphere around $x$.
The *index of* $V$ at $x$ is,
$$\operatorname{ind}_x(V) : \deg\left(V/\|V\| : S^{n-1} \to S^{n-1}\right).$$

**Theorem 3.52: Poincaré–Hopf Theorem**

Let $V$ be a vector field on a smooth closed manifold $M$ with isolated zeroes. Then,
$$\sum_{x \in \operatorname{crit}(V)} \operatorname{ind}_x(V) = \chi(M).$$
If $M$ is compact with boundary, the same holds provided $V$ points outward along the boundary.

Proof. We use Morse homology. One notes that the sum of the indices is independent of which vector field we use. There are a few ways to see this, using homotopy invariance, embedding $M$ into Euclidean space and studying the degree of the Gauss map, or realizing this sum as the pairing $\langle e(TM), [M] \rangle$. Whatever way, we may take $V$ to be a gradient vector field of a function $f$.

By locality, we may assume locally at a critical point $c$ of $f$ and a zero of $V$, $f$ looks like $f(c) - x_1^2 + \ldots - x_i^2 + x_{i+1}^2 + \ldots + x_n^2$. We may also take our metric to be standard with respect to these coordinates, so that the gradient vector field looks like,

$$V = -2\sum_{j=1}^{i} x_j \frac{\partial}{\partial x_j} + 2\sum_{k=i+1}^{n} x_k \frac{\partial}{\partial x_k}.$$

This is a linear vector field and hence it has degree plus or minus one depending on if it preserves orientation or not. It preserves orientation precisely when $i$ is even. Hence, using the definition of the Morse complex,

$$\sum_{x\in \mathrm{crit}(V)} \mathrm{ind}_x(V) = \sum_{x\in \mathrm{crit}(f)} (-1)^{\mathrm{ind}(x)} = \sum_{i=0}^{n} (-1)^i \dim(CM_i(M)) = \chi(CM_*(M)).$$

But note Morse and cellular homology agree, so $\chi(CM_*(M)) = \chi(M)$.

The boundary case follows similarly from the Morse complex for manifolds with boundary. □

Corollary 3.53 (Hairy Ball Theorem). *Every smooth vector field on an odd dimensional sphere has a zero. In particular, the odd dimensional spheres are non-parallelizable.*

**Definition 3.54**

Let $X$ be a space with finitely generated homology and $f : X \to X$ continuous. The *Lefschetz number* of $f$ is defined as,

$$\Lambda(f) = \sum_i (-1)^i \operatorname{tr}\big(f_* : H_n^{\mathrm{free}}(X) \to H_n^{\mathrm{free}}(X)\big).$$

Note this is invariant under homotopies of $f$. And if $f$ is homotopic to the identity, $\Lambda(f) = \chi(X)$.

We have the following generalization of Brouwer's fixed point theorem.

**Theorem 3.55: Lefschetz Fixed Point Theorem**

If $X$ is a finite simplicial complex and $f : X \to X$ is continuous with $\Lambda(f) \neq 0$, then $f$ has a fixed point. In particular if $X$ has the rational homotopy type of a point, then every $f : X \to X$ has a fixed point.

Proof. Assume $f$ has no fixed points. By compactness, we know that $d(x, f(x))$ obtains a positive minimum $\varepsilon$. By subdivision, we may assume that every simplex in $X$ has diameter smaller than $\varepsilon/2$ and so $f$ maps every point

out of its simplex. Analogous to the cellular approximation theorem, we prove by the Lebesgue number lemma that after subdivision of the codomain, we may approximate a continuous map $f$ of simplices by a simplicial map $g$ so that for a simplex $\sigma$, $f(\sigma)$ lies in the simplices containing $g(\sigma)$.

Apply this simplicial approximation to approximate $f$ by a simplicial map $g$. Note that for any simplex $\sigma$ and some $x \in \sigma$, all of $g(\sigma)$ lies within $\varepsilon/2$ from $f(x)$ and all $\sigma$ lies within $\varepsilon/2$ from $x$. Since $d(x, f(x)) \geq \varepsilon$, we conclude $g(\sigma) \cap \sigma = \varnothing$.

The induced map $g_*$ on the simplicial complex is described by the matrix of coefficients in our simplicial map. But since $g(\sigma) \cap \sigma = \varnothing$, the diagonal coefficients are zero. Hence the trace of $g_*$ on the cell complex is zero. As with the Euler characteristic, the alternating sum of traces of $g_*$ is invariant under passing to homology and hence $\Lambda(f) = \Lambda(g) = 0$. □

There is the following theorem which generalizes both Lefschetz fixed point and Poincaré–Hopf. We state it for interest, without proof.

**Theorem 3.56: Lefschetz–Hopf**

Let $X$ be a finite simplicial complex and $f : X \to X$ have finitely many fixed points. Then,

$$\Lambda_f = \sum_{x \in \mathrm{Fix}(f)} \mathrm{ind}_x(f),$$

where the index of $f$ at a fixed point is defined as for the zero of a vector field.

We prove one more very classical result using the full structure of the cohomology ring.

**Theorem 3.57: Borsuk–Ulam**

Any continuous map $f : S^n \to \mathbb{R}^n$ agrees on some pair of antipodes. That is there is $x \in S^n$ so that $f(-x) = f(x)$.

Proof. The case $n = 1$ is an easy consequence of the intermediate value theorem.

Now suppose for contradiction there is an $f : S^n \to \mathbb{R}^n$ that does not agree on any antipodes. Setting $g(x) = f(x) - f(-x)$, we obtain an odd function on $S^n$ with no zeroes. Thus $h(x) = g(x)/|g(x)|$ is a well defined odd continuous function $S^n \to S^{n-1}$. We claim such a map cannot exist.

In the case $n = 2$, we have an odd continuous function $S^2 \to S^1$. Restricting to the equator, gives an odd map $t : S^1 \to S^1$. Note that $t(\pi)$ is antipodal to $t(0)$ and so $t$ must traverse an angle of $\pi + 2\pi k$ on $[0, \pi]$ for some $k \in \mathbb{Z}$. Since $t$ is odd, it traverses the same path on $[\pi, 2\pi]$ as on $[0, \pi]$ and so $t$ traverses an angle of $(1 + 2k)2\pi$ after a full revolution. Hence the map $t$ has an odd winding number. But on the other hand, $t$ extends to a map on the upper hemisphere of $S^2$ and hence is null homotopic. This is a contradiction with the fact $t$ represents an odd element in $\pi_1(S^1) \cong \mathbb{Z}$.

For $n > 2$, we can use the fact $h$ is odd to quotient by antipodal identification to get a map $h' : \mathbb{R}P^n \to \mathbb{R}P^{n-1}$. Note that $H_1(\mathbb{R}P^n; \mathbb{Z}_2)$ is generated by $[\mathbb{R}P^1]$, which we can represent by a standard embedding $S^1 \to \mathbb{R}P^n$. Lifting to $S^n$, this map is a path half way along an equator. Composing with the map $h$, since it's odd, we get a path in $S^{n-1}$ connecting antipodes. Hence, under quotienting, it becomes a non-contractible loop in $\mathbb{R}P^{n-1}$ which thus represents $[\mathbb{R}P^1]$ the generator of $H_1(\mathbb{R}P^{n-1}; \mathbb{Z}_2)$.

We conclude that $h'$ induces an isomorphism on $H_1$. And thus by dualizing its pullback induces an isomorphism on $H^1$. Recall the structure of the cohomology rings,

$$H^*(\mathbb{R}P^n) = \mathbb{Z}[a]/(a^{n+1}) \quad \text{and} \quad H^*(\mathbb{R}P^{n-1}) = \mathbb{Z}[b]/(b^n).$$

We have just shown $(h')^* b = a$. By naturality, since $b^n = 0$,

$$0 = (h')^* b^n = ((h')^* b)^n = a^n.$$

This is a contradiction. □

**3.5. Cohomology of Some Common Spaces.** For reference, we list the homology and cohomology rings for some important spaces. These can be computed either using cellular homology, or spectral sequences. We'll revisit some of the spectral sequence computations when we discuss them later.

Some Cohomology Groups

**Spheres:**

$$H^*(S^n) = \mathbb{Z}[\alpha]/\alpha^2 \quad |\alpha| = n, \alpha = [S^n].$$

**Projective Spaces:** As we found,

$$H^*(\mathbb{C}P^n) = \mathbb{Z}[a]/(a^{n+1}) \quad |a| = 2, a = PD[\mathbb{C}P^{n-1}]$$
$$H^*(\mathbb{R}P^n; \mathbb{Z}_2) = \mathbb{Z}_2[b]/(b^{n+1}) \quad |b| = 1, b = PD[\mathbb{R}P^{n-1}].$$

We find over $\mathbb{Z}$,

$$H^i(\mathbb{R}P^n) = \begin{cases} \mathbb{Z} & i = 0 \text{ and } n \text{ if } n \text{ is odd} \\ \mathbb{Z}_2 & 0 < i < n \text{ even} \\ 0 & \text{else} \end{cases}$$

**Grassmannians:** Over $\mathbb{C}$, the differentials in the cell complex for the Schubert cell structure are trivial. So $H^{2i}(\mathbb{C}G(n,k))$ is the free abelian group generated by the $i$ square Young diagrams confined in the $k \times (n-k)$ rectangle and the odd cohomology groups are zero.

Over $\mathbb{R}$, we can deduce the homology from the following description of attaching maps. Suppose $\Delta, \Delta'$ are two Young diagrams of size $i$ and $i-1$ respectively which we identify with Schubert cells in $G(n,k)$. If $\Delta' \not\subset \Delta$, then the attaching map has degree zero. Otherwise, if $\Delta$ is $\Delta'$ union the square $(s,t)$, then the attaching map has degree,

$$\langle \partial\Delta, \Delta' \rangle = \begin{cases} 0 & : s+t \equiv 0 \pmod 2 \\ \pm 2 & : s+t \equiv 1 \pmod 2. \end{cases}$$

Whether the coefficient is plus or minus two can I think usually/always be fixed by the condition that $\partial^2 = 0$. In particular, the $\mathbb{Z}_2$ homology is just freely generated by Young diagrams as in the complex case.

**Surfaces:** By Mayer–Vietoris, we can deduce the effect on homology of attaching a handle. We can geometrically understand the cup product. We deduce for an orientable surface of genus $g$ $\Sigma_g$,

$$H^*(\Sigma_g) = \mathbb{Z}[a_1, b_1, \ldots, a_g, b_g, c]/(a_i a_j, b_i b_j, a_i b_j - \delta_{ij} c),$$

where $|a_i| = |b_i| = 1, |c| = 2$. The homology of the non-orientable surface $S_k$, the connect sum of $k$ copies of $\mathbb{R}P^2$, is

$$H^0(S_k) = \mathbb{Z}, \quad H^1(S_k) = \mathbb{Z}^{k-1} \oplus \mathbb{Z}_2, \quad \text{and} \quad H^2(S_k) = 0.$$

**Lie Groups:** From spectral sequences we will deduce,

$$H^*(\mathrm{SU}(n)) = \Lambda^*_{\mathbb{Z}}[x_3, x_5, \ldots, x_{2n-1}]$$
$$H^*(\mathrm{U}(n)) = \Lambda^*_{\mathbb{Z}}[x_1, x_3, \ldots, x_{2n-1}]$$
$$H^*(\mathrm{Sp}(n)) = \Lambda^*_{\mathbb{Z}}[x_3, x_7, \ldots, x_{4n-1}].$$

The cohomology of $\mathrm{SO}(n)$ is harder. Here are some results,

$$H^{*,\text{free}}(\mathrm{SO}(2k+1)) = \Lambda^*_{\mathbb{Z}}[x_3, x_7, \ldots, x_{4k-1}]$$
$$H^{*,\text{free}}(\mathrm{SO}(2k)) = \Lambda^*_{\mathbb{Z}}[a_1, a_7, \ldots, a_{4k-5}, a_{2k-1}]$$
$$H^*(\mathrm{SO}(n); \mathbb{Z}_2) = \bigotimes_{0<i<n \text{ odd}} \mathbb{Z}_2[b_i]/(b_i^{p_i}),$$

where $|b_i| = i$ and $p_i = \min\{2^k : |b_i|^{2^k} \geq n\}$.

## 4. Homotopy Theory

We now move on to the other main set of algebro-topological invariants of spaces: their homotopy groups. While conceptually simpler, the homotopy groups will not admit the same kind of algorithmic formulation as homology; even the homotopy groups of spheres are not fully known! Nevertheless, the ideas of homotopy theory will vastly expand our understanding of topology and the tools we have for the classification of spaces.

### 4.1. Homotopy Groups.

**Definition 4.1**

Given a pointed space $(X, x_0)$, its $n$th *homotopy group* as a set is given by pointed homotopy classes of maps from the $n$-sphere, called *spheroids*:

$$\pi_n(X, x_0) := [(S^n, *), (X, x_0)].$$

If $X$ is path connected, this is independent of basepoint, and so we usually ignore $x_0$ in our notation. The group structure on $\pi_n$ can be described as follows. Given two pointed maps $f, g : S^n \to X$, we may define $f * g$ by taking $S^n$, contracting the equator to a point, and mapping the two hemispheres by $f$ and $g$ respectively.

The group $\pi_n$ has two equivalent descriptions. We can instead view $\pi_n(X)$ as maps $[(S^n, S^n \setminus B), (X, x_0)]$, where $B$ is a ball inside $S^n$. Then the group operation is given by cutting out two balls from $S^n$ and identifying each with one of the composed elements. We may also view $\pi_n(X)$ as $[(I^n, \partial I^n), (X, x_0)]$. Then the group operation is given by gluing two cubes $I^n$ together and mapping on each by one of the composed elements. Our three perspectives are shown below in Figure 2.

Proposition 4.2. *For $n \geq 2$, the homotopy group $\pi_n(X)$ is abelian.*

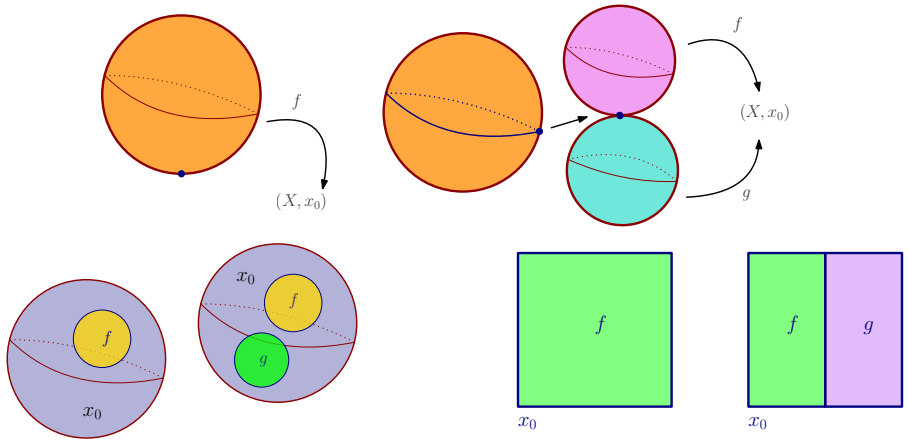



FIGURE 2. Three Models of $\pi_n(X, x_0)$

PROOF. If we use the second picture, it is tautological. There is no preferred ordering of the balls on which $f$ and $g$ are defined inside $f * g$. The argument using our cube picture is important and worth seeing. An illustration of the proof is shown below in Figure 3. □

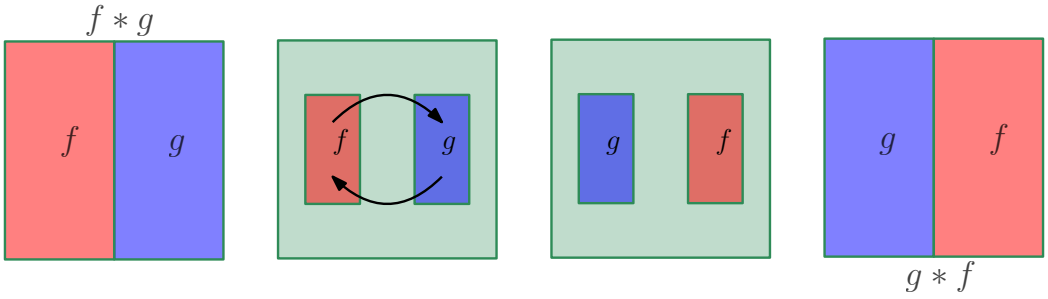



FIGURE 3. Proof that $\pi_n$ is abelian.

EXERCISE 4.3. Prove $\pi_n(X \times Y) = \pi_n(X) \times \pi_n(Y)$. ▲

Given a continuous map $f : X \to Y$, we can precompose it with a spheroid in $X$ to obtain one in $Y$. Thus $f$ induces a map $f_* : \pi_n(X) \to \pi_n(Y)$. It is immediate that this is a group homomorphism, natural with respect to composition, and only depends on the homotopy class of $f$. In particular we obtain the following.

**Theorem 4.4**

Homotopy equivalent spaces have the same homotopy groups.

Because the definition of the homotopy groups is so simple, it does allow for some easy computations.

PROPOSITION 4.5. *For $n < m$, $\pi_n(S^m) = 0$.*

PROOF. Suppose we have a map $f : S^n \to S^m$. By the cellular approximation theorem, we may homotope $f$ to be cellular with respect to the usual cell structures on the spheres. Hence $S^n$ must land in the $n$-skeleton of $S^m$, which is just a point since $n < m$. So $f$ is null-homotopic and $\pi_n(S^m)$ is trivial. □

PROPOSITION 4.6. *If $p : \widetilde{X} \to X$ is a covering map , then $\pi_n(\widetilde{X}) = \pi_n(X)$ for $n > 1$.*

PROOF. We have a map $p_* : \pi_n(\widetilde{X}) \to \pi_n(X)$. Because $S^n$ is simply connected for $n > 1$, the map lifting lemma implies that any map $f : S^n \to X$ lifts to a map $\widetilde{f} : S^n \to \widetilde{X}$ with chosen basepoint, hence $p_*$ is surjective. And moreover, if $p_* f$ is null homotopic, the map lifting lemma implies we may lift the homotopy $S^n \times I \to X$ to $\widetilde{X}$. The uniqueness part of the map lifting lemma tells us that this homotopy must connect $f$ to the constant map and so $f$ is null homotopic as well. Hence $p_*$ is injective and thus an isomorphism. □

COROLLARY 4.7. *Any space with contractible universal cover has trivial higher homotopy groups. In particular, this applies to $S^1$ and all surfaces of genus at least one.*

**Definition 4.8**

Given a pointed pair $(X, A, x_0)$, we may define its $n$th *relative homotopy group* as,

$$\pi_n(X, A) := [(I^n, \partial I^n, \partial I^n \setminus \{0\} \times I^{n-1}), (X, A, x_0)],$$

with the same product as in the absolute case. Note absolute homotopy groups $\pi_n(X)$ are relative homotopy groups for the pair $(X, x_0)$.

We have a couple ways to visualize this as in Figure 4.

EXERCISE 4.9. Prove that $\pi_n(X, A)$ is abelian for $n > 2$ and show it need not be for $n = 2$. Show the map $\pi_2(X) \to \pi_2(X, A)$ induced by inclusion of pairs lands in the centre of $\pi_2(X, A)$. ▲

REMARK 4.10. We can make sense of $\pi_0(X)$ as homotopy classes of pointed maps $S^0 \to X$ and hence homotopy classes of maps $* \to X$. That is, $\pi_0(X)$

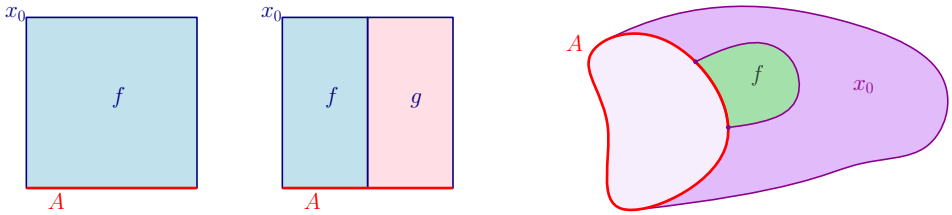



FIGURE 4. Two models for $\pi_n(X, x_0)$

is the *set* (not group) of path components of $X$. We may similarly define $\pi_1(X, A)$ as a set, although its meaning is less transparent. These sets are still compatible with the framework of homotopy groups, and they will play a role in the exact sequence we define in the next subsection.

Note that one can also define $\pi_{-1}(X)$, not as a set but as a Boolean variable, recording if $X$ is empty or not. △

### 4.2. Long Exact Sequences in Homotopy.

*Long Exact Sequence of a Pair.* We have a homotopical analogue of the exact sequence of a pair in homology.

**Theorem 4.11: Homotopy Long Exact Sequence of a Pair**

Let $(X, A)$ be a topological pair. There is a long exact sequence of homotopy groups,

$$\cdots \to \pi_n(A) \xrightarrow{i_*} \pi_n(X) \xrightarrow{j_*} \pi_n(X, A) \xrightarrow{\partial} \pi_{n-1}(A) \to \cdots .$$

Here $i_*$ and $j_*$ are induced by inclusion, and $\partial$ is the map induced by restricting relative spheroids $I^n \to X$ to the face $\{0\} \times I^{n-1} \to A$.

PROOF. Let's check exactness at $\pi_n(X)$. If $f \in \pi_n(A)$, then $i_*(f)$ lies only within $A$. But then the relative spheroid $j_* i_*(f)$ can be pushed down along the first axis to become constant. If $g \in \pi_n(X)$ has $j_*(g) = 0$, then there is a map $I^n \times I \to X$ which is constant on all the faces except $\{0\} \times I^{n-1} \times I$, where it lies in $A$. We can consider a one parameter family of angled cubes beginning from $I^n \times \{0\}$ and ending with $\{0\} \times I^n$ which defines a homotopy from $g$ to a spheroid contained in $A$. So $g$ comes from the image of $i_*$. Hence $\operatorname{im} i_* = \ker j_*$.

Let's check exactness at $\pi_n(X, A)$. If $f \in \pi_n(X)$, then $j_*(f)$ is a relative spheroid which is constantly $x_0$ on its base face. Hence $\partial j_*(f)$ is constant. If $g \in \pi_n(X, A)$ has $\partial g = 0$, then the base face of $g$ can be homotoped to be constant. By Borsuk's theorem, this homotopy can extend to all of $I^n$, and

hence $g$ is homotopic to a spheroid which is constantly $x_0$ on its base, and thus lies in $\pi_n(X)$. So $\operatorname{im} j_* = \ker \partial$.

Lastly, we check exactness at $\pi_n(A)$. If $f \in \pi_{n+1}(X, A)$, then $f : I \times I^n \to X$ defines a homotopy between the base $\partial f : \{0\} \times I^n \to X$ and the constant spheroid $\{1\} \times I^n \to X$. Hence $\partial f$ is trivial under the inclusion map $i_*$. If $g \in \pi_n(A)$ has $i_*(g) = 0$, then there is a homotopy of $I \times I^n \to X$ connecting $g$ to a constant map. This homotopy is precisely the data of a relative spheroid $I^{n+1} \to X$. Hence $\operatorname{im} \partial = \ker i_*$. □

REMARK 4.12. The tail end of this sequence looks like,

$$\cdots \to \pi_1(A) \to \pi_1(X) \to \pi_1(X, A) \to \pi_0(A) \to \pi_0(X) \to 0.$$

After the first map, these are just maps of pointed sets, not groups. But this still makes sense as an exact sequence of pointed sets. So for example a path component $A$ (an element of $\pi_0(A)$) is in the basepoint component of $X$ ($0 \in \pi_0(X)$) precisely when there is a path in $X$ from the basepoint to an element of that path component of $A$ (i.e. it's in the image of $\pi_1(X, A)$). △

EXERCISE 4.13. Prove a homotopy long exact sequence for triples. ▲

*Long Exact Sequence of a Fibration.* The homotopy long exact sequence of a pair is largely of instrumental use to give us the long exact sequence of a fibration. This is the one point on which homotopy has homology beat computationally, and we will try to milk it for all it's worth.

LEMMA 4.14. *If $p : (E, x_0) \to (B, b_0)$ is a Serre fibration and $\pi^{-1}(b_0) = F$ then,*

$$p_* : \pi_n(E, F, x_0) \xrightarrow{\cong} \pi_n(B, b_0)$$

*is an isomorphism for each $n$.*

PROOF. We first prove injectivity. Suppose $f \in \pi_n(E, F)$ has $p_* f = 0$. Then there is a homotopy $(I^n, \partial I^n) \times I \to (B, b_0)$ from $p_* f$ to a constant spheroid. By definition of a Serre fibration, this homotopy lifts to relative to $f$ and a chosen homotopy within $F$ on $\{0\} \times I^{n-1}$ to a homotopy $(I^n, \partial I^n, \partial I^n \setminus \{0\} \times I^{n-1}) \times I \to (E, F, b_0)$ of relative spheroids. By construction this homotopy is between $f$ and a map lying in $p^{-1}(b_0) = F$. Hence, $f$ is null-homotopic in $\pi_n(E, F)$.

Now we prove surjectivity. Suppose $g \in \pi_n(B, b_0)$. Again by defining property of a Serre fibration, $g$ can be lifted relatively to a map $(I^n, \partial I^n, \partial I^n \setminus \{0\} \times I^{n-1}) \to (E, F, b_0)$ which gives an element of $\pi_n(E, F)$. Since this is a lift, it maps to $g$ under $p_*$. □

Our central result now falls out for free.

**Theorem 4.15: Homotopy Long Exact Sequence of a Fibration**

Let $p : E \to B$ be a Serre fibration with fibre $F$ over the base-point. Then there is a long exact sequence,

$$\cdots \to \pi_n(F) \xrightarrow{i_*} \pi_n(E) \xrightarrow{p_*} \pi_n(B) \xrightarrow{\partial} \pi_{n-1}(F) \to \cdots,$$

where $i_*$ is induced by inclusion and $p_*$ by projection.

PROOF. Consider the homotopy exact sequence of the pair $(E, F)$. Using Lemma 4.14, we may replace the term $\pi_n(E, F)$ with $\pi_n(B)$. Based on the proof of the lemma we see the map $p_*$ is as described.

One should also check the end of the sequence:

$$\cdots \to \pi_1(B) \to \pi_0(F) \to \pi_0(E) \to \pi_0(B) \to 0$$

is exact at $\pi_0(E)$, since the term $\pi_0(B)$ is replacing something ill-defined in the long exact sequence of a pair. But this is clear, since a path component of $E$ projects to the path component of the basepoint in $B$ precisely when it contains a path component of $F$. □

REMARK 4.16. What is the map $\partial$ in the long exact sequence? We may view a spheroid $S^n \to B$ in $\pi_n(B)$ as a homotopy of spheroid $S^{n-1} \times I \to B$ from the constant spheroid to itself by sweeping out cross sections of the sphere. We may lift this homotopy to $E$ relative to one endpoint to obtain a homotopy of spheroids $S^{n-1} \times I \to E$ for which $S^{n-1} \times \{1\}$ maps to $F$ and hence defines an element of $\pi_{n-1}(F)$. This is essentially just repeating the proof of the above lemma.

EXAMPLE 4.17. This long exact sequence is our main computational tool to find homotopy groups. Right away we can see a few important applications of this result.

(a) Consider the Hopf fibration $S^1 \hookrightarrow S^3 \to S^2$. One way to see this is identifying $S^2 = \mathbb{C}P^1$ and then this fibration is just the defining quotient map of $\mathbb{C}P^1$ as points of $S^3$ modulo rescaling by complex units. We have, remembering $\pi_2(S^3) = 0$ by cell approximation,

$$\cdots \to \cancel{\pi_2(S^3)} \xrightarrow{0} \pi_2(S^2) \xrightarrow{\cong} \pi_1(S^1) \to \cancel{\pi_1(S^3)} \xrightarrow{0} \cdots.$$

Hence $\pi_2(S^2) = \mathbb{Z}$. We proved earlier that the higher homotopy groups of $S^1$ vanish and so for $n > 2$ we find,

$$\cdots \to \cancelto{0}{\pi_n(S^1)} \to \pi_n(S^3) \xrightarrow{\cong} \pi_n(S^2) \to \cancelto{0}{\pi_{n-1}(S^1)} \to \cdots .$$

Hence,

$$\boxed{\pi_n(S^3) = \pi_n(S^2) \quad \text{for} \quad n \geq 3.}$$

We will show shortly that $\pi_n(S^n) = \mathbb{Z}$ for all $n$, with homotopy classes of maps $S^n \to S^n$ classified by their degree. Hence in particular $\pi_3(S^2) = \mathbb{Z}$. The fibration long exact sequence show the unit of this group is exactly the Hopf fibration, and other elements are given by precomposing the Hopf fibration with higher degree maps of $S^3$.

(b) One has a generalized Hopf fibration $S^1 \hookrightarrow S^{2n+1} \to \mathbb{C}P^n$. And the same long exact sequence argument shows $\pi_k(S^{2n+1}) = \pi_k(\mathbb{C}P^n)$ for all $k \geq 3$ and $\pi_2(\mathbb{C}P^n) = \mathbb{Z}$.

Taking a direct limit gives the fibration $S^1 \hookrightarrow S^\infty \to \mathbb{C}P^\infty$. Because $S^\infty$ is contractible, one obtains $\pi_n(S^1) \cong \pi_{n+1}(\mathbb{C}P^\infty)$ for all $n$. Hence $\pi_2(\mathbb{C}P^\infty) = \mathbb{Z}$ and all its other homotopy groups vanish.

From quaternions and octonions one has analogous Hopf bundles $S^3 \hookrightarrow S^7 \to S^4$ and $S^7 \hookrightarrow S^{15} \to S^8$ which give some non-trivial relations of homotopy groups.

(c) Given any pointed space $(X, x_0)$, it has a "path space fibration," $\Omega X \hookrightarrow EX \to X$ given by sending a path in $X$ beginning at a fixed endpoint $x_0$ to its other endpoint. We see that the fibre over $x_0$ is given by $\Omega X$. Note $EX$ is contractible by performing a homotopy retracting paths to their first endpoint $x_0$. Hence the long exact sequence of this fibration gives for all $n$,

$$\pi_n(X) \cong \pi_{n-1}(\Omega X).$$

In the case $n = 1$, this is just the statement that homotopy classes of loops are the same as path components of the loop space. The full statement is actually just a special case of the *loop-suspension adjunction*:

$$[\Sigma X, Y] = [X, \Omega Y],$$

which holds for any pair of spaces $X, Y$. To see this is true, note that a pointed map $\Sigma X \to Y$ associates to each $x \in X$, a one-parameter family of points in $Y$ which begins and ends with the basepoint, i.e. a loop, and so we can interpret it as a map $X \to \Omega Y$. One then just has to easily check this association is compatible with homotopies.

Applying this adjunction to $S^n$ and our original space $X$ and noting $\Sigma S^n = S^{n+1}$ gives the equality of homotopy groups.

This is in turn a special case of the adjunction,

$$[X \wedge A, Y] = [X, [A, Y]].$$

EXERCISE 4.18. Check this adjunctive isomorphism of homotopy classes is compatible with the group structure of $\pi_n$. ▲

(d) The determinant map gives a fibration $\mathrm{SU}(2) \xrightarrow{\det} \mathrm{U}(2) \to S^1$. Since $\mathrm{SU}(2) \cong S^3$, we deduce from the fibration long exact sequence that $\pi_3(\mathrm{U}(2)) \cong \pi_3(S^3) = \mathbb{Z}$, where we are again using the yet unproved fact $\pi_n(S^n) \cong \mathbb{Z}$.

The map given by taking an $n \times n$ unitary matrix and mapping to its first column defines a fibration $\mathrm{U}(n-1) \hookrightarrow \mathrm{U}(n) \to S^{2n-1}$. As long as $n \geq 3$, $\pi_3(S^{2n-1}) = 0$. Hence we inductively find from the long exact sequence of the fibration that for all $n \geq 2$,

$$\pi_3(\mathrm{U}(n)) \cong \pi_3(\mathrm{U}(n-1)) \cong \cdots \cong \pi_3(\mathrm{U}(2)) = \mathbb{Z}.$$

(e) There are many more fibrations of Lie groups and homogeneous spaces that can be exploited to find homotopy groups, some we will need to apply later. Here are a few of note:

$$\begin{aligned} \mathrm{O}(n-1) &\hookrightarrow \mathrm{O}(n) &&\to S^{n-1} \\ \mathrm{U}(n-1) &\hookrightarrow \mathrm{U}(n) &&\to S^{2n-1} \\ \mathrm{Sp}(n-1) &\hookrightarrow \mathrm{Sp}(n) &&\to S^{4n-1} \\ V(n-k, m-k) &\hookrightarrow V(n,m) &&\to V(n,k) \\ \mathrm{O}(n) &\hookrightarrow V(n,k) &&\to \mathrm{Gr}(n,k). \end{aligned}$$

▲

**4.3. The Freudenthal Suspension Theorem and Stable Homotopy Theory.** Suppose $f : X \to Y$ is a continuous map. There is a natural map $\Sigma f : \Sigma X \to \Sigma Y$ induced by quotienting from the trivial product map $f \times \mathrm{id} : X \times I \to Y \times I$. Clearly we may suspend homotopies as well, so that if $f, g$ are homotopic, then $\Sigma f, \Sigma g$ are homotopic.

In the particular case of maps $S^n \to X$, we may suspend to a map $S^{n+1} \to \Sigma X$. Thus we obtain a map,

$$\Sigma : \pi_n(X) \to \pi_{n+1}(\Sigma X).$$

It is simple enough to check $\Sigma(f+g)$ is homotopic to $\Sigma(f)+\Sigma(g)$ so that this is a group homomorphism.

**Definition 4.19**

We say a space $X$ is *$n$-connected* if $\pi_k(X) = 0$ for $0 \leq k \leq n$.

**Theorem 4.20: Freudenthal Suspension Theorem**

If $X$ is $n$-connected, the suspension homomorphism,

$$\Sigma : \pi_k(X) \to \pi_{k+1}(\Sigma X)$$

is an isomorphism for $k \leq 2n$ and surjective for $k = 2n+1$. In particular,

$$\Sigma : \pi_{n+k}(S^n) \to \pi_{n+k+1}(S^{n+1})$$

is an isomorphisms for $n > k+1$ and surjective for $n = k+1$.

We will not that prove this theorem since it is a bit of an involved argument. One method is by smoothly approximating. Another is through homotopy excision. We will prove the special case for the sphere where $k = 0$, since it is necessary for what we do later.

**Theorem 4.21**

For every $n \geq 0$, $\pi_n(S^n) = \mathbb{Z}$. These homotopy classes are labelled by the degree of maps $S^n \to S^n$.

Proof. Since homotopic maps have the same degree, we have a map,

$$\pi_n(S^n) \xrightarrow{\deg} \mathbb{Z}.$$

Note given two maps $f, g : S^n \to S^n$, their product $f+g$ as an element of $\pi_n$ is given by $S^n \to S^n \vee S^n \to S^n$ where we contract the equator and then map on the pair of spheres by $f$ and $g$.

We recall $H_n(S^n_a \vee S^n_b) = \mathbb{Z}^2$ is generated by $[S^n_a]$ and $[S^n_b]$, the fundamental classes under inclusion. It is clear that the equatorial contraction map $S^n \to S^n \vee S^n$ induces a map on homology that will send $[S^n]$ to $[S^n_a]+[S^n_b]$. Almost by definition,

$$f_*[S^n_a] = \deg(f)[S^n] \quad \text{and} \quad g_*[S^n_b] = \deg(g)[S^n].$$

Hence,

$$(f+g)_*[S^n] = f_*[S^n_a] + g_*[S^n_b] = (\deg(f) + \deg(g))[S^n].$$

Thus the degrees are additive under the group operation of $\pi_n$, and so the above degree map we wrote down is a group homomorphism.

We know 1 is in the image of deg, since the identity map has degree one, and so the homomorphism is surjective.

EXERCISE 4.22. Consider $S^n \subset \mathbb{C} \times \mathbb{R}^{n-1}$ with coordinates $(z, \vec{x})$. Show for $k \geq 0$ that the maps,

$$f(z, \vec{x}) = (z^k, \vec{x}) \quad \text{and} \quad g(z, \vec{x}) = (\overline{z}^k, \vec{x})$$

have $\deg(f) = k$ and $\deg(g) = -k$. ▲

It only remains to show the homomorphism is injective. Suppose $f : S^n \to S^n$ has degree zero. We can homotope $f$ to be smooth, then a generic point $p$ with have preimages $q_1, \ldots, q_{2n}$ so that near $q_i$, $f$ is a local diffeomorphism which is orientation preserving at half of the $q_i$'s and orientation reversing on the others. Now pair up the $q_i$'s in pairs with opposite orientations and draw paths connecting the pairs. In a tubular neighbourhood of each path, we may homotope $f$ so it corresponds in this neighbourhood to the sum of 1 and $-1$ in the fundamental group of $S^n$. Hence we may homotope $f$ to be constant on this neighbourhood, and in particular not equal $p$. Doing this on each tubular neighbourhood, $p$ will not be in the image of a function homotopic to $f$. But then this map factors as $S^n \to S^n \setminus \{p\} \hookrightarrow S^n$. Since $S^n \setminus \{p\}$ is contractible, we conclude $f$ is null homotopic. □

If we take the Freudenthal Suspension Theorem on faith, it leads us to the following definition,

**Definition 4.23**

The *kth stable homotopy group* of a space $X$ is,

$$\pi_k^{\text{st}}(X) := \lim_{n \to \infty} \pi_{n+k}(\Sigma^n X).$$

Later results will imply that the $(n+1)$st suspension of any space is $n$-connected. And hence this limit stabilizes after $n = k + 2$,

$$\pi_{2k+2}(\Sigma^{k+2} X) = \pi_{2k+3}(\Sigma^{k+3} X) = \cdots = \pi_k^{\text{st}}(X).$$

In particular, we have the *stable homotopy groups of the sphere*,

$$\pi_k^{\text{st}} := \pi_{2k+2}(S^{k+2}).$$

It turns out that $\pi_k^{\text{st}}(\cdot)$ is an extraordinary homology theory in the sense of Definition 3.20. The homology groups of a point are the stable homotopy

groups of the sphere. Stable homotopy is in some sense the most fundamental extraordinary cohomology theory and largely initiated their study.

We have already computed the 0th stable homotopy group of the sphere $\pi_0^{\mathrm{st}} = \pi_2(S^2) = \mathbb{Z}$. The first stable homotopy group $\pi_1^{\mathrm{st}} = \pi_4(S^3) = \mathbb{Z}_2$ will be computed later using spectral sequences. The stable homotopy groups of a sphere are difficult to calculate, but some tools exist, for example the Adams spectral sequence and chromatic homotopy theory. The *unstable homotopy groups* $\pi_{n+k}(S^n)$ with $n \leq k+1$ are even harder. In Table 1, the first several homotopy groups of low dimensional spheres are shown; in gray are groups we know to be trivial, and the coloured diagonals are the stable groups. While there are patterns, the possible groups are varied and sometimes complicated.

| | $\pi_1$ | $\pi_2$ | $\pi_3$ | $\pi_4$ | $\pi_5$ | $\pi_6$ | $\pi_7$ | $\pi_8$ | $\pi_9$ | $\pi_{10}$ | $\pi_{11}$ | $\pi_{12}$ | $\pi_{13}$ | $\pi_{14}$ | $\pi_{15}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $S^1$ | $\mathbb{Z}$ | | | | | | | | | | | | | | |
| $S^2$ | | $\mathbb{Z}$ | $\mathbb{Z}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}_{12}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}_3$ | $\mathbb{Z}_{15}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2^2$ | $\mathbb{Z}_{12} \times \mathbb{Z}_2$ | $\mathbb{Z}_{84} \times \mathbb{Z}_2^2$ | $\mathbb{Z}_2^2$ |
| $S^3$ | | | $\mathbb{Z}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}_{12}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}_3$ | $\mathbb{Z}_{15}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2^2$ | $\mathbb{Z}_{12} \times \mathbb{Z}_2$ | $\mathbb{Z}_{84} \times \mathbb{Z}_2^2$ | $\mathbb{Z}_2^2$ |
| $S^4$ | | | | $\mathbb{Z}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z} \times \mathbb{Z}_{12}$ | $\mathbb{Z}_2^2$ | $\mathbb{Z}_2^2$ | $\mathbb{Z}_{24} \times \mathbb{Z}_3$ | $\mathbb{Z}_{15}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2^3$ | $\mathbb{Z}_{120} \times \mathbb{Z}_{12} \times \mathbb{Z}_2$ | $\mathbb{Z}_{84} \times \mathbb{Z}_2^5$ |
| $S^5$ | | | | | $\mathbb{Z}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}_{24}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}_{30}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2^3$ | $\mathbb{Z}_{72} \times \mathbb{Z}_2$ |
| $S^6$ | | | | | | $\mathbb{Z}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}_{24}$ | 0 | $\mathbb{Z}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_{60}$ | $\mathbb{Z}_{24} \times \mathbb{Z}_2$ | $\mathbb{Z}_2^3$ |
| $S^7$ | | | | | | | $\mathbb{Z}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}_{24}$ | 0 | 0 | $\mathbb{Z}_2$ | $\mathbb{Z}_{120}$ | $\mathbb{Z}_2^3$ |
| $S^8$ | | | | | | | | $\mathbb{Z}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}_{24}$ | 0 | 0 | $\mathbb{Z}_2$ | $\mathbb{Z} \times \mathbb{Z}_{120}$ |
| $S^9$ | | | | | | | | | $\mathbb{Z}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}_{24}$ | 0 | 0 | $\mathbb{Z}_2$ |

TABLE 1. Homotopy Groups of Spheres

In Table 2, is a list of the first few stable homotopy groups. Memorizing these is probably high on the list of bar tricks to impress a mathematician.

| $k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\pi_k^{\mathrm{st}}$ | $\mathbb{Z}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}_{24}$ | 0 | 0 | $\mathbb{Z}_2$ | $\mathbb{Z}_{240}$ | $\mathbb{Z}_2^2$ | $\mathbb{Z}_2 \times \mathbb{Z}_4$ | $\mathbb{Z}_6$ | $\mathbb{Z}_{504}$ | 0 | $\mathbb{Z}_3$ | $\mathbb{Z}_4$ |

TABLE 2. Stable Homotopy Groups

One important result which one could maybe guess from the above tables is the following.

PROPOSITION 4.24 (Serre's Finiteness Theorem). *The homotopy groups of the spheres $\pi_{n+k}(S^n)$ for $k \geq 0$ are finite abelian groups except for:*

*(1) $k = 0$ in which case $\pi_n(S^n) = \mathbb{Z}$.*

*(2) $k = 2m$ and $n = 2m-1$ in which case,*

$$\pi_{4m-1}(S^{2m}) = \mathbb{Z} \oplus F_m,$$

*for $F_m$ finite abelian.*

*In particular, the stable homotopy groups $\pi_k^{\mathrm{st}}$ are finite for $k \geq 1$.*

We will prove this result later using spectral sequences and describe a free generator of $\pi_{4m-1}(S^{2m})$.

This is the beginning of *stable homotopy theory*, in which spaces are *stabilized* by repeated suspension and homotopy groups are studied in this direct limit. In the modern perspective on algebraic topology, spaces are replaced by *spectra*, which encapsulate this direct limit under stabilization. Spectra turn out to be extremely important because it turns out extraordinary cohomology theories are all representable functors and their representing objects are precisely spectra; this is the Brown representability theorem. We will discuss some other generalized cohomology theories and their associated spectra later in these notes.

### 4.4. Homotopy and CW Complexes.

*Whitehead Theorem.* At this stage, a reasonable question to ask is: how much of the homotopy type of a space is captured by its homotopy groups? A partial answer is that for CW complexes the homotopy groups are almost enough.

Recall we said that spaces $X$ and $Y$ weak homotopy equivalent if there is a natural bijection $[Z, X] \to [Z, Y]$ for any CW complex $Z$. A special case of this, when the bijection is induced by a map, is described in the following theorem.

**Theorem 4.25**

Let $f : X \to Y$ be a continuous map. The following are equivalent.

(i) $f_* : [Z, X] \to [Z, Y]$ is a bijection for any CW complex $Z$.

(ii) $f_* : \pi_n(X) \to \pi_n(Y)$ is an isomorphism for all $n$.

(iii) For any CW pair $(Z, W)$ with maps $h, g$ satisfying the following commutative diagram,

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ {\scriptstyle h}\uparrow & & \uparrow{\scriptstyle g} \\ W & \hookrightarrow & Z, \end{array}$$

there exists a map $\widetilde{h} : Z \to X$ extending $h$ so that $f \circ \widetilde{h} \sim g$.

In this case we call $f$ a *weak homotopy equivalence.*

Proof. $(i) \implies (ii)$: Taking $Z = S^n$ gives the result.

$(iii) \implies (i)$: Take pair $(Z, \varnothing)$. This shows $f_*$ is surjective. Take pair $(Z \times I, Z \times \{0\} \sqcup Z \times \{1\})$. Given $\varphi, \psi : Z \to X$ homotopic in $Y$, we thus obtain a homotopy in $X$. So, $f_*$ is injective.

$(ii) \implies (iii)$: Let $C_f$ be the mapping cylinder of $f$. This is homotopy equivalent to $Y$ and so from the long exact sequence of the pair $(C_f, X)$ and the fact $f_*$ is an isomorphisms on homotopy groups, we conclude $\pi_n(C_f, X) = 0$ for all $n$.

Suppose we have a CW pair $(Z, W)$ and maps $h : W \to X$ and $g : Z \to Y$ so that $g|_W = f \circ h$. Because $(Z, W)$ is built out of cells and because $\pi_n(C_f, X) = 0$, the map $i \circ g : Z \to C_f$ on cells of $Z \setminus W$ can be homotoped to a map $\widetilde{h}$ lying in $X \subset C_f$. By construction, composing $\widetilde{h}$ with $f$ gives a map homotopic to $g$. □

This has the following crucial corollary.

**Theorem 4.26: Whitehead Theorem**

Let $f : X \to Y$ be a weak homotopy equivalence between CW complexes. Then $f$ is a homotopy equivalence.

Proof. This proof is the same as Theorem 1.3. By Theorem 4.25, $f$ induces a bijection $[Z, X] \to [Z, Y]$. Applying this to $X$ and $Y$, since they are CW complexes, shows $f_* : [X, X] \to [X, Y]$ and $f_* : [Y, X] \to [Y, Y]$ are bijections. Let $g = (f_*)^{-1}(\mathrm{id}_Y)$. This implies $f_*(g) = f \circ g \sim \mathrm{id}_Y$. On the other hand, $f_*(g \circ f) = (f_* g) \circ f \sim \mathrm{id}_Y \circ f \sim f_*(\mathrm{id}_X)$. Since $f_*$ is a bijection, $g \circ f \sim \mathrm{id}_X$. So $f$ and $g$ determine a homotopy equivalence. □

Note that it is not enough that $X$ and $Y$ have the same homotopy groups, we need a map realizing this isomorphism in order to conclude the spaces are homotopy equivalent. The following exercise demonstrates this.

Exercise 4.27. Show that $S^2$ and $S^3 \times \mathbb{C}P^\infty$ have the same homotopy groups but are not weak homotopy equivalent. Show the same for $\mathbb{R}P^3 \times S^2$ and $S^3 \times \mathbb{R}P^2$. ▲

We can do even worse than this. The spaces $S^2 \times S^2$ and $\mathbb{C}P^2 \# \overline{\mathbb{C}P^2}$ have the same homotopy and homology groups, but are not homotopy equivalent (which can be seen by studying the cohomology ring). Worse still, there are two $S^3$ bundles on $S^2$. They have the same homotopy groups and cohomology ring but are not homotopy equivalent (which can be seen by showing the non-trivial bundle is not spin and applying the Wu formula).

*Attaching Cells.* We need several preliminary results dealing with how attaching cells to CW complexes affects homotopy. These will be applied in the next few sections.

PROPOSITION 4.28. *If $X$ is an $n$-connected CW complex, it is homotopy equivalent to a CW complex with one 0-cell and no cells of dimension 1 through $n$.*

PROOF. We work by induction. In the base case, if $n$ is 0-connected, it is path connected. Thus for any pair of 0-cells, there is a path connecting them, which by cellular approximation can be taken to be a one cell joining the 0-cells. We may identify this 1-cell to a point and the resulting cell structure will be homotopy equivalent with our two 0-cells identified. Doing this inductively for all pairs of 0-cells and using the weak topology gives a homotopy equivalent complex with only one 0-cell.

Now suppose that $X$ is $n$-connected with one 0-cell and no cells of dimension 1 through $n-1$. Thus the $n$-skeleton of $X$ is a wedge sum of $n$-spheres. Because $X$ is $n$-connected, each such $n$-sphere is contractible in $X$. By cellular approximation, this homotopy $S^n \times I \to X$ to a point defines a disk $D^{n+1}$ lying in the $n+1$ skeleton of $X$. We can form a CW complex $\widetilde{X}$ by adding a disk $D^{n+2}$ bounded by two copies of $D^{n+1}$ for each $n$-sphere in the $n$-skeleton of $X$. By construction, $D^{n+2}$ is contractible and so $\widetilde{X}$ is homotopy equivalent to $X$. If $A$ is the wedge-product of these $(n+2)$-disks, we have that $X \sim \widetilde{X}/A$ and $\widetilde{X}/A$ has no $n$-cells. □

As we probably should have mentioned before, but only use now, there is a $\pi_1(X)$ action on the higher homotopy groups of $X$ given by composing a spheroid with some loop.

**Theorem 4.29: Cell-Attaching Lemma**

Let $X$ be a CW complex and $Y$ a complex obtained from $X$ by attaching a disk $D^{n+1}$ with $n \geq 2$ via $\varphi : \partial D^{n+1} \to X_n$. Then the inclusion $i : X \to Y$ induces an isomorphism on homotopy groups $\pi_q$ for $q < n$. For $q = n$, it induces a surjection with kernel generated by the $\pi_1(X)$ orbit of the homotopy class of the attaching map $[\varphi]$.

PROOF. (Sketch) Given a spheroid $S^q \to Y$, if $q \leq n$, cellular approximation implies it can be homotoped to lie in $X$. Hence the map $\pi_q(X) \to \pi_q(Y)$ is surjective. If a spheroid $S^q \to X$ is contractible under inclusion into $Y$ and $q < n$, then cellular approximation implies the contraction $D^q \to Y$ can be

homotoped to lie in $X$. Hence the spheroid $S^q \to X$ was contractible to begin with and $\pi_q(X) \to \pi_q(Y)$ is injective.

It only remains to check the kernel condition for $q = n$. Suppose $f : S^n \to X$ is non-trivial and in the kernel of $i_*$ so that it extends to a contraction $\widetilde{f} : D^{n+1} \to Y$. We may homotope $\widetilde{f}$ to be smooth. Since $f$ is not null-homotopic, $\widetilde{f}$ must land at some point in $D^{n+1} \subset Y$. Looking at a preimage of a neighbourhood of regular value of $\widetilde{f}$ in $D^{n+1}$ gives a collection of neighbourhoods in $D^{n+1}$. We can homotope $\widetilde{f}$ so the domain is a bouquet of spheres each containing one of the preimage neighbourhoods. By retracting this neighbourhood onto the entire sphere, and simultaneously onto all of $D^{n+1} \subset Y$, $\widetilde{f}$ becomes a connected sum of maps homotopic to $\varphi$ up to composition with a loop on the boundary. □

Corollary 4.30. *Let $X$ and $Y$ be CW complexes.*

*(i)* $\pi_n(X) = \pi_n(X_{n+1})$.

*(ii) If $X$ and $Y$ are $p$ and $q$ connected respectively for $p, q \geq 1$, then,*

$$\pi_n(X \vee Y) = \pi_n(X) \oplus \pi_n(Y) \quad \text{for} \quad n \leq p + q.$$

*(iii)* $\pi_n(\bigvee_m S^n) = \mathbb{Z}^m$ *for $n \geq 2$ and $m \in \mathbb{N} \cup \infty$, with generators corresponding to the inclusion map of spheres in the bouquet.*

Proof. (i) This is immediate from the cell-attaching lemma.

(ii) By Proposition 4.28, we may suppose $X \vee Y$ has one 0-cell at the joint vertex and $X, Y$ have no cells of dimensions 1 through $p$ and 1 through $q$ respectively. Note that $X \times Y \setminus X \vee Y$ consists only of cells of dimension at least $p + q + 2$. Hence the result follows from cell-attaching.

(iii) This is immediate from (ii). □

*CW Approximation.* We have seen that homotopy captures almost all the information about homotopy type of CW complexes. We now have a partial converse in that homotopy groups and weak homotopy equivalence captures only as much about a space as can be encoded in a CW complex. The upshot of this is that it means often one can conclude a result about homotopy or homology holds for all spaces just by checking for CW complexes.

**Theorem 4.31: CW Approximation Theorem**

For any space $X$ there is a CW complex $X^{CW}$ with a weak homotopy equivalence $f : X^{CW} \to X$ called a *CW approximation* to $X$. The CW approximation of $X$ is unique up to (homotopically unique) homotopy equivalence. Moreover, two space are weak homotopy equivalent if and only if they have the same CW approximation.

PROOF. We construct $X^{CW}$ inductively as a direct limit,

$$X_0^{CW} \subset X_1^{CW} \subset X_2^{CW} \subset \cdots \subset X^{CW}.$$

Let $X_0^{CW}$ be a point and $f_0 : X_0^{CW} \to X$ the map to the basepoint. Suppose we have $f_{n-1} : X_{n-1}^{CW} \to X$ which induces an isomorphism on homotopy group $\pi_q$ for $q \leq n-1$.

Let $Y_n = X_{n-1}^{CW} \bigvee_\alpha S_\alpha^n$, with $\alpha$ indexing a generating set for $\pi_n(X)$. We may define a map $g_n : Y_n \to X$ which acts as $f_{n-1}$ on $X_{n-1}^{CW}$ and as a spheroids $\varphi_\alpha$ representing the indexed generators of $\pi_n(X)$ on each $S_\alpha^n$. Then, by the cell attaching lemma, $\pi_q(X)$ and $\pi_q(Y_n)$ are the same for $q \leq n-1$ with isomorphisms induced by $g_n$. Also the map,

$$(g_n)_* : \pi_n(Y_n) \to \pi_n(X)$$

is surjective since its image contains a generating set $\{\varphi_\alpha\}$ for $\pi_n(X)$. Now let $\{\psi_\beta : \partial D^{n+1} \to Y_n\}$ be a generating set for the kernel of $(g_n)_*$ so that $g_n \circ \psi_\beta$ extends to a map $\widetilde{\psi_\beta} : D^{n+1} \to X$ for each $\beta$. Set $X_n^{CW} = Y_n \cup_\beta D_\beta^{n+1}$ glued by $\psi_\beta$. Define a map $f_n : X_n^{CW} \to X$ which is $g_n$ on $Y_n$ and $\widetilde{\psi_\beta}$ on $D_\beta^{n+1}$. By the cell-attaching lemma, $f_n$ induces isomorphisms on $\pi_q$ for $q \leq n-1$, and actually also for $q = n$, since the generators of $\ker((g_n)_*)$ are trivial in $\pi_n(X_n^{CW})$.

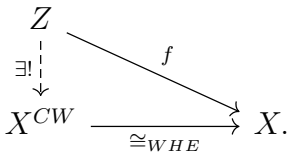
To check uniqueness, suppose $f_1 : X_1^{CW} \to X$ and $f_2 : X_2^{CW} \to X$ are two cellular approximations. These define weak homotopy equivalences, which by transitivity, tells us $X_1^{CW}$ and $X_2^{CW}$ are weak homotopy equivalent. It is an easy extension of Whitehead's theorem that weak homotopy equivalent CW complexes are homotopy equivalent, as desired.

Lastly, suppose $X$ and $Y$ have the same CW approximation $Z$. Then $X$ and $Z$ are weak homotopy equivalent and $Y$ and $Z$ are weak homotopy equivalent. By transitivity of weak homotopy equivalence, so are $X$ and $Y$. Conversely, suppose $X$ and $Y$ are weak homotopy equivalent. And consider CW approximations $Z_1 \to X$ and $Z_2 \to Y$. These define weak homotopy equivalences and

so again by transitivity, $Z_1$ and $Z_2$ are weak homotopy equivalent, and since they are CW, they are homotopy equivalent. □

Remark 4.32. From this theorem we can infer another definition of a CW approximation in terms of a universal property.

The CW approximation $X^{CW} \to X$ of a space $X$ is the unique up to homotopy equivalence space so that any map $f : Z \to X$ for $Z$ a CW complex uniquely factors in **hTop** through $X^{CW}$. That is, we obtain the following diagram commuting up to homotopy:




In more categorical language, we can say that CW approximation defines a functor **hTop** $\to$ **hCW** which is the right adjoint to the inclusion functor **hCW** $\hookrightarrow$ **hTop**. △

*Homotopy and Homology: The Hurewicz Theorem.* We will now study the first non-trivial homotopy group of a CW complex and compare it to homology. We need the following proposition first.

Proposition 4.33. *A weak homotopy equivalence $f : X \to Y$ induces an isomorphism of homology groups.*

Proof. Let $\Sigma$ be an $n$ dimensional CW complex representing a given $n$ dimensional singular cycle in $Y$ under a map $\sigma : \Sigma \to Y$. By definition of weak homotopy equivalence, there is a unique homotopy class of map $\widetilde{\sigma} : \Sigma \to X$ so that $\sigma = f \circ \widetilde{\sigma}$. This map defines an $n$-cycle in $X$ mapping to $\sigma$ under $f_*$, and so $f_*$ is surjective.

Now suppose $\sigma : \Sigma \to X$ is an $n$-cycle so that $f_*(\sigma) = \partial\tau$ for $\tau$ a singular $n$-chain. We can thus extend $\Sigma$ to an $n+1$ complex $T$ so that $T \to Y$ extends $f \circ \sigma$ and represents $\tau$. By Proposition 4.25 in its relative incarnation, we can lift $\Sigma \to X$ to $T \to X$ and so $\sigma$ is a boundary in $X$. □

**Definition 4.34**

Given a space $X$ and a class of spheroid $\varphi : S^n \to X$, we may define an element

$$h([\varphi]) = \varphi_*([S^n]) \in H_n(X).$$

This only depends on the homotopy class of $\varphi$, and it's easy to see that adding spheroids in $\pi_n$ will sum their image under $h$ in $H_n$. Hence we obtain a group homomorphism,

$$h : \pi_n(X) \to H_n(X),$$

called the *Hurewicz homomorphism*.

**Theorem 4.35: Hurewicz Theorem**

For any space $X$, the Hurewicz homomorphism $h : \pi_1(X) \to H_1(X)$ is surjective, with kernel the commutator subgroup of $\pi_1(X)$. Moreover, if $X$ is simply connected and $n \geq 2$ then the following are equivalent:

(i) $\pi_k(X) = 0$ for $k = 1, \ldots, n-1$,

(ii) $H_k(X) = 0$ for $k = 1, \ldots, n-1$.

In this case, $h : \pi_n(X) \to H_n(X)$ is an isomorphism.

PROOF. Without loss of generality, $X$ is a CW complex. If not, we can CW approximate and use that homology and homotopy are invariants of weak homotopy equivalence.

Suppose $X$ is $(n-1)$-connected for $n > 1$. By Proposition 4.28, we may assume $X$ has a single 0-cell and no other cells of dimension less than $n$. In this case, the chain groups $C_k(X)$ are zero for $k < n$. So, $H_n(X)$ is the free abelian group generated by the $n$-cells of $X$ quotiented by the image of the boundary of $(n+1)$-cells. Meanwhile by the cell attaching lemma, $\pi_n(X)$ is the free abelian group generated by the $n$-cells of $X$ quotiented by the $\pi_1(X)$ orbit of the attaching spheroids of $(n+1)$-cells. Since $\pi_1(X) = 0$, this is just the free abelian group generated by attaching maps. But expressing the attaching map of an $(n+1)$-cell as a sum of copies of the inclusion of each $n$-cell is easily seen to be the same as the boundary map on that $(n+1)$-cell in homology. Hence our descriptions of $\pi_n(X)$ and $H_n(X)$ are the same and so these groups are isomorphic. Since the Hurewicz homomorphism sends the generators of $\pi_n(X)$ to the generators of $H_n(X)$ it induces this isomorphism.

In the case where $n = 1$, everything holds verbatim except that $\pi_1(X)$ is generated instead of abelian generated by 1-cells. Hence to recover $H_1(X)$, one has to abelianize, after which we obtain an isomorphism. The Hurewicz homomorphism must factor through the abelianization of $\pi_1(X)$ by universal properties, after which it sends generators to generators and thus again must give an isomorphism.

Finally, to show (ii) implies (i), suppose $X$ is simply connected $H_k(X) = 0$ for $k = 1, \dots, n-1$. If $\pi_k(X) \neq 0$ for any of these groups $k > 1$, what we have just established shows $H_k(X) = \pi_k(X) \neq 0$. Since we know $H_k(X) = 0$, we conclude $\pi_k(X) = 0$ for $k = 1, \dots, n-1$. □

This result allows us to prove a homological version of the Whitehead theorem (assuming we know the Serre spectral sequence).

**Theorem 4.36: Homological Whitehead Theorem**

Suppose $f : X \to Y$ is a continuous map between simply connected spaces. If $f_* : H_*(X) \to H_*(Y)$ is an isomorphism, then $f$ is a weak homotopy equivalence. More generally, the following are equivalent:

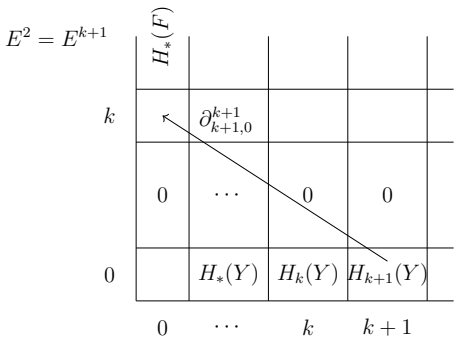
(i) $f_* : \pi_k(X) \to \pi_k(Y)$ is an isomorphism for $k < n$ and surjective for $k = n$.

(ii) $f_* : H_k(X) \to H_k(Y)$ is an isomorphism for $k < n$ and surjective for $k = n$.

Proof. We can homotope $f : X \to Y$ to a fibration with homotopy fibre $F$. From the long exact sequence of a fibration we conclude that condition (i) in the theorem is equivalent to $\pi_{k-1}(F) = 0$ for $k \leq n$.

Because $\pi_1(Y) = 0$, our fibration is homologically simple and so we can consider the associated Serre spectral sequence. Suppose $f_* : H_k(X) \to H_k(Y)$ is an isomorphism for $k < n$ and surjective for $k = n$. We work inductively to show $H_{k-1}(F) = 0$ for $2 \leq k \leq n$.

Suppose $H_0(F) = \mathbb{Z}$ and $H_1(F) = \cdots = H_{k-1}(F) = 0$. There are no non-trivial differentials on the spectral sequence until the page $k+1$:




This is the last time the illustrated squares participate in a non-trivial differential. We conclude $E^\infty_{k,0} = H_k(Y)$, $E^\infty_{k+1,0} = \ker \partial^{k+1}_{k+1,0}$ and $E^\infty_{0,k} = \operatorname{coker} \partial^{k+1}_{k+1,0}$.

We have that the $f_* : H_k(X) = E^\infty_{p+q=k} \to E^\infty_{k,0} = H_k(Y)$ is a surjection with kernel $E^\infty_{0,k}$. We also know that $E^\infty_{k+1,0} = \operatorname{im} f_* : H_{k+1}(X) \to H_{k+1}(Y)$ . In particular, we have an exact sequence,

$$H_{k+1}(X) \xrightarrow{f_*} H_{k+1}(Y) \xrightarrow{\partial^{k+1}_{k+1,0}} H_k(F) \xrightarrow{i_*} H_k(X) \xrightarrow{f_*} H_k(Y) \to 0.$$

It follows that if $H_0(F) = \mathbb{Z}$ and $H_1(F) = \cdots = H_{n-1}(F) = 0$, then $f_* : H_k(X) \to H_{k+1}(Y)$ is an isomorphism for $k < n$ and surjective for $k = n$. Conversely, if $f_* : H_k(X) \to H_{k+1}(Y)$ is an isomorphism for $k < n$ and surjective for $k = n$, we can inductively apply versions of the above exact sequence for increasing $k$ to show $H_k(F) = 0$ for $0 < k < n$. Hence the condition (ii) in the theorem is equivalent to $H_{k-1}(F) = 0$ for $1 < k \leq n$.

Thus we have reduced our theorem to showing $\pi_{k-1}(F) = 0$ for $k \leq n$ if and only if $H_{k-1}(F) = 0$ for $1 < k \leq n$. From the long exact sequence of a fibration,

$$\pi_2(Y) \to \pi_1(F) \to \cancel{\pi_1(X)} \to 0 \to \cancel{\pi_1(Y)}, \to 0$$

we conclude $\pi_1(F)$ is abelian and hence equal to $H_1(F)$. The rest of the equivalence is then just the Hurewicz theorem. □

Remark 4.37. A more standard proof goes the other way. We homotope $f : X \to Y$ to a cofibration and then prove a relative version of Hurewicz.

Also note that we did not quite use the full strength of the assumption that $X, Y$ are simply connected. If we only assume they are *homotopically simple* so that their fundamental groups are abelian and act trivially on the higher homotopy groups, then the result still holds. △

**4.5. Obstruction Theory.** One very beautiful synergy of homotopy and homology is to ask about whether a pair $(X, A)$ admits extensions of maps $A \to F$ to maps $X \to F$. Or more generally, we can ask about extending sections over $A$ of a bundle with fibre $F$ to sections over $X$.

We will deal with a fibration $F \hookrightarrow E \to B$ with $B$ a CW complex and try to extend sections over some subcomplex of $B$ to additional cells. We will need some assumptions on the fibration to make this work. We assume $F$ is *homotopically trivial* so that the action of $\pi_1(F)$ on its homotopy groups is trivial (in particular $\pi_1(F)$ should be abelian). We also assume the fibration is *homotopically simple* so that $E$ trivializes over any loop in $B$ (this holds for example if $\pi_1(B) = 0$ or if $E$ is a trivial fibration).

Suppose we have a section $s : B_n \to E$ defined over the $n$ skeleton of $B$ which we wish to extend to an $\widetilde{s}$ defined on an $n+1$ cell $D^{n+1}$. Since Feldbau's lemma

implies our bundle trivializes over a disk, we are dealing with the following diagram:

$$\begin{array}{ccc} F \times \partial D^{n+1} & \hookrightarrow & F \times D^{n+1} \\ \downarrow \quad \nwarrow s & & \downarrow \quad \nwarrow \widetilde{s}? \\ \partial D^{n+1} & \hookrightarrow & D^{n+1}. \end{array}$$

**Definition 4.38**

The map $s$ projected to $F$ defines an $n$-spheroid $c_s(D^{n+1}) \in \pi_n(F)$. Doing this for each $(n+1)$-cell of $B$ defines an element of the cochain complex $c_s \in \mathcal{C}^{n+1}(B; \pi_n(F))$ called the *obstruction cochain* of $s$.

Given two sections $s, s'$ on $B_n$ that agree on $B_{n-1}$, they together define for each $n$-cell $D^n$ an $n$-spheroid $d_{s,s'}(D^n) \in \pi_n(F)$. Doing this for each $n$-cell of $B$ defines an element $d_{s,s'} \in \mathcal{C}^n(B; \pi_n(F))$ called the *difference cochain* of $s, s'$.

Remark 4.39. Note that we used homtopic simplicity of $F$ to ignore base-points and we used homotopical triviality of the fibration to identitfy $\pi_n(F)$ over different fibres. △

Given a fibration $E \to B$ and a map $\varphi : \widetilde{B} \to B$, we can pullback the bundle to given $\varphi^* E \to \widetilde{B}$. A section $s$ of $E$ has a pullback section $\varphi^* s$ of $\varphi^* E$. The obstruction cochains we have defined are natural in the sense that,

$$c_{\varphi^* s} = \varphi^* c_s.$$

To see this, note that for any $n+1$ cell $\widetilde{D}$ of $\widetilde{B}$

$$\varphi^* c_s(\widetilde{D}) = c_s(\varphi_* \widetilde{D}) = c_s(\sum_\beta \deg_\beta D_\beta),$$

where $\deg_\beta$ is the degree of the map $\widetilde{B}/\widetilde{B} \setminus \widetilde{D} \xrightarrow{\varphi} B/B \setminus D_\beta$ for each $n$-cell $d_\beta$ of $B$. Approximating $\varphi$ smoothly and looking at the preimage of a regular value, $\varphi^* s$ on $\widetilde{D}$ is homotopic to a map on the preimages of the cells $D_\beta$ with appropriate coefficients given by the degrees, as above.

**Theorem 4.40:** [Adh, §16.3]

Suppose we have sections $s, s'$ defined on the $n$-skeleton of the base of a fibration $E \to B$. The following hold.

(i) The obstruction cochains are cocycles: $\delta c_s = 0$.

(ii) Obstruction cochains agreeing on $B_{n-1}$ are cohomologous via the difference cochain: $c_s - c_{s'} = \delta d_{s,s'}$.
(iii) The *obstruction class* $[c_s] \in H^{n+1}(B; \pi_n(F))$ is zero if and only if $s$ extends to $B^{n+1}$, possibly after modifying $s$ on $B^n \setminus B^{n-1}$.

PROOF. (i) Consider a $(n+2)$-cell $D^{n+2}$ of $B$. We wish to show $\delta c_s(D^{n+2}) = c_s(\partial D^{n+2}) = 0$. Denote by $\Gamma \subset B$ the subcomplex $\partial D^{n+2}$. Consider the following diagram, where $Z$ denotes the space of cellular cycle,

$$\begin{aligned}\mathcal{C}_{n+1}(\Gamma; \pi_n(F)) \xrightarrow{\partial} Z_n(\Gamma; \pi_n(F)) &= Z_n(\Gamma_n; \pi_n(F)) \\ &= H_n(\Gamma_n; \pi_n(F)) \xleftarrow{h} \pi_n(\Gamma_n) \xrightarrow{s_*} \pi_n(F).\end{aligned}$$

Here $h$ is the Hurewicz homomorphism and $s_*$ is induced by $s : D^{n+2} \to F \times D^{n+2} \to F$. For $n > 1$, $\Gamma_n$ is $n-1$ connected and $h$ is an isomorphism by Hurewicz, and so $\psi = s_* \circ h^{-1}$ is well defined. If $n = 1$, $s_*$ factors through the abelianization of $\pi_n(\Gamma_n)$, and so we can still make sense of $\psi : Z_n(\Gamma; \pi_n(F)) \to \pi_n(F)$.

Since $\mathcal{C}_{n-1}(\Gamma; \pi_n(F))$ is free abelian, we may express $Z_n(\Gamma; \pi_n(F))$ as a direct summand of $\mathcal{C}_n(\Gamma; \pi_n(Y))$. In particular $\psi$ can be extended to a map,

$$\widetilde{\psi} : \mathcal{C}_n(\Gamma; \pi_n(F)) \to \pi_n(F).$$

We have $c_s(D^{n+1}) = s_* \circ h^{-1}(\partial D^{n+1}) = \widetilde{\psi}(\partial D^{n+1})$. Hence,

$$c_s(\partial D^{n+2}) = \widetilde{\psi}(\partial \partial D^{n+2}) = 0,$$

as desired.

(ii) We want to show that for any given $(n+1)$-cell $D^{n+1} \subset B$,

$$c_s(D^{n+1}) - c_{s'}(D^{n+1}) = d_{s,s'}(\partial D^{n+1}).$$

But this is clear, since the difference of the homotopy classes of maps $s, s' : \partial D^{n+1} \to F$ is the homotopy class of the the glued map $s \cup -s' : \partial D^{n+1} \to F$.

(iii) We know that $c_s = 0$ if and only if $s$ extends to $B_{n+1}$ essentially by definition. Then changing $s$ to $s'$ by altering it on $B_n$ in any way gives, $c_{s'} = c_s + \delta d_{s',s} = \delta d_{s',s}$. Hence $[c_{s'}] = 0$ and so $[c_s] = [c_{s'}] = 0$.

Conversely, if $c_s = \delta f$ for $f \in H^n(B; \pi_n(F))$ so that $[c_s] = 0$, we can define $s'$ by altering $s$ on each $n$-cell $D^n$ to be a spheroid representing $s(D^n) + f(D^n)$. Then,

$$c_{s'} = c_s - \delta d_{s,s'} = c_s - \delta f = 0.$$

So $s$ can be altered on $B_n$ to $s'$ which extends to $B_{n+1}$. □

We will not elaborate on the details, but one can easily extend this to a relative version. Suppose $E \to B$ is a fibration and $(B, A)$ a CW pair. If we want to extend a section $s$ from $B_n \cup A$ to $B_{n+1} \cup A$, we can define a relative obstruction cochain $c_s \in \mathcal{C}^{n+1}(B, A; \pi_n(F))$ which turns out to be a cocycle and for two sections agreeing on $B_{n-1} \cup A$, a difference cochain $d_{s,s'} \in \mathcal{C}^n(B, A; \pi_n(F))$. Then $[c_s] = 0$ if and only if $s$ extends from $B_n \cup A$ to $B_{n+1} \cup A$ after altering it on $B_n \setminus B_{n-1} \cup A$.

REMARK 4.41. One can remove the assumption of homotopic triviality using local coefficients. Given a homotopically simple fibration $\pi : E \to B$, we define for each $k \geq 0$ the local coefficient system $\{\pi_k(E_x)\}$ on $B$. This local system is trivial for each $k$ and identified with $\pi_k(F)$ exactly when the fibration is homotopically trivial. Given a section $s$ of $E$ over the $k$-skeleton of $B$, we can define its obstruction class,

$$c_s \in H^{k+1}(B; \{\pi_k(E_x)\}).$$

All our results are exactly analogous to the homotopically simple case. △

EXAMPLE 4.42. The relative obstruction case has the following important example for constructing homotopies. Suppose we have sections $s, s'$ defined on a CW complex $B$ and a homotopy between them on $B_{n-1}$, i.e. a map $B_{n-1} \times I \to E$ agreeing with $s$ and $s'$ at the endpoints of $I$. We wish to extend this to a homotopy on $B_n$.

The obstruction class $c_{s \cup s'}$ lives in $\mathcal{C}^{n+1}(B \times I, B \times \{0\} \cup \{1\}; \pi_n(F))$ which is the same as $\mathcal{C}^n(B; \pi_n(F))$. One can check explicitly that $c_{s \cup s'} = d_{s,s'}$. We have,

$$\delta c_{s \cup s'} = \delta d_{s,s'} = c_s - c_{s'} = 0,$$

since $s$ and $s'$ extend to $B$. We have that $[c_{s \cup s'}] = 0$ in $H^n(B; \pi_n(F))$ if and only if this homotopy extends to $B_n \times I$ without alteration on $B_{n-2} \times I$. ▲

EXAMPLE 4.43. There is another generalization that we may be useful. Suppose $p : E \to B$ is a fibration and there is a map $f : X \to B$. We can ask about the existence of a lift $g : X \to E$. Suppose such a lift $g$ exists on the $n$-skeleton of $X$ and we wish to extend it to the $(n+1)$-skeleton. For a given $(n+1)$-cell $D^{n+1}$, its boundary defines an attaching map $S^n \xrightarrow{\alpha} X \xrightarrow{g} E$. Composing with $p$, we get a map $S^n \to B$ that we know extends to the disk $D^{n+1}$ via $f$. Hence we have a homotopy $F_t : S^n \times I \to B$ between $f \circ \alpha$ and a constant. Since we are dealing with a fibration, we may lift $F_t$ to a map $G_t : S^n \times I \to E$ so that $G_0 = g \circ \alpha$ and $p \circ G_1$ is constant. Hence we may homotope $g \circ \alpha : S^n \to E$ to a map $S^n \to F$ lying in a single fibre; to extend this to $D^{n+1} \to F$ requires the spheroid be contractible. Under our standard topological assumptions,

the obstruction to lifting $g$ to a map $X_{n+1} \to F$ thus defines an obstruction cochain $c_g \in H^{n+1}(X; \pi_n(F))$. Similarly we define difference cochains and our same results will hold. The case of sections is just this construction applied when the map $f$ we begin with on the base is $\mathrm{id}_B : B \to B$. ▲

*Another Proof of Homological Whitehead.* We can use obstruction theory to reprove the homological case of the Whitehead theorem without the need for spectral sequences.

**Theorem 4.44: Homological Whitehead Theorem**

Let $X$ and $Y$ be homotopically simple CW complexes and $f : X \to Y$ a continuous map inducing isomorphisms in homology. Then $f$ is a homotopy equivalence.

PROOF. We may homotope $f$ to be a cofibration $X \hookrightarrow Y$. From the homology long exact sequence of a pair, because $f_*$ is an isomorphism, $H_*(Y, X) = 0$. We conclude from the universal coefficient theorem that,

$$H^*(Y, X; \pi_n(X)) = 0 \quad \text{for all} \quad n.$$

We would like to extend the identity map $X \to X$ to a map $g : Y \to X$ completing the following diagram which commutes up to homotopy,

$$\begin{array}{ccc} X & \xhookrightarrow{f} & Y \\ {\scriptstyle \mathrm{id}}\downarrow & \swarrow_{g} & \\ X. & & \end{array}$$

Because $X$ is homotopically simple we can do obstruction theory, The relative obstruction to extending $g$ to the $i$ skeleton of $Y$ lives in $H^i(Y, X; \pi_{i-1}(X)) = 0$. Hence there is no obstruction to finding such a map $g : Y \to X$ so that $g \circ f \cong \mathrm{id}_X$.

Now note since id and $f$ both induce isomorphisms on homology groups, the above diagram implies that $g_* : H_*(Y) \to H_*(X)$ is also an isomorphism.

But now, since $Y$ is homotopically simple, we can apply the same obstruction theory argument to $g$ as we did to $f$ and find a map $h : X \to Y$ so that $h \circ g \cong \mathrm{id}_Y$.

But then,

$$f \circ g \cong h \circ g \circ f \circ g \cong h \circ g \cong \mathrm{id}_Y.$$

Hence $f$ is a homotopy equivalence between $X$ and $Y$. □

*First Obstructions.* If our fibre $F$ is highly connective, there may be no obstruction to extending a section to some higher skeleta of the base $B$. The first time we do run into an obstruction will carry some intrinsic data about the fibration independent of the choice of section. This concept of "first obstructions" will give important geometric realizations of characteristic classes which we study in depth in a later chapter.

Consider a homotopically trivial fibration $p : E \to B$ with CW base $B$ and fibre $F$. Suppose that,

$$\pi_0(F) = \pi_1(F) = \cdots = \pi_{n-1}(F) = 0,$$

so that $\pi_n(F)$ is the first non-trivial homotopy group of the fibre. From obstruction theory, we know that we can always obtain sections of $E$ over the $n$-skeleton $B_n$ of the base.

**Theorem 4.45**

Given any two sections $s, s' : B_n \to E$, we have,

$$[c_s] = [c_{s'}] \in H^{n+1}(B; \pi_n(F)).$$

Proof. Note that if $s|_{B_{n-1}}$ and $s''|_{B_{n-1}}$ agree then $c_s - c_{s''} = \delta d_{s,s''}$ and so $c_s$ and $c_{s''}$ are cohomologous. Thus it suffice to show $s'$ is homotopic to a section $s''$ agreeing with $s$ on $B_{n-1}$. Then we will have $[c_{s'}] = [c_{s''}] = [c_s]$.

We can prove this inductively. Say $s|_{B_{k-1}} = s'|_{B_{k-1}}$ for some $0 \leq k < n$. We would like to homotope $s'$ to agree with $s$ on $B_k$. The obstruction to such a homotopy is precisely,

$$d_{s|_{B_k}, s'|_{B_k}} \in \mathcal{C}^k(B; \pi_k(F)) = 0.$$

Hence they are homotopic on $B_k$. We can extend this homotopy to a map on all of $B \times I$ by a simple extension of Brosuk's theorem. Hence $s'$ is homotopic to a section agreeing with $s$ on $B_k$. Repeating this argument, we obtain our result. □

**Definition 4.46**

In this setting, we define the *first obstruction class* of the fibration $E$ to be the cohomology class,

$$C(E) := c_s \in H^{n+1}(B; \pi_n(F)),$$

for any section $s$ of $E$ defined on the $n$-skeleton of $B$ (this is independent of $s$ by the above theorem).

Suppose we have a cellular map $f : B' \to B$ and a fibre bundle $E$ on $B$. We showed above that obstruction classes are natural with respect to pullback. Hence,

$$C(f^*E) = f^*C(E). \tag{4.1}$$

By cellular approximation this extends to non-cellular maps. In particular, $C(E)$ does not depend on the CW structure we put on $B$. If $B$ is not CW, we may take a CW approximation $\varphi : B^{CW} \to B$, and pullback our bundle $E$ and then define $C(E)$ as $(\varphi^*)^{-1}C(\varphi^*E)$.

An association of a cohomology class in $H^*(B)$ to any bundle $E \to B$ that satisfies the naturality property 4.1 is called a *characteristic class*.

Thus, the first obstruction class defines a characteristic class of our fibre bundle sometimes called the *primary characteristic class*.

We will have much more to say about characteristic classes in the case of vector bundles later, but as a taste we will see one example now.

**Definition 4.47**

Let $\xi \to B$ be an oriented rank $n$ vector bundle and $S\xi$ be its associated sphere bundle, defined by picking some metric on $\xi$ and restricting to vectors of norm one in each fibre. The fibres $F = S^{n-1}$ of $S\xi$ are $(n-2)$-connected and $\pi_{n-1}(F) = \mathbb{Z}$. The fact the bundle is oriented means $S\xi$ is homotopically trivial. Thus we obtain a first obstruction,

$$e(\xi) := C(S\xi) \in H^n(B),$$

called the *Euler class* of $\xi$.

Here is some motivation for the name of the Euler class.

Proposition 4.48. *Let $M$ be a closed oriented $n$-manifold. Then,*

$$\langle e(TM), [M] \rangle = \chi(M).$$

*Equivalently, $e(TM)$ is Poincaré dual to $\chi(M) \in H_0(M)$.*

Proof. Pick a cell structure on $M$ by use of Morse theory. Pick a generic section $V$ of $TM$, i.e. a vector field, so that the zeroes of $V$ are discrete and non-degenerate. We can assume the zeroes all lie in the interior of $n$-cells of $M$. Hence $V$ defines a section of the sphere bundle $STM$ over the $n-1$ skeleton of $M$. Our proof of the fundamental class theorem shows $[M]$ can be written as the sum of all $n$ cells of $M$, oriented compatibly with the orientation of $M$. Given a single $n$-cell $D^n$, the value of $\langle e(TM), [D^n] \rangle$ is the value of the

obstruction class $c_V(D^n)$. The spheroid $V : \partial D^n \to STM|_{D^n} \cong S^{n-1} \times D^n$ extends over $D^n$ minus some small balls around the zeroes of $V$ and so the homotopy class of this spheroid (i.e. the degree of this map $V : \partial D^n \to S^{n-1} \times D^n$) is the sum of the homotopy classes around each zero. Near each of these zeroes, $V$ determines a map $S^n \to S^n$ whose degree is by definition the index of $V$ at the zero. Hence, $\langle e(TM), [D^n] \rangle$ is the sum of the indices of all interior zeroes. Summing over all cells,

$$\langle e(TM), [M] \rangle = \sum_{V(x)=0} \mathrm{ind}_x(V) = \chi(M),$$

where the last equality is just Poincaré–Hopf. □

We can give a more general geometric description of the Euler class.

PROPOSITION 4.49. *Let $\xi$ be a rank $r$ oriented vector bundle on $B$ and let $t_\xi \in H^r(\xi, \xi \setminus B) \cong H^r(D_\xi, \partial D_\xi)$ be its Thom class. The pullback of $t_\xi$ under the inclusion of the zero section $i : B \hookrightarrow \xi$ is $i^*(t_\xi) = e(\xi) \in H^r(B)$.*

*Moreover, suppose $B$ is an oriented $n$-manifold and $s$ is a section of $\xi$ that transversely intersects the zero section. Let $Z$ be the zero locus of $s$, which is a submanifold of $B$ of dimension $n - r$. Then $e(\xi)$ is Poincaré dual to $[Z] \in H_{n-r}(B)$.*

PROOF. Pick some non-vanishing section $s$ of $\xi$ over the $(k-1)$-skeleton of $B$. The Euler class measure the first obstruction to extending the section to the $k$-skeleton of $B$. Extend $s$ to a section on $B_k$ transversely intersecting the zero section on each cell. The Euler class evaluated on each $k$-cell $D^k$ is an oriented count of the zeroes of $s$ on $D^k$ (see the proof of the previous proposition).

Because $s$ is homotopic to the zero section on $D^k$, we can compute our restriction of $t_\xi$ by pulling back to $s(D^k)$. Since $s$ is non-zero on the $(k-1)$-skeleton, we can pullback $t_\xi$ to define an element of $H^k(B_k, B_{k-1})$. The section $s$ considered in $\xi$ modulo $\xi \setminus B$ is homotopic to an oriented sum of the fibres of $\xi$ over which $s$ vanishes. The Thom class on the other hand represents the fundamental class of each fibre. So restricting the Thom class to the section and then considering this as an element in the homology of the base gives the cohomology class counting the oriented zeroes of $s$ on each $k$-cell $D^k$. But this is exactly what the Euler class does. Hence $t_\xi|_B = e(\xi)$.

For the second statement, recall from intersection theory that $t_\xi$ is Poincaré dual to $[B]$, the fundamental class of the base in the total space of $\xi$. Pulling back $t_\xi$ to the base gives on the one hand $e(\xi)$ by above, while on the other

hand it is Poincaré dual to the class of the self intersection of $[B]$ in $\xi$. The self intersection of $[B]$ can be found by perturbing $B$ to a section of $\xi$ that has transverse intersection with the zero section and looking at the class of this intersection. This is exactly what we claimed. □

EXAMPLE 4.50. The above proposition for a tangent bundle is a special case of this result. There is another important case of the normal bundle $\nu_N$ of an embedding $N \hookrightarrow M$ of smooth oriented manifolds. The Euler class $e(\nu_N)$ of this embedding will be Poincaré dual to the self-intersection of $N$, $[N] \cdot [N]$ inside $N$. In the case where $\dim M = 2 \dim N$, the Euler class is identified with an integer which is the *self-intersection number* of $N \subset M$. As the proposition shows, this is computed by an oriented count of intersections of $N$ with some generic isotoped copy of itself. The submanifold $N \subset M$ can be displaced from itself only if its self-intersection number is zero. ▲

There is one final easy but important corollary of what we have said.

COROLLARY 4.51. *The Euler class is* unstable. *If $\xi$ is an oriented vector bundle on $B$ and $\underline{\mathbb{R}}$ is the trivial line bundle on $B$, then $e(\xi \oplus \underline{\mathbb{R}}) = 0$.*

PROOF. The section $s(x) = (0, 1)$ of $\xi \oplus \mathbb{R}$ is non-vanishing on $B$ and so there is no obstruction to extending a non-vanishing section of the bundle to all the skeleta of $B$. □

REMARK 4.52. Even if $\xi \to B$ is not oriented, we can use the construction of obstruction classes with local coefficients to define an Euler class,

$$e(\xi) \in H^{n+1}(B; \{H_n(E_x)\}).$$

Here we use that $H_n(E_x) \cong \pi_n(E_x) \cong \mathbb{Z}$ for each $x \in B$.

### 4.6. Eilenberg–MacLane Spaces, Messing Around with Homotopy Groups, and Applications to Cohomology.

*Eilenberg–MacLane Spaces.* Another very reasonable question to ask is whether there exist spaces realizing any sets of homotopy groups. This can easily be seen to be true.

EXERCISE 4.53. Let $G_1, G_2, \ldots$ be a collection of groups, all abelian except possibly $G_1$. Let $\rho_i$ for $i = 2, 3, \ldots$ be a collection of group actions of $G_1$ on $G_i$. Construct a connected CW complex whose homotopy groups are $\pi_i(X) = G_i$ and whose action of $\pi_1(X)$ on $\pi_i(X)$ is $\rho_i$. ▲

The general spaces constructed in the above exercise are usually not so interesting. But, they turn out to be very important in the case when only one of the homotopy groups is non-zero.

**Definition 4.54**

Let $n$ be a positive integer and $G$ a group, abelian if $n > 1$. A space is called an *Eilenberg–MacLane space* and denoted $K(G, n)$ if it satisfies,

$$\pi_i(K(G,n)) = \begin{cases} G & i = n \\ 0 & \text{otherwise} \end{cases}.$$

While in general the homotopy groups are not enough to determine the weak homotopy type of a space, they are sufficient in the case of Eilenberg–MacLane spaces.

**Theorem 4.55: Eilenberg–MacLane Spaces**

There exists a unique Eilenberg–MacLane space $K(G, n)$ up to weak homotopy equivalence for any $G$ and $n$. In particular, all CW Eilenberg–MacLane spaces $K(G, n)$ are homotopy equivalent.

Proof. Let's construct a CW complex which is an Eilenberg–MacLane space $K(G, n)$. Let $\{g_\alpha\}$ be a generating set for $G$. Consider $X_n$ a bouquet of $n$-spheres indexed by the generators $g_\alpha$. We have that $\pi_i(X_n) = 0$ for $i < n$ and $\pi_n(X_n)$ is the free abelian group (or just free group if $n = 1$) generated by the $g_\alpha$. We can thus find a subgroup $H$ of $\pi_n(X_n)$ so that $\pi_n(X_n)/H \cong G$. Let $\{h_\beta\}$ be a generating set for $H$. We may attach $n + 1$ cells $D_\beta$ to $X_n$ using $h_\beta : S^n \to X_n$ as the attaching maps, to obtain a $n+1$ dimensional CW complex $X_{n+1}$. Note by the cell attaching lemma that $\pi_i(X_{n+1}) = \pi_i(X_n) = 0$ for $i < n$. Moreover, $\pi_n(X_{n+1}) = \pi_n(X_n)/H = G$.

Now, we finish inductively. Suppose we have an $m+1$ dimensional CW complex $X_{m+1}$ for $m \geq n$ so that,

$$\pi_i(X_{m+1}) = \begin{cases} G & i = n \\ 0 & i \leq m, i \neq n. \end{cases}$$

Set $\pi_{m+1}(X_{m+1}) = G_{m+1}$. Suppose this group is generated by spheroids $f_\gamma : S^{m+1} \to X_{m+1}$. We may attach a collection of $m + 2$ cells $D_\gamma$ to $X_{m+1}$ using $f_\gamma$ as the attaching maps, to obtain a $m + 2$ dimensional CW complex $X_{m+2}$. By the cell attaching lemma, $\pi_i(X_{m+2}) = \pi_i(X_{m+1})$ for $i \leq m$. Moreover, by the cell attaching lemma, $\pi_{m+1}(X_{m+1}) = 0$.

Hence taking the direct limit of these CW complexes with the weak topology gives a space $X = \lim_{n\to\infty} X_n$ whose homotopy groups all vanish except for $\pi_n(X) = G$. That is $X$ is an Eilenberg–MacLane space.

Now suppose $Y$ is another CW version of $K(G, n)$. From our construction of $X$, we know $\pi_n(X) \cong G$ is generated by the inclusions of spheres $g_\alpha : S^n_\alpha \hookrightarrow \bigvee_\beta S^n_\beta = X_n \subset X$. Let $s_\alpha : S^n \to Y$ a collection of spheroids representing the same generators under an isomorphism $\pi_n(Y) \cong G$. Now define a map on the $n$-skeleta $F : X_n \to Y$ which on each sphere in the bouquet $S^n_\alpha \subset X_n$ restricts to the map $s_\alpha$.

Note from obstruction theory, the series of obstructions to extending $F$ to a map $\widetilde{F} : X \to Y$ live in $H^{i+1}(X, \pi_i(X))$ for $i \geq n$. Note that every $n+1$ cell we attach to $X_n$ bounds a spheroid $h_\alpha$ which is null homotopic after composition with $F$ (since $F_*$ sends generators of $\pi_n(X_n)$ to the corresponding generators of $G$, and $h_\alpha$ was chosen as a relation on the generators of $G$). Hence the first obstruction $c_F \in H^{n+1}(X, \pi_n(X))$ vanishes and we can extend $F$ to the $n + 1$ skeleton. For any $i > n$, the obstruction class must vanish since $X$ is a $K(G, n)$ and so $F$ extends to a map $\widetilde{F} : X \to Y$ with $\widetilde{F}|_{X_n} = F$.

It is clear that under isomorphism with $G$, $\widetilde{F}_* : \pi_n(X) \to \pi_n(Y)$ acts as the identity on generators and hence is an isomorphism. For all other $i$, $\widetilde{F} : \pi_i(X) \to \pi_i(Y)$ is necessarily the trivial isomorphism.

Hence $\widetilde{F}$ is a weak homotopy equivalence, and by the Whitehead theorem is a homotopy equivalence. Given any space $Z$ which is a $K(G, n)$, it has a CW approximation $Z^{CW} \to Z$. Note $Z^{CW}$ is a CW $K(G, n)$ which must be homotopy equivalent to $X$ by above. Hence every Eilenberg–MacLane space is weak homotopy equivalent to $X$. □

Example 4.56. While our proof of existence was constructive, it was far from explicit, as in general we do not know what the $(m + 1)$st homotopy group of our $(m + 1)$-skeleton that we need to kill will be.

For example, suppose we want to build a CW model for $K(\mathbb{Z}, n)$. We know we may take its $n$-skeleton to be $S^n$. But already, we do not in general know what $\pi_{n+1}(S^n)$ is so we do not know what $(n + 2)$-cells to attach to ensure the $(n + 1)$st homotopy group vanishes.

We can make a few Eilenberg–MacLane space constructions explicit, as we discuss now.

(i) If we wish to find an Eilenberg–MacLane space $K(G, 1)$, one method is to find a space with fundamental group $G$ and a contractible universal

cover. Conversely, if $G$ is discrete, we can find a contractible space with a free $G$-action and take the quotient.

For example, quotienting $\mathbb{R}^n$ by a lattice gives, $K(\mathbb{Z}^n, 1) = T^n$, the $n$-torus. In particular, $K(\mathbb{Z}, 1) = S^1$.

(ii) Recall the sphere $S^\infty$ is contractible. The quotient by antipodal identification gives $K(\mathbb{Z}_2, 1) = \mathbb{R}P^\infty$.

(iii) More generally, $S^\infty \subset \mathbb{C}^\infty$ has a $\mathbb{Z}_m$ action by,

$$q \cdot (z_1, z_2, \ldots) = (z_1 e^{2\pi i q/m}, z_2 e^{2\pi i q/m}, \ldots).$$

The quotient $L^\infty_m = S^\infty/\mathbb{Z}_m$ is called an *infinite lens space* and is a model for $K(\mathbb{Z}_m, 1)$.

(iv) Any surface $\Sigma$ which is not $S^2$ or $\mathbb{R}P^2$ has $\mathbb{R}^2$ as its universal cover and hence is a model for $K(\pi_1(\Sigma), 1)$.

(v) Recall from the infinite Hopf fibration $S^1 \to S^\infty \to \mathbb{C}P^\infty$ that $\pi_i(\mathbb{C}P^\infty) = \pi_{i-1}(S^1)$. We conclude that $\mathbb{C}P^\infty$ is a model for $K(\mathbb{Z}, 2)$. This is essentially the only space $K(G, n)$ for $n > 1$ that we can make totally explicit.

(vi) Let $\mathrm{Sym}^n(X)$ denote the $n$th symmetric product of $X$, i.e. the quotient $X^n/S_n$ where $S_n$ is the symmetric group acting by permutations of factors. Note if $X$ has a basepoint $x_0$, there is an inclusion map $\mathrm{Sym}^n(X) \hookrightarrow \mathrm{Sym}^{n+1}(X)$ sending $(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_n, x_0)$. Let $\mathrm{Sym}^\infty(X)$ be the direct limit as $n \to \infty$ of this construction. One can construct a fibration to show for any path-connected Hausdorff space $X$ that $\pi_n(\mathrm{Sym}^\infty X) = \pi_{n+1}(\mathrm{Sym}^\infty \Sigma X)$. In particular we deduce, $K(\mathbb{Z}, n) = \mathrm{Sym}^\infty(S^n)$.

(vii) More generally, one has the Dold–Thom theorem, which says that for a connected CW complex $X$,

$$\pi_n(\mathrm{Sym}^\infty(X)) \cong \widetilde{H}_n(X).$$

One concludes that we can obtain any Eilenberg–McLane space $K(G, n)$ by finding a space $X$ whose reduced homology is concentrated in degree $n$ and equals $G$ there. Such a space is called a *Moore space* $M(G, n)$. These are not unique like Eilenberg-MacLane spaces, for example there are homology spheres which have the same homology as a sphere but a non-trival perfect fundamental group. A CW Moore space $M(\mathbb{Z}_k, n)$ can be constructed easily. Simply take an $n$-sphere and glue on an $n + 1$ dimensional disk by a degree $k$ map.

(viii) Let $\mathbf{Top}(n)$ and $\mathbf{PL}(n)$ be respectively the topological groups of homeomorphisms and piecewise linear homeomorphisms of $\mathbb{R}^n$ fixing the origin. Let $\mathbf{Top}$ and $\mathbf{PL}$ be the direct limit of these groups under inclusions into one dimension higher by the identity on the last factor. It is a result of Kirby and Siebenmann that $\mathbf{Top}/\mathbf{PL}$ is a model for $K(\mathbb{Z}_2, 3)$. This has the following consequence in differential topology.

Associated to a manifold $M$ with an $\mathcal{M}$-structure, there is a classifying map for its tangent bundle $M \to B\mathcal{M}$, where $B\mathcal{M}$ is the classifying space of the group $\mathcal{M}$ (we will talk about these later). So, to make a topological manifold $M$ piecewise linear (and hence smooth, since those categories are equivalent) is to ask for a lift of the classifying map $M \to B\mathbf{Top}(n)$ to $M \to B\mathbf{PL}(n)$.

As for any classifying spaces, there is a fibration,

$$\mathbf{Top}(n)/\mathbf{PL}(n) \hookrightarrow B\mathbf{PL}(n) \to B\mathbf{Top}(n).$$

By obstruction theory (our example 4.43), the obstructions to lifting live in the groups $H^{i+1}(M; \pi_i \mathbf{Top}(n)/\mathbf{PL}(n))$. It is known for $n \geq 5$ that the map $\mathbf{Top}(n)/\mathbf{PL}(n) \to \mathbf{Top}/\mathbf{PL}$ is $(n+1)$-connected and hence the only obstruction to lifting is a characteristic class $ks(M) \in H^4(M; \mathbb{Z}_2)$. This obstruction is called the *Kierby–Siebenmann class* and by what we've claimed topological manifolds of dimension at least 5 are smoothable precisely when it vanishes. ▲

Exercise 4.57. Here are some properties of Eilenberg–MacLane spaces.

(1) Prove that $K(\pi_1, n) \times K(\pi_2, n) \cong K(\pi_1 \oplus \pi_2, n)$. Construct as explicitly as possible $K(F, n)$ for any finitely generated abelian group.

(2) Prove that $K(\pi, n-1) \cong \Omega K(\pi, n)$ and conclude $\Omega \mathbb{C}P^\infty$ is weak homotopy equivalent to $S^1$.

(3) Prove that a group homomorphism $G \to H$ induces a map of Eilenberg–MacLane spaces $K(G, n) \to K(H, n)$. Prove a short exact sequence of abelian groups $0 \to A_1 \to A_2 \to A_3 \to 0$ induces a long exact sequence of Eilenberg–MacLane spaces,

$$\cdots \to K(A_1, n) \to K(A_2, n) \to K(A_3, n) \to K(A_1, n+1) \to \cdots .$$

What does the beginning of this sequence look like for the exact sequence $0 \to \mathbb{Z} \xrightarrow{\times 2} \mathbb{Z} \to \mathbb{Z}_2 \to 0$? ▲

*An Application to Homology.* One very important application of Eilenberg–MacLane spaces is to give a homotopical formulation of homology groups. This can be useful in proving some properties of homology. But its main benefit is reducing some homotopical questions, for example the classification of principal and vector bundles which we will intensively study later, to homological considerations.

For some space $X$ and some Eilenberg–MacLane space $K(G,n)$, consider the homotopy classes of pointed maps $[X, K(G,n)]$. Note this has a group structure described as follows. By the exercise above, $K(G,n) \cong \Omega K(G,n+1)$. There is a map,

$$\Omega K(G,n+1) \times \Omega K(G,n+1) \to \Omega K(G,n+1),$$

given by composition of loops. Applying the $[X,-]$ functor and the above isomorphism, this induces a map,

$$[X, K(G,n)] \times [X, K(G,n)] \to [X, K(G,n)].$$

This gives a group operation. The identity is the constant map and the inverse is given by taking a map in $[X, K(G,n)] \cong [X, \Omega K(G,n+1)]$ and composing with the map reversing directions of loops.

There is also a graded ring structure on $\bigoplus_{n\in\mathbb{N}}[X, K(\mathbb{Z},n)]$ described as follows. Consider the smash product $K(\mathbb{Z},n) \wedge K(\mathbb{Z},m)$. Using our CW model of Eilenber–MacLane spaces, we deduce this space is $(n+m-1)$ connected and its $\pi_{n+m}$ group is $\mathbb{Z}$. Hence we may glue on cells of dimension $n+m+2$ and higher to obtain a cofibration $K(\mathbb{Z},n) \wedge K(\mathbb{Z},m) \hookrightarrow K(\mathbb{Z},n+m)$. Given maps $f : X \to K(\mathbb{Z},n)$ and $g : X \to K(\mathbb{Z},m)$, we obtain a map $f \wedge g$ from $X \times X$ into the smash product which we may precompose with diagonal $\Delta : X \to X \times X$ and post-compose with inclusion to get a map $X \to K(\mathbb{Z}, n+m)$. Thus we obtain a map,

$$[X, K(\mathbb{Z},n)] \times [X, K(\mathbb{Z},m)] \to [X, K(\mathbb{Z}, m+n)].$$

Exercise 4.58. Show this map distributes over our addition map. Show this defines a ring multiplication on $\bigoplus_{n\in\mathbb{N}}[X, K(\mathbb{Z},n)]$. Extend this to a map,

$$[X, K(G,n)] \times [X, K(H,m)] \to [X, K(G \otimes_{\mathbb{Z}} H, n+m)],$$

for $\mathbb{Z}$-modules $G, H$. In particular for $G = H = G \otimes H$ a field. ▲

As the final ingredient we need for the upcoming theorem, we note there is a canonical element of $H^n(K(G,n); G)$. Namely, note by the Hurewicz theorem and the universal coefficient theorem that $H^n(K(G,n); G) \cong \mathrm{Hom}(G,G)$. Consider the element $F_G \in H^n(K(G,n); G)$ which under this isomorphism represents the identity map in $\mathrm{Hom}(G,G)$.

This element has a couple equivalent definitions. Recall that we constructed $K(G,n)$ as a CW complex whose $n$-skeleton was a bouquet of $n$-spheres $S^n_\alpha$ indexed by generators $g_\alpha$ of $G$. We can define a cellular cochain $F_G \in \mathrm{Hom}(\mathcal{C}_n(K(G,n)), G)$ so that $F_G(S^n_\alpha) = g_\alpha$ and then extend linearly. For any $n+1$ cell $D^{n+1}$,

$$\delta F_G(D^{n+1}) = F_G(\partial D^{n+1}).$$

But the boundary of $(n+1)$-cells we attach correspond to relations in $G$, and so $F_G$ must evaluate to zero on them. Hence $F_G$ defines a cohomology class and will agree with $F_G$ above.

For an obstruction theory perspective, we can view $F_G$ as the difference cochain $d_{\mathrm{id},*} \in \mathcal{C}^n(K(G,n); G)$ between the identity and the constant maps $K(G,n) \to K(G,n)$ on the $n$-skeleton. We know that both id and $*$ extend to all of $K(G,n)$ and so $\delta d_{\mathrm{id},*} = c_{\mathrm{id}} - c_* = 0 - 0$. Hence $d_{\mathrm{id},*} \in H^n(K(G,n); G)$ is a cohomology class. Studying this explicitly on the $n$-skeleton, we see it again coincides with $F_G$.

**Theorem 4.59: Homotopical Formulation of Cohomology**

For any CW complex $X$ and any abelian group $G$ there is a bijection,

$$[X, K(G,n)] \to H^n(X; G) \quad \text{given by} \quad f \mapsto f^*(F_G),$$

which is natural with respect to $X$. Moreover, this is a group isomorphism using the loop composition map on the left and addition on the right. Taking the direct sum over all $n$ and $G = \mathbb{Z}$ or a field $F$, this is a graded ring isomorphism using the smash product on the left and the cup product on the right.

Proof. Consider a cellular cocycle representative of a cochain $c \in H^n(X; G)$. Consider the constant map $f_0 : X_{n-1} \to * \subset K(G,n)$. We can extend this to $f_n : X_n \to K(G,n)$ as follows. Because $f_0$ is constant, we may factor $f_n$ as $X_n \to X_n/X_{n-1} \to K(G,n)$. $X_n/X_{n-1}$ is a bouquet of spheres $S^n_\alpha$ corresponding to the $n$-cells $D^n_\alpha$ of $X$. We define $f_n$ on $S^n_\alpha$ to represent the spheroid $c(D^n_\alpha) \in G \cong \pi_n(K(G,n))$. We have by construction and naturality,

$$c = d_{f_n,*} = f_n^*(d_{\mathrm{id},*}) = f_n^* F_G.$$

So $c$ is in the image of our map at least on $X_n$. Also, since $c$ is cocycle,

$$0 = \delta c = \delta d_{f_n,*} = c_{f_n} - c_* = c_{f_n}.$$

This implies that $f_n$ extends to a map $f_{n+1} : X_{n+1} \to K(G,n)$. But then $\pi_k(K(G,n)) = 0$ for $k > n$ and so the map extends to all higher skeleta and

we obtain $f : X \to K(G,n)$ with $f|_{X_n} = f_n$. Hence $f^*F_G = c$. Thus, our map is surjective.

Now suppose $f, g : X \to K(G,n)$ satisfy $[f^*F_G] = [g^*F_G]$. By cellular approximation we may assume $f, g$ are constant on the $(n-1)$-skeleton. On $X_n$,

$$d_{f,g} = d_{f,*} - d_{g,*} = f^*F_G - g^*F_G \in \mathrm{im}(\delta).$$

Hence $\delta d_{f,g} = 0$. By example 4.42, we can find a homotopy between $f$ and $g$ on $X_n$. Since the higher homotopy groups of $K(G,n)$ vanish, obstruction theory implies this homotopy extends to $X$. Hence $f, g$ are in the same homotopy class and our map is injective.

The naturality of the map with respect to $X$ is clear, since if we have a map $g : Y \to X$ and $f \in [X, K(G,n)]$ then $(f \circ g)^*F_G = g^*(f^*F_G)$.

To show the map is a group homomorphism, note that the inclusion $\Omega K \xhookrightarrow{\mathrm{id}\times *} \Omega K \times \Omega K \xrightarrow{\circ} \Omega K$ is homotopic to the identity, as is including by $* \times \mathrm{id}$ instead. Applying this to $K(G,n) = \Omega K(G, n+1)$, our group operation precomposed with inclusion from either factor is homotopic to the identity.

Let $i_1, i_2$ denote the two factor inclusions and $\mu$ our multiplication. We thus have that $i_1^*\mu^*F_G = i_2^*\mu^*F_G = F_G$. By the universal coefficient theorem and the Künneth formula,

$$H^n(K(G,n) \times K(G,n); G) = \mathrm{Hom}(G \otimes \mathbb{Z} \oplus \mathbb{Z} \otimes G, G).$$

We thus have, $\mu^*F_G = F_G \oplus F_G$. But then for any $f_1, f_2 : X \to K$,

$$(f_1 \times f_2)^*\mu^*F_G = (f_1 \times f_2)^*(F_G \oplus F_G) = f_1^*(F_G) + f_2^*(F_G).$$

Now finally we show it is a ring homomorphism, beginning with some notation. Let $p_1, p_2$ denote the projections from $K(G,n) \times K(G,m)$ to the two factors. Let $q$ be the composition of quotient and inclusion $K(G,n) \times K(G,m) \to K(G,n) \wedge K(G,m) \to K(G, n+m)$. Let $F_G^i$ denote our canonical element in $H^i(K(G,i); G)$. It is simple enough to check explicitly in terms of cellular descriptions using Künneth and otherwise that,

$$p_1^*F_G^n \smile p_2^*F_G^m = q^*F_G^{n+m}.$$

We can rephrase this as follows. Applying the bijection we have just proved to all the terms in the cup product gives a map,

$$[X, K(G,n)] \times [X, K(G,n)] \to [X, K(G, n+m)].$$

If we plug in $X = K(G,n) \times K(G,m)$, then we just concluded that the image of $(p_1, p_2)$ under this map is the map $q$. Now to verify that this cup product

map always agrees with the one coming from smash products is essentially just a matter of naturality.

Given $f \in [X, K(G,n)]$ and $g \in [X, K(G,m)]$, the smash product we have described maps this pair to $q \circ (f \times g) \circ \Delta \in [X, K(G, n+m)]$. Hence under our isomorphism, this product is sent to the cohomology class,

$$\Delta^*(f \times g)^* q^* F_G^{n+m} = \Delta^*(f \times g)^*(p_1^* F_G^n \smile p_2^* F_G^m).$$

Note that $p_i \circ f \times g = f \circ \pi_i$ where $\pi_i$ are the projections $X \times X \to X$,

$$= \Delta^*(\pi_1^* f^* F_G^n \smile \pi_2^* g^* F_G^m).$$

Recalling the definition of the cup product in terms of the diagonal embedding into the product,

$$= f^* F_G^n \smile g^* F_G^m$$

Thus the cup product applied after the isomorphism is the same as the smash product applied before the isomorphism. Hence, we have a ring map. □

We have the following corollary.

COROLLARY 4.60. *Reduced homology of spaces is stable under suspension:*

$$\widetilde{H}_i(X) \cong \widetilde{H}_{i+1}(\Sigma X).$$

PROOF. By CW approximating, we may take $X$ to be a CW complex so our theorem applies. One has group isomorphisms for $i \geq 1$,

$$H^i(X) \cong [X, K(\mathbb{Z}, i)] \cong [X, \Omega K(\mathbb{Z}, i+1)] \cong [\Sigma X, K(\mathbb{Z}, i)] \cong H^{i+1}(\Sigma X),$$

where we used loop-suspension adjunction in the middle. By the universal coefficient theorem, we get an isomorphism of homology as well. For $i = 0, -1$, this is easy to prove directly.

The isomorphism can be more explicitly realized by taking a singular simplex $\sigma : \Delta^n \to X$ and suspending to $\Sigma\sigma : \Sigma\Delta^n \to \Sigma X$. We may identify $\Sigma\Delta^n$ with two copies of $\Delta^{n+1}$ glued along a pair of faces so that $\Sigma\sigma$ defines a singular $n+1$ chain.

One could also prove the isomorphism with Mayer–Vietoris. □

The fact we mentioned much earlier that the Freudenthal suspension theorem implies stability of the homotopy groups of repeated suspensions can now be proved.

COROLLARY 4.61. *The $(n+1)$st suspension of a space $X$ is $n$-connected. Hence, $\pi_k^{\mathrm{st}}(X) = \pi_{2k+2}(\Sigma^{k+2} X)$.*

PROOF. We have $\widetilde{H}_i(\Sigma^{n+1}X) = \widetilde{H}_{i-n-1}(X)$ is zero as long as $i < n+1$. By Hurewicz, we conclude $\pi_i(\Sigma^{n+1}X) = 0$ for $i \leq n$. The second statement follows directly from Freudenthal suspension. □

REMARK 4.62. We should have probably discussed this earlier, but better late than never. We define a homotopical generalization of topological groups (and a dual object which doesn't exist in **Top**!).

**Definition 4.63**

An *H-space* (H for Hopf) is a space $X$ equipped with a multiplication $\mu : X \times X \to X$ so that the following diagram commutes up to homotopy,

$$\begin{array}{ccc} X \vee X & & \\ \downarrow & \searrow^{\mathrm{id}\vee\mathrm{id}} & \\ X \times X & \xrightarrow{\mu} & X. \end{array}$$

A *co-H-space* is a space $Y$ equipped with a comultiplication $c : Y \to Y \vee Y$ so that the following diagram commutes up to homotopy,

$$\begin{array}{ccc} & & Y \vee Y \\ & \overset{c}{\nearrow} & \downarrow \\ Y & \xrightarrow[\Delta]{} & Y \times Y. \end{array}$$

Every loop space $\Omega X$ is an H-space via loop composition. Every topological group (or even monoid) is by group multiplication (this is actually a special case of loop spaces since $G \cong \Omega BG$ as we will prove later).

Every suspension $\Sigma Y$ is a co-H-space by collapsing the equator. More generally $Y \wedge A$ is a co-H-space for any co-H-space $A$.

Note that if either $X$ is a co-H-space or $Y$ is an $H$-space, then the set $[X, Y]$ obtains a multiplication (and if we're lucky a group structure). In the case $X = S^n = \Sigma S^{n-1}$ or $Y = K(G, n) = \Omega K(G, n+1)$, these are just the familiar group operations on the homotopy and homology groups. If both $X$ is a co-H-space and $Y$ is an $H$-space, we will obtain two multiplications, but it turns out these must coincide and moreover are commutative and associative by the Eckmann–Hilton argument. So for example the isomorphisms,

$$\pi_n(K(G, n)) \cong [S^n, K(G, n)] \cong H^n(S^n; G)$$

are all compatible with group structures.

Similarly if $X$ is an H-space or $Y$ is a co-H-space, then $[X, Y]$ obtains a co-multiplication structure. For example, the homotopy groups $H^*(G; K)$ of a topological group $G$ have a co-multiplication. It turns out this is compatible with the multiplication structure of the cup product, so that $H^*(G)$ is a *Hopf algebra*. Dualizing, the homology $H_*(X)$ of an H-space $X$ has a multiplication; this multiplicative structure is called the *Pontryagin ring* of $X$.

As one last note motivating the definition of $H$-spaces, the smash product can be seen as a topological version of the tensor product. And the definition of an $H$-space really says we have something like an operation on the smash product $X \wedge X$ up to homotopy. △

Combining this description of homology with the cellular approximation theorem and our knowledge about attaching cells may allow explicit descriptions of homotopy classes $[X, Y]$ for certain pairs $X, Y$. This will be fruitful when we wish to study principal $G$-bundles, which are known to be classified by $[X, BG]$. Here is a preliminary example of this.

COROLLARY 4.64 (Hopf Theorem). *Let $X$ be an $n$ dimensional connected CW complex. Then,*

$$[X, S^n] \cong H^n(X; \mathbb{Z}).$$

PROOF. By cellular approximation, any map $X \to K(\mathbb{Z}, n)$ can be assumed to lie in the $n$-skeleton of $K(\mathbb{Z}, n)$ and any homotopy $X \times I \to K(\mathbb{Z}, n+1)$ can be assumed to lie in the $(n+1)$-skeleton. So restricting the codomain to the $(n+1)$-skeleton gives the same set of maps. From our Eilenberg–MacLane space construction we see that a CW $K(\mathbb{Z}, n)$ can be constructed by starting from $S^n$ and gluing cells of dimension at least $n+2$, so that $K(\mathbb{Z}, n)$ has a cell structure with $(n+1)$-skeleton $S^n$. Hence,

$$H^n(X; \mathbb{Z}) \cong [X, K(\mathbb{Z}, n)] \cong [X, \mathrm{sk}_{n+2} K(\mathbb{Z}, n)] \cong [X, S^n].$$

□

*Cohomology Operations.* The results we just proved lend themselves to a compact description of the functorial operations one can perform on cohomology.

**Definition 4.65: Cohomology Operation**

A *cohomology operation* of type $(q, \pi, G)$ and degree $n$, for integers $n, q$ and abelian groups $\pi, G$ is a natural transformation of the functors,

$$\Phi : H^q(-; \pi) \to H^{q+n}(-; G).$$

Equivalently, $\Phi$ assigns to each space $X$ a map $\Phi_X : H^n(X;\pi) \to H^q(X;G)$ which is natural with respect to $X$.
A *stable cohomology operation* $\Phi$ of degree $n$ is a sequence of cohomology operations $\{\Phi^q\}$ of type $(q,\pi,G)$ ranging over $q \in \mathbb{N}$ so that for every space $X$ the diagram,

$$\begin{array}{ccc} H^q(X;\pi) & \xrightarrow{\Phi^q} & H^{q+n}(X;G) \\ \downarrow{\scriptstyle \Sigma} & & \downarrow{\scriptstyle \Sigma} \\ H^{q+1}(\Sigma X;\pi) & \xrightarrow{\Phi^{q+1}} & H^{q+n+1}(\Sigma X;G) \end{array}$$

is commutative.

Recall a functor $F : \mathbf{C} \to \mathbf{Set}$ is *representable* if there is an element $X \in \mathbf{C}$ so that $F(Y) = \mathrm{Hom}(X,Y)$ for all $Y \in \mathbf{C}$. Similarly for contravariant functors. The possible cohomology operations are characterized by the following foundational result of category theory.

LEMMA 4.66 (Yoneda). *Let $F, F'$ be functors from a category $\mathbf{C}$ to $\mathbf{Set}$. Suppose $F$ is representable: $F(-) = \mathrm{Hom}(-,Z)$. Then there is a canonical bijection between natural tranformations $\Phi : F \to F'$ and elements $\varphi \in F'(Z)$.*

PROOF. Consider $\mathrm{id}_Z \in F(Z) = \mathrm{Hom}(Z,Z)$. Given $\Phi$, let $\varphi = \Phi(\mathrm{id}_Z)$. Conversely, given $\varphi$, define $\Phi_X : F(X) \to F'(X)$ by $\Phi_X(f) = f^*(\varphi)$ for any $f \in \mathrm{Hom}(X,Z)$, where $f^* = F'(f) : F'(Z) \to F'(X)$. One can check this is natural and these are inverse constructions. □

COROLLARY 4.67. *Given representable functors $\mathrm{Hom}(-,Z)$ and $\mathrm{Hom}(-,Z')$, there is a canonical bijection between natural transformations $\Phi : \mathrm{Hom}(-,Z) \to \mathrm{Hom}(-,Z')$ and elements $\varphi \in \mathrm{Hom}(Z,Z')$.*

**Theorem 4.68**

Cohomology operations of type $(q,\pi,G)$ and degree $n$ are in canonical bijection with elements of $H^{q+n}(K(\pi,q);G)$.

PROOF. Such a cohomology operation is a natural transformation $H^q(-;\pi) \to H^{q+n}(-;G)$. These are both representable by Theorem 4.59. So this is the same as a natural transformation,

$$[-, K(\pi,q)] \to [-, K(G,q+n)].$$

By the above corollary, these are in bijection with $[K(\pi,q), K(G,q+n)]$. Again by Theorem 4.59, this is isomorphic to $H^{q+n}(K(\pi,q);G)$. □

COROLLARY 4.69. *All non-zero cohomology operations have degree $n \geq 0$.*

PROOF. By Hurewicz, the first non-trivial cohomology group of $K(\pi,q)$ is in degree $q$. By our theorem, we need to have $n \geq 0$ to have any non-zero cohomology operations. □

This gives us a good reason to want to understand the homology of Eilenberg-MacLane spaces. We will study some of this with spectral sequences. For example, it will be an exercise later to prove that for $\pi$ a finitely generated abelian group of rank $r$,

$$H^*(K(\pi,n);\mathbb{Q}) = \begin{cases} \Lambda_{\mathbb{Q}}(x_1,\ldots,x_r),\ |x_i| = n & : n \text{ odd} \\ \mathbb{Q}[x_1,\ldots,x_r],\ |x_i| = n & : n \text{ even.} \end{cases}$$

From this one can deduce for example if $q$ is odd then the only cohomology operations $H^q(-,\mathbb{Z}) \to H^{q+n}(-,\mathbb{Q})$ have $n = 0$ and come induced from the inclusion $\mathbb{Z} \to \mathbb{Q}$ multiplied by a rational constant.

As another example, one can show any degree one cohomology operation of type $(n,C,A)$ is a Bockstein homomorphism (i.e. the connecting homomorphism in the induced long exact sequence) from a short exact sequence $0 \to A \to B \to C \to 0$. Over a field, there are cohomology operations given by taking cup product powers of a cohomology class : $x \mapsto x \smile \cdots \smile x$.

Over $\mathbb{Z}_2$ there are many novel cohomology operations.

**Definition 4.70: Steenrod Squares**

For $n \geq 0$ there are stable cohomology operations

$$Sq^n : H^q(X;\mathbb{Z}_2) \to H^{q+n}(X;\mathbb{Z}_2),$$

called *Steenrod squares*. These are uniquely determined by the following properties.

(i) $Sq^0$ is the identity.

(ii) $Sq^n(x) = x^2$ when $x$ has degree $n$ and $Sq^n(x) = 0$ if $x$ has degree less than $n$.

(iii) The Steenrod squares applied to a product satisfy the Cartan formula:

$$Sq^n(xy) = \sum_{i+j=n} Sq^i(x)Sq^j(y).$$

There is a multiplication on all $\mathbb{Z}_2$ cohomology operations under composition forming the *Steenrod algebra*. This Steenrod squares are not free in this algebra. The multiplicative structure of the Steenrod squares is determined by the *Adem relations*:

$$Sq^i Sq^j = \sum_{0 \leq k \leq \lfloor i/2 \rfloor} \binom{j-k-1}{i-2k} Sq^{i+j-k} Sq^k.$$

It turns out however that the Steenrod squares multiplicatively generate the Steenrod algebra, so that all $\mathbb{Z}_2$ cohomology operations can be described as linear combinations of iterated Steenrod squares. Using the Adem relations, we can obtain a subset of Steenrod squares that freely generate the cohomology of $K(\mathbb{Z}_2, q)$. Let $F_n \in H^n(K(\mathbb{Z}_2, n); \mathbb{Z}_2)$ be the fundamental/tautological class defined in the previous section.

**Theorem 4.71: (Serre)**

$H^*(K(\mathbb{Z}_2, q); \mathbb{Z}_2)$ is a polynomial algebra generated by certain iterates of Steenrod operations applied to $F_q$. More precisely, if $Sq^I = Sq^{i_1} \cdots Sq^{i_j}$ for $I = (i_1, \ldots, i_j)$ a multi-index, the generators are $Sq^I F_q$, where the multi-index $I$ satisfies $i_r \geq 2i_{r+1}$ and $i_1 < i_2 + \ldots + i_j + q$.

Remark 4.72. Note that these generators $Sq^I$ do not *freely* generate the Steenrod algebra (they do as a $\mathbb{Z}_2$-basis). It is only when applied to $F_q$ that they give a free generating set of the $q$th part of the Steenrod algebra. △

We will later give a construction of the Steenrod squares from spectral sequences. Although we will not prove their properties, we will give some intuition for this result of Serre.

Remark 4.73. Steenrod also constructed corresponding stable operations on $\mathbb{Z}_p$ homology for primes $p > 2$ called the *reduced p-th power operations*

$$P^n : H^q(X; \mathbb{Z}_p) \to H^{q+2n(p-1)}(X; \mathbb{Z}_p).$$

These have the same characterizing properties as $Sq^n$ except that (ii) is replaced with the fact $P^n(x) = x^p$ for $|x| = 2n$ and 0 for $|x| < 2n$. They also satisfy similar but more complicated Adem relations. It was proved by Cartan that $H^*(K(\mathbb{Z}_p, q); \mathbb{Z}_p)$ is multiplicatively generated by the reduced power operations along with the Bockstein homomorphism $\beta_p$ coming from the short exact sequence $0 \to \mathbb{Z}_p \to \mathbb{Z}_{p^2} \to \mathbb{Z}_p \to 0$, and the cohomology is in fact a polynomial algebra generated by some collection of compositions of the $P^n$ possibly composed with $\beta_p$ (for $p = 2$, $\beta_2 = Sq^1$). △

*Capping and Killing Spaces.* Recall in our construction of $K(G,n)$ we glued on cells of dimensions $n+2$ and higher to kill all higher homotopy groups. None of this was special to the Eilenberg–MacLane space and we can do the same to kill the higher homotopy groups of any CW complex $X$.

PROPOSITION 4.74. *For any CW complex $X$ and any $n \in \mathbb{N}$, there is a CW complex $X^n$ and a cofibration $f_n : X \to X^n$ so that $(f_n)_* : \pi_i(X) \to \pi_i(X^n)$ is an isomorphism for $i \leq n$ and so that $\pi_j(X^n) = 0$ for $j > n$.*

**Definition 4.75**

The space $X^n$ given in the above proposition is called the $n$th *capping space* of $X$. By construction,

$$\pi_i(X^n) = \begin{cases} \pi_i(X) & i \leq n \\ 0 & i > n. \end{cases}$$

EXERCISE 4.76. Show that the capping space $X^n$ as described in the above proposition is homotopically unique. Moreover for any pair of capping space $X_1^n, X_2^n$, the cofibration $X \to X_2^n$ factors up to homotopy through the cofibration $X \to X_1^n$ and a homotopy equivalence $X_1^n \to X_2^n$. ▲

Let $X$ be $(n-1)$-connected and consider our capping space $f_n : X \to X^n$. By definition, $X^n$ is a cellular $K(\pi_n(X), n)$. We may homotopy $f_n$ into a fibration with homotopy fibre $X|_{n+1}$. One can iteratively apply the construction to kill off the $n+2$ homotopy group and so on, and obtain a space $X|_m$ for any $m \geq n$.

**Definition 4.77**

The space $X|_m$ above is called the $m$th *killing space* of $X$. By the homotopy long exact sequence of a fibration, it must satisfy,

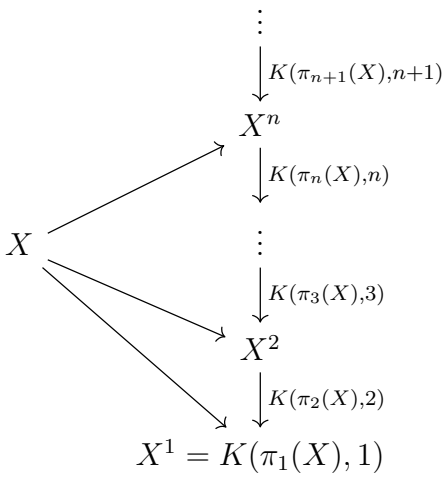
$$\pi_j(X|_m) = \begin{cases} \pi_j(X) & j \geq m \\ 0 & j \leq m. \end{cases}$$

Because $X|_{n+1}$ is the fibre of $f_n$, we obtain a map $X|_{n+1} \to X$. We can homotope this to a fibration and the fibre will be $K(\pi_n(X), n-1)$. Similarly there is a map $X|_{m+1} \to X|_m$ for each $m$ and composing them gives a map $X|_m \to X$.

EXAMPLE 4.78. Consider $S^3|_4 \to S^3$, a fibration with fibre $K(\mathbb{Z}, 2) = \mathbb{C}P^\infty$. By the Hurewicz theorem, $\pi_4(S^3) = \pi_4(S^3|_4) = H^4(S^3|_4)$. Once we have the

Serre spectral sequence, we will be able to use the homology of $S^3$ and $\mathbb{C}P^\infty$ to deduce the homology of $S^3|_4$ and hence compute $\pi_3(S^4)$, the second stable homotopy group. ▲

Given a capping space $X \to X^n$ and $m > n$, we can consider the capping space $X^n \to (X^n)^m$. Looking at the homotopy groups, we conclude by exercise 4.76 that $(X^n)^m$ is homotopy equivalent to $X^m$. Hence we conclude that we can assemble all the capping spaces into a sequence of maps. Further, we can homotope the maps to be fibrations for which the fibres will be Eilenberg–MacLane spaces. We obtain a diagram like the following.




This is called a *Postnikov system* or *Postnikov tower* for $X$. It can be seen as an attempt to build $X$ from Eilenberg–MacLance spaces through a series of fibrations. Indeed $X$ looks something like a limit of this tower, which we can make precise with the aid of the following lemma.

LEMMA 4.79. *Given a sequence of fibrations $\cdots \to X_2 \to X_1$, the natural map*

$$\pi_i(\varprojlim X_n) \to \varprojlim \pi_i(X_n),$$

*is surjective. It is injective if the maps $\pi_{i+1}(X_n) \to \pi_{i+1}(X_{n-1})$ are surjective for large $n$.*

PROOF. An element of $\varprojlim \pi_i(X_n)$ is given by a collection of spheroids $f_n : S^i \to X_n$ compatible in the sense that the fibration $p_n : X_n \to X_1$ carries $f_n$ to $f_{n-1}$ up to homotopy. By definition of a fibration, we may homotope $f_n$ so that $p_n \circ f_n = f_{n-1}$. We may do this inductively for all $n$, which gives the surjectivity of our map.

Now we prove injectivity. We may ignore a finite set of terms from our inverse limit, and so we may assume all the maps $\pi_{i+1}(X_n) \to \pi_{i+1}(X_{n-1})$ are surjective. Suppose we have a spheroid $f : S^i \to \varprojlim X_n$ so that each map $f_n : S^i \to X_n$ is null-homotopic via some $F_n : D^{i+1} \to X_n$.

We have $p_n \circ F_n = F_{n-1}$ when restricted to $S^i$, hence we may glue them together to a map $g_{n-1} : S^{i+1} \to X_{n-1}$. Since $\pi_{i+1}(X_n) \to \pi_{i+1}(X_{n-1})$ is surjective, we may pick $F_n$ so that $g_{n-1}$ is null homotopic and thus $p_n \circ F_n \cong F_{n-1} \mathrm{rel} S^i$. By the relative homotopy lifting property, we can ensure $p_n \circ F_n = F_{n-1}$. Doing this inductively for all $n$ to obtain a null homotopy of $f$. □

COROLLARY 4.80. *The inverse limit of the Postnikov tower for $X$ is weak homotopy equivalent to $X$. Hence $X$ is the CW approximation for $\varprojlim X^n$.*

PROOF. The lemma above implies $\varprojlim X^n$ has the correct homotopy groups and moreover the map $X \to \varprojlim X^n$ induces an isomorphism on all these homotopy groups (since it does for $n$ large enough). □

The homotopy fibres of the maps $X \to X^n$ are the killing spaces $X|_n$. We can form a sequence out of them, and homotope all the maps to be fibrations,

$$* \to \cdots \to X|_{n+1} \xrightarrow{K(\pi_{n+1}(X),n)} X|_n \to \cdots \to X|_2 \xrightarrow{K(\pi_2(X),1)} X|_1 \to X.$$

This is called the *Whitehead tower* of $X$. It is dual to the Postnikov tower and gives $X$ as a colimit of a tower of spaces again built from fibrations with Eilenberg–MacLane spaces as the fibres. Note that $X|_1$ has the same homotopy groups as the universal cover of $X$ so we can also see this tower as a generalization of the universal covering construction where we find $n$-connected spaces $X|_n$ fibered over $X$.

We will not prove it, but it turns out that if (and only if) $X$ is homotopically simple, then each of the fibrations $X^{n+1} \to X^n$ in the Postnikov tower can be extended to fibre sequences,

$$K(\pi_{n+1}(G), n+1) \to X^{n+1} \to X^n \to K(\pi_{n+1}(G), n+2).$$

In this case, $X^{n+1}$ is the pullback under $X^n \to K(\pi_{n+1}(G), n+2)$ of the pathspace fibration of $K(\pi_{n+1}(G), n+2)$ and so the homotopy type of $X^{n+1}$ is determined by the homotopy class of the map $X^n \to K(\pi_{n+1}(G), n+2)$.

This defines an element $k^n \in [X^n, K(\pi_{n+1}(G), n+2] \cong H^{n+2}(X^n; \pi_{n+1}(X))$ called the *nth Postnikov $k$ invariant*. Because $k^n$ tells us the homotopy type of $X^{n+1}$ and the homotopy class of the fibration $X^{n+1} \to X^n$, we can recover the weak homotopy type of $X$ from its homotopy groups along with some obstruction-type classes which tell us how to fibre the spaces $K(\pi_n(X), n)$ over

each other to re-obtain $X$. For example, if all the $k$ invariants vanish, then these fibrations are trivial and $X$ is the product of the Eilenberg–MacLane spaces corresponding to each of its homotopy groups.

## 5. Spectral Sequences

Spectral sequences are powerful tools to compute the homology of some graded (or more generally filtered) complex (or even more generally differential group) in homological algebra. There are several important spectral sequences in algebraic topology: the Serre and Eilenberg–Moorse spectral sequences for fibrations, the Adams spectral sequence for stable homotopy, and the Atiyah–Hirzebruch spectral sequence for K-theory and generalized homology. We will focus on the Serre spectral sequence, which is the easiest and of the greatest importance. The initial notational framework of spectral sequences is very daunting, but most of it can be ignored, and once understood they become an extremely useful tool to have in attempting computations. We will mostly omit proofs of the basic homological results as they are unenlightening and of little importance in practice to computations.

As some elementary motivation for the construction, consider a complex $C_*$ of abelian groups with a subcomplex $D_*$. Then we obtain a short exact sequence,

$$0 \to D_* \to C_* \to C_*/D_* \to 0,$$

which induces a long exact sequence in homology,

$$\cdots \to H_n(D_*) \to H_n(C_*) \to H_n(C_*/D_*) \xrightarrow{\partial_n} H_{n-1}(D_*) \to \cdots .$$

Suppose we know the homologies $H_*(D_*)$ and $H_*(C_*/D_*)$. We can consider the homology of the two term chain complex,

$$H_*(C_*/D_*) \xrightarrow{\partial_*} H_*(D_*),$$

which will be two groups $G_1H_*$ and $G_0H_*$. Then, decomposing the long exact above, we obtain short exact sequences,

$$0 \to G_0H_* \to H_*(C_*) \to G_1H_* \to 0.$$

This determines $H_*(C_*)$ up to an extension problem.

The goal of spectral sequences will be to generalize this from a subcomplex to a sequence of nested subcomplexes, i.e. a filtration. Computing the homology of the associated graded groups through a sequence of approximations will yield the homology of our original complex, up to a series of extension problems.

**5.1. Spectral Sequences in Homological Algebra.** We introduce spectral sequences in full generality. This discussion can be freely skimmed.

**Definition 5.1**

A *differential abelian group* $(F, \mathrm{d})$ is an abelian group $F$ equipped with a differential $\partial : F \to F$ so that $\mathrm{d}^2 = 0$. A *filtration* on $F$ is a nested sequence of subgroups $0 \subset F_0 \subset F_1 \subset \cdots$ which union to all of $F$ and so that $\mathrm{d} : F_p \to F_p$ preserves the filtration. In this case, $F$ has an *associated graded group*,

$$\mathrm{Gr}F := \bigoplus_p F_p/F_{p-1}, \quad \mathrm{Gr}_n F := F_n/F_{n-1}.$$

In the cases we are interested in $F = \bigoplus_n C_n$ will be a chain complex. A *graded filtered differential group* $F$ is a filtered differential abelian group so that $F = \bigoplus_n C_n$ and so that the differential restricts to a map $\mathrm{d} : C_n \to C_{n-1}$. Further, we ask the filtration and grading are compatible in the sense that,

$$F_p = \bigoplus_n F_p \cap C_n.$$

For ease of notation, we abbreviate $F_p \cap C_n$ by $F_p C_n$.

*Spectral Sequence of a Filtered Differential Group.* Associated to a filtered differential group is its *spectral sequence*, which is a sequence of graded differential groups $(E^r, \mathrm{d}_r)$ for $r = 0, 1, \ldots, \infty$. In terms of the grading, $E^r = \bigoplus_p E^r_p$ and the differential restricts to a map $\mathrm{d}^r_p : E^r_p \to E^r_{p-r}$.

The first term in the sequence is $E^0_p = \mathrm{Gr}_p F$ with differential $\mathrm{d}^0$ given by the overall differential $\mathrm{d}$ on $F$ descended to the quotient. We recursively define $(E^{r+1}, \mathrm{d}_{r+1})$ to be the homology of the previous term in the sequence: $E^{r+1} = H_p(E^r_p, \mathrm{d}^r)$. The differential $\mathrm{d}^{r+1}$ is inherited from the original $\mathrm{d}$ by whatever "remnant" is left in the quotient group we have defined.

More explicitly, we have,

$$E^r_p = \frac{F_p \cap \mathrm{d}^{-1} F_{p-r}}{F_{p-1} \cap \mathrm{d}^{-1} F_{p-r} + F_p \cap \mathrm{d} F_{p+r-1}}.$$

And the differential,

$$\frac{F_p \cap \mathrm{d}^{-1} F_{p-r}}{F_{p-1} \cap \mathrm{d}^{-1} F_{p-r} + F_p \cap \mathrm{d} F_{p+r-1}} \xrightarrow{\mathrm{d}^r} \frac{F_{p-r} \cap \mathrm{d}^{-1} F_{p-2r}}{F_{p-r-1} \cap \mathrm{d}^{-1} F_{p-2r} + F_{p-r} \cap \mathrm{d} F_{p-1}} = E^r_{p-r},$$

is just the differential d with its induced action on the quotient. It is elementary but tedious to check this is all well defined and matches the description above that $E^{r+1}$ is the homology of $E^r$.

What happens as $r \to \infty$? Using the expression for $E^r$ above, we see heuristically,

$$E_p^\infty = \frac{F_p \cap \mathrm{d}^{-1}(0)}{F_{p-1} \cap \mathrm{d}^{-1}(0) + F_p \cap \mathrm{d}F} = \mathrm{Gr}_p(H(F, \mathrm{d})).$$

One often speaks of the spectral sequence *converging* to the homology of $F$ and one writes,

$$E_p^1 = H(\mathrm{Gr}_p(F)) \implies \mathrm{Gr}_p H(F).$$

Sometimes one omits the Gr on the right side and says the spectral sequence converges to $H(F)$, even if this is technically not true.

We say that the spectral sequence *degenerates* or *collapses* at index $k$ if $\mathrm{d}^i = 0$ for $i \geq k$, and hence $E^k = E^{k+1} = \cdots$. In this case we will obtain $E^\infty$ after finitely many steps. If the filtration is finite, then the spectral sequence always collapses after finitely many pages. In general, the sequence may never collapse, although the $E^\infty$ term can always be realized as a direct limit since the filtered differential $\mathrm{d}_p^r$ will always collapse for fixed $p$ and large enough $r$. We will not be so interested in this subtlety as in the cases we apply the spectral sequence will almost always degenerate.

*Spectral Sequence of a Graded Filtered Differential Group.* In the cases we are interested in, our differential abelian group will be a chain complex, and hence will be graded by degree. So suppose that $F = \bigoplus_q C_q$ is a graded filtered differential group. Because the filtration and grading are compatible, the associated spectral sequence $E^r$ of $F$ now inherits the $q$ grading from the $C_q$ and hence will be bigraded,

$$E_{p,q}^r := \frac{F_p C_{p+q} \cap \mathrm{d}^{-1}(F_{p-r} C_{p+q-1})}{[F_{p-1} C_{p+q} \cap \mathrm{d}^{-1}(F_{p-r} C_{p+q-1})] + [F_p C_{p+q} \cap \mathrm{d}(F_{p+r-1} C_{p+q+1})]}.$$

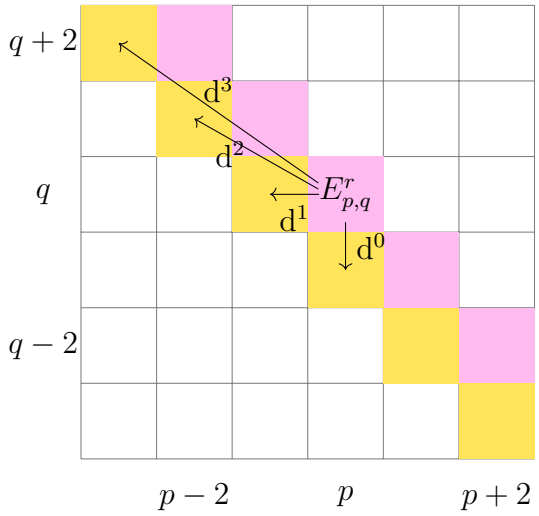
The differential is also compatible with this bigrading and restricts to a map,

$$\mathrm{d}_{p,q}^r : E_{p,q}^r \to E_{p-r,q+r-1}^r.$$

It is again elementary but tedious to check this all works.

This is a lot of indices to keep track of in our spectral sequence. Fortunately, there is a graphical way to efficiently organize all this information. We refer to the total group $\bigoplus_{p,q} E_{p,q}^r$ as the $r$th *page* of the spectral sequence. Each page of the spectral sequence is represented by a grid of the terms $E_{p,q}^r$, with $p$ increasing along the horizontal axis and $q$ increasing along the vertical axis. The terms along a downwards diagonal will all have the same *total degree*

$n = p + q$. See the following picture showing the differentials on the first few pages. Diagonals of fixed total degree are shown in pink and yellow.




On each page the differentials all point in the same direction. We obtain the next page by taking the homology of all the diagonal complexes on the page. To help remember how the differentials go, you can remember the second differential $\mathrm{d}^2$ (usually the first one we care about) is a knight's move and later differentials are extended knight's moves, where all differentials need to decrease the total degree by one.

What happens as $r \to \infty$ in this case? We have,

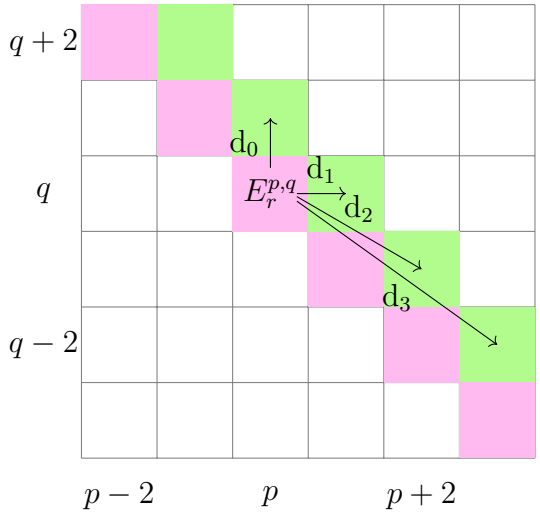
$$E^\infty_{p,q} = \frac{F_p C_{p+1} \cap \mathrm{d}^{-1}(0)}{F_{p-1} C_{p+q} \cap \mathrm{d}^{-1}(0) + F_p C_{p+q} \cap \mathrm{d}F} = \frac{F_p H_n}{F_{p-1} H_n} = \mathrm{Gr}_p H_{p+q},$$

where $n = p+q$ is the total degree and $H_n$ is the $n$th homology group of $F$ with respect to the grading of $F$. Again, in the cases we care about, the sequence usually degenerates after finitely many pages and we converge to $E^\infty$.

The relation between the actual homology of $F$ and these associated graded pieces can be a little subtle. But if on the $E^\infty$ page the diagonal of total degree $n$ is empty, then we must have $H_n(F) = 0$. And if that diagonal has only one non-zero entry, that entry must equal $H_n(F)$.

*Spectral Sequences in Cohomology.* Of course we can do a very similar construction in cohomology with a differential that increases the graded degree. A trick to make the construction especially easy is to notice that by flipping the indexing of the grading, the cohomological differential becomes a homological differential. So, we may flip the indexing, apply the techniques above, and the flip back to get our cohomological spectral sequence. The result will

be diagrams that look the same, except with the differentials pointing in the opposite direction. Because we are dealing with cohomology, the grading indices should now be superscripts. Correspondingly, we move our page indices to subscripts. See the picture below.




*Naturality.* Suppose we have a homomorphism of filtered differential groups $C \to C'$ which is compatible with the filtrations. If our groups are graded, we ask the homomorphism be compatible with that as well.

Then it is clear we will induce morphisms of the corresponding spectral sequences $E \to E'$ on each page. This extends to the page at infinity, where the morphism $\mathrm{Gr}H(C) \to \mathrm{Gr}H(C')$ should be the usual one induced on homology (and then to the associated graded quotient).

**5.2. Basic Examples of Spectral Sequences.** We will begin with a few simple spectral sequences to get a sense for the subject. One can also refer back to our usage of spectral sequences in proving the universal coefficient theorem and Künneth formula.

*The Hochschild–Serre Spectral Sequence.* This is a spectral sequence for computing group cohomology of a group $G$ by studying a normal subgroup $H$ and the quotient $G/H$, which may be considerably simpler. We will look at the particular case of Lie algebra cohomology.

Let $\mathfrak{g}$ be a complex Lie algebra and $M$ a $\mathfrak{g}$-module, which will be our coefficient space. We define a cochain complex with $C^n(\mathfrak{g}; M) = \mathrm{Hom}_{\mathbb{C}}(\Lambda^n \mathfrak{g}, M)$ and

differential $\mathrm{d} : C^n(\mathfrak{g}; M) \to C^{n+1}(\mathfrak{g}; M)$ given by,

$$\mathrm{d}c(g_1 \wedge \cdots \wedge g_{n+1}) = \sum_{s<t} (-1)^{s+t-1} c([g_s, g_t] \wedge g_1 \wedge \cdots \widehat{g_s} \cdots \widehat{g_t} \cdots \wedge g_{n+1}) - \sum_u (-1)^u g_u c(g_1 \wedge \cdots \widehat{g_u} \cdots \wedge g_{n+1}),$$

where $\widehat{\cdot}$ denotes an omitted term.

The *Lie algebra cohomology* $H^n(\mathfrak{g}; M)$ of $\mathfrak{g}$ is the cohomology of the complex $C^n(\mathfrak{g}; M)$.

To apply spectral sequence techniques, consider a Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$. Define a filtration on our cochain spaces by,

$$F^p C^{p+q}(\mathfrak{g}; M) = \left\{ c \in C^{p+q}(\mathfrak{g}; M) : c(g_1 \wedge \cdots \wedge g_{p+q}) = 0 \quad \text{if} \quad g_1, \ldots, g_{q+1} \in \mathfrak{h} \right\}.$$

It is simple enough to see that this filtration is compatible with the differential and grading. Hence we obtain a spectral sequence converging to $\mathrm{Gr} H^n(\mathfrak{g}; M)$ called the *Hochschild–Serre spectral sequence*.

On the $E_0$ page we have,

$$E_0^{p,q} = \frac{F^p C^{p+q}}{F^{p-1} C^{p+q}} = \mathrm{Gr}_p(C^{p+q}) = C^q(\mathfrak{h}; \mathrm{Hom}_{\mathbb{C}}(\Lambda^p(\mathfrak{g}/\mathfrak{h}), M)),$$

where the verification of the last equality is left to the reader. The effect of the differential $\mathrm{d}_0$ is to increase the degree of $q$ by one, and so it is no surprise that, $E_1^{p,q} = H^q(\mathfrak{h}; \mathrm{Hom}_{\mathbb{C}}(\Lambda^p(\mathfrak{g}/\mathfrak{h}), M))$.

In the case where $\mathfrak{h}$ is an ideal, we can rewrite this as, $E_1^{p,q} = C^p(\mathfrak{g}/\mathfrak{h}; H^q(\mathfrak{h}; M))$. Hence taking homology,

$$E_2^{p,q} = H^p(\mathfrak{g}/\mathfrak{h}; H^q(\mathfrak{h}; M)) \implies H^{p+q}(\mathfrak{g}; M) = E_\infty^{p+q}.$$

So the cohomology of $\mathfrak{g}$ can be deduced from the homology of $\mathfrak{h}$ and a quotient $\mathfrak{g}/\mathfrak{h}$ using a spectral sequence. For proofs of these claims, see the original paper [**HS**].

*Spectral Sequence of a Filtered Topological Space.* Suppose we have a topological space $X$ filtered by subspaces:

$$\varnothing \subset X_0 \subset X_1 \subset \cdots \subset X_n \subset \cdots = X.$$

If the filtration is infinite, $X$ should have the weak topology with respect to the filtration (so $X$ is the colimit of inclusion maps).

The space of singular chains on $X$ has a filtration with $p$th part $F_p C_n := C_n(X_p)$, consisting of singular chains living in $X_p$. This makes the singular

chains on $X$ into a graded filtered group and we can study the corresponding spectral sequence.

Note that $E^0_{p,q} = \mathrm{Gr}_p(C_{p+q}(X)) = C_{p+q}(X_p, X_{p-1})$. The differential $\mathrm{d}_0$ is just inherited from the usual differential map on homology and so,

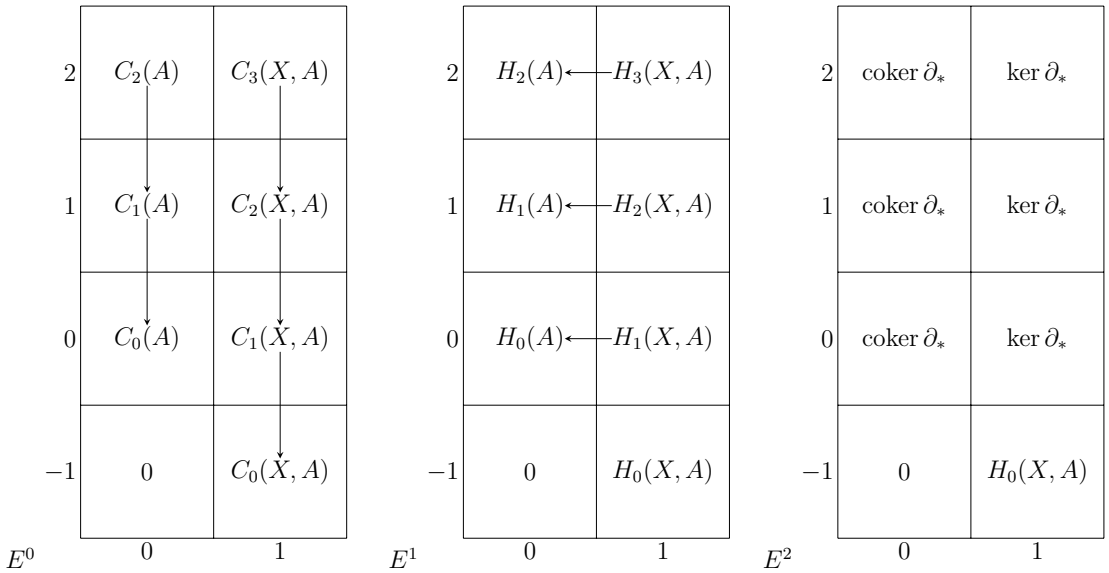
$$E^1_{p,q} = H_{p+q}(X_p, X_{p-1}).$$

The differential on the first page,

$$\mathrm{d}^1 : E^1_{p,q} = H_{p+q}(X_p, X_{p-1}) \to E^1_{p-1,q} = H_{p+q-1}(X_{p-1}, X_{p-2}),$$

is the connecting homomorphism in the long exact sequence of the triple $(X_p, X_{p-1}, X_{p-2})$. This spectral sequence will converge to the associated graded group of $H(X)$.

Let us briefly consider the special case of a pair, where our filtration of $X$ is $\varnothing \subset A \subset X$. Then the first three pages in our spectral sequence have the following form.




After this, the sequence degenerates. We conclude that $H_n(X)/\operatorname{coker}(\partial_*) \cong \ker \partial_*$. This is just equivalent to the long exact sequence of a pair,

$$\cdots \to H_{n+1}(X, A) \xrightarrow{\partial_*} H_n(A) \to H_n(X) \to H_n(X, A) \xrightarrow{\partial_*} H_{n-1}(A) \to \cdots .$$

*Spectral Sequence of Cellular Complex.* Let $X$ be a CW complex. As before, consider a filtration of $X$ but now specifically given by the successive skeleta of $X$:

$$\varnothing \subset X_0 \subset X_1 \subset \cdots \subset X_n \subset \cdots = X.$$

As before, we get a spectral sequence converging to the homology of $X$. We see that,

$$E^1_{p,q} = H_{p+q}(X_p, X_{p-1}) = \widetilde{H}_{p+q}(X_p/X_{p-1}) = \begin{cases} \mathcal{C}_p(X) & q = 0 \\ 0 & q \neq 0. \end{cases}.$$

Hence the $E^1$ page of our spectral sequence looks like the following.

| | | | | |
|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 |
| 0 | $\mathcal{C}_0(X)$ | $\mathcal{C}_1(X)$ | $\mathcal{C}_2(X)$ | $\mathcal{C}_3(X)$ |
| $E^1$ | 0 | 1 | 2 | 3 |

Referring back to our discussion of cellular homology, one sees the map here is precisely the boundary map in the cellular chain complex (alternatively we could use this spectral sequence to define the complex). Thus the $E^2$ page looks like the following.

| | | | | |
|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 |
| 0 | $H_0^{\text{cell}}(X)$ | $H_1^{\text{cell}}(X)$ | $H_2^{\text{cell}}(X)$ | $H_3^{\text{cell}}(X)$ |
| $E^2$ | 0 | 1 | 2 | 3 |

The spectral sequence degenerates here. There is only one entry in each diagonal and hence,

$$H_i(X) = H_i(\mathcal{C}_*(X), \partial_*),$$

which gives another proof that singular and cellular homology agree.

**5.3. The Serre Spectral Sequence.** We now move on to the main subject of this chapter. The Serre spectral sequence allows us to relate the homology of a total space of a fibration with the homology of the base and the fibre.

*The Homological Sequence.* Suppose $\pi : E \to B$ is a Serre fibration with base $B$ a CW complex and fibre $F$. We will need to assume our fibration is *homologically simple*. By this we mean that the homotopy equivalence from Theorem 1.10 between fibres of $\pi$ is independent of the path used to define it.

More succinctly, we might say we require $\pi_1(B)$ to act trivially on $H_*(F)$. In practice it is easiest to verify this by assuming $B$ is simply connected.

There is a filtration of $B$ by its skeleta: $\varnothing \subset B_0 \subset B_1 \subset \cdots = B$. This in turn determines a filtration of $E$ by $E_p = \pi^{-1}(B_p)$. From this we get a filtration of the singular chain complex of $E$:

$$0 \subset C_*(E_0) \subset C_*(E_1) \subset \cdots \subset C_*(E),$$

by singular chains landing in $E_p$. This makes $C_*(E)$ a graded filtered differential group, and hence it has a spectral sequence. The zeroth page of the sequence is $E^0_{p,q} = C_{p+q}(E_p, E_{p-1})$ with standard differential,

$$\partial : C_{p+q}(E_p, E_{p-1}) \to C_{p+q-1}(E_p, E_{p-1}).$$

Hence $E^1_{p,q} = H_{p+q}(E_p, E_{p-1})$.

Now, for each $p$-cell of $B_p$, pick a small ball $D^p_\alpha$. The inclusion,

$$B^{p-1} \hookrightarrow B^p \setminus \bigsqcup_\alpha D^p_\alpha$$

is a homotopy equivalence. Hence,

$$\pi^{-1}(B^{p-1}) = E_{p-1} \hookrightarrow \pi^{-1}\left(B^p \setminus \bigsqcup_\alpha D^p_\alpha\right)$$

is a weak homotopy equivalence from the long exact sequence of a fibration and the five lemma. Hence,

$$H_{p+q}(E_p, E_{p-1}) = H_{p+q}\left(E_p, E_p \setminus \pi^{-1}(\bigsqcup_\alpha D^p_\alpha)\right).$$

By excision,

$$\begin{aligned} &= H_{p+q}\left(\pi^{-1}(\bigsqcup_\alpha D^p_\alpha), \pi^{-1}(\bigsqcup_\alpha \partial D^p_\alpha)\right) \\ &= \bigoplus_\alpha H_{p+q}\big(\pi^{-1}(D^p_\alpha), \pi^{-1}(\partial D^p_\alpha)\big). \end{aligned}$$

By Feldbau's lemma, the fibration trivializes on each disk $D^p_\alpha$. Using our homological simplicity to speak unambiguously about the homology of fibres on each path component and identifying fibres on different path components by an arbitrary isomorphism,

$$= \bigoplus_\alpha H_{p+q}(F \times D^p_\alpha, F \times \partial D^p_\alpha).$$

By the Künneth formula,

$$\begin{aligned} &= \bigoplus_{\alpha} H_q(F) \\ &= \mathcal{C}_p(B; H_q(F)). \end{aligned}$$

The boundary map on the $E^1$ page is,

$$E^1_{p,q} = \mathcal{C}_p(B; H_q(F)) \xrightarrow{\partial} \mathcal{C}_{p-1}(B; H_q(F)) = E^1_{p-1,q}.$$

Comparing the explicit construction of this boundary map to the cellular boundary map reveals they are the same. Hence,

$$E^2_{p,q} = H_p(B; H_q(F)).$$

This spectral sequence should converge to the associated graded $\mathrm{Gr}_p H_{p+q}(E)$ of the homology of the total space $E$. If the spaces $B$ or $F$ are finite dimensional, then the spectral sequence will degenerate by the page $\min\{\dim B, \dim F\} + 1$. If the spaces are infinite, the sequence may not degenerate at a finite page. But at a square $E^r_{p,q}$, all differentials involving this square are zero for $r > \max\{p, q+1\}$, and so on any fixed diagonal $\bigoplus_{p+q=n} E^r_{p,q}$ the sequence is constant once $r = n+2$. Hence there is no ambiguity in what it means to converge to the homology of $E$. We thus have a proof of the following central result.

**Theorem 5.2: Serre Spectral Sequence**

Let $\pi : E \to B$ be a homologically simple Serre fibration with a connected CW base and a fibre $F$. Then there is a spectral sequence,

$$E^2_{p,q} = H_p(B; H_q(F)) \implies H_{p+q}(E).$$

That is, it has $E^2$ page given by the homology of the base with homology of the fibre coefficients and converges to (the associated graded of) the homology of $E$.

As we noted before, the relationship between $H_n(E)$ and $\mathrm{Gr}_p(H_{p+q}(E))$ may be difficult to work out, since entries $E^\infty_{p,q}$ of the $E^\infty$ page are quotients $F_p H_{p+q}(E)/F_{p-1}H_{p+q}(E)$ of intermediate subgroups of $H_{p+q}(E)$ (i.e. there is still a sequence of extension problems involved in determining $H_n(E)$). But the fact that our spectral sequence is contained in the first quadrant of each page does give us some important information. We know that the entries $E^\infty_{0,n}$ in the zeroth column of the infinity page are equal to $F_0 H_n(E)$ and in particular are a subgroup of $H_n(E)$. Correspondingly, the entries $E^\infty_{n,0}$ in the zeroth row of the infinity page are equal to $H_n(E)/F_{n-1}H_n(E)$ and in particular are

quotient groups of $H_n(E)$. Thus we get canonical injections and surjections for each $n$,

$$E^\infty_{0,n} \hookrightarrow H_n(E) \twoheadrightarrow E^\infty_{n,0}. \tag{5.1}$$

Here is an important immediate corollary of our spectral sequence.

COROLLARY 5.3. *If $E \to B$ is a homologically simple fibration with fibre $F$ so that $B$ is a finite CW complex and each homology group of $F$ is finite rank then,*

$$\chi(E) = \chi(B)\chi(F).$$

PROOF. We compute from the Künneth formula,

$$\sum_{p,q}(-1)^{p+q}\dim E^1_{p,q} = \sum_{p,q}(-1)^{p+q}\dim \mathcal{C}_p(B)\dim H_q(F) = \chi(B)\chi(F).$$

But passing to homology does not affect the Euler characteristic. Hence going from the $E^1$ page to the $E^\infty$ page will not affect the answer,

$$\chi(B)\chi(F) = \sum_{p,q}(-1)^{p+q}\dim E^\infty_{p,q} = \sum_n (-1)^n \dim H_n(E) = \chi(E).$$

□

REMARK 5.4. If our fibration is not homologically simple, there is still a way to interpret the second page of our spectral sequence. Instead of considering the homology of $B$ with coefficients in $H_*(F)$, we should use a *local coefficient system* where the homology has coefficients in the locally constant sheaf $H_q(\pi^{-1}(x))$ on $B$. The details can be found in [**FF**, §22.2].

If $p : E \to B$ is a map, not necessarily a fibration, we can still get a sheaf of abelian groups on $B$ as coefficients. Namely, $\check{H}^q(p^{-1}(U);\underline{\mathbb{Z}})$ for $U$ an open set of $B$, where $\check{H}^*$ denotes sheaf or Čech cohomology and $\underline{\mathbb{Z}}$ is the usual locally constant sheaf (this is a right derived functor of the direct image $p_*$ of $\underline{\mathbb{Z}}$). Denote this sheaf by $\check{H}^q(F)$. Then there is a cohomological spectral sequence,

$$E_2^{p,q} = \check{H}^p(B;\check{H}^q(F)) \implies \check{H}^{p+q}(E;\underline{\mathbb{Z}}).$$

This is called the *Leray–Serre spectral sequence* and it directly generalizes the Serre spectral sequence as well as the local coefficient case. It is an important tool in algebraic geometry and more advanced topics in algebraic topology like equivariant cohomology. △

*Cohomology.* As we saw for the homology of any spectral sequence of a filtered complex, we can define analogously a spectral sequence for cohomology. The diagrams will look the same will differential arrows pointing in the opposite direction. We conclude by the same proof as above that we have a cohomological Serre spectral sequence,

$$E_2^{p,q} = H^p(B; H^q(F)) \implies H^{p+q}(E).$$

As with (5.1), we obtain important injection and surjection maps on the $E^\infty$ page but now reversed due to the change in direction of the filtration,

$$E_\infty^{n,0} \hookrightarrow H^n(E) \twoheadrightarrow E_\infty^{0,n}.$$

*Morphisms.* Suppose that one has a morphism $f : E \to E'$ of fibrations, meaning there is a commutative diagram of the following form.

$$\begin{array}{ccc} E & \xrightarrow{f} & E' \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi'} \\ B & \xrightarrow{g} & B' \end{array} \tag{5.2}$$

Further suppose the map on the base $g : B \to B'$ is cellular. In this case we get homomorphism between the chain complexes of $E$ and $E'$ compatible with the filtration by skeleta of $B, B'$. Thus we get a morphism of spectral sequences $f_* : E^r_{p,q} \to E'^r_{p,q}$. This has the following properties.

(1) $f_* : E^1_{p,q} \to E'^1_{p,q}$ is the usual induced map on relative homology $H_{p+1}(E_p, E_{p-1}) \to H_{p+1}(E'_p, E'_{p-1})$.

(2) The map $f_*$ on the $(r+1)$st page is the induced map on homology from the map $f_*$ on the $r$th page.

(3) The map $f_*$ on the $E^\infty$ page is the usual induced map $f_* : H_n(E) \to H_n(E')$ on the associated graded groups.

We also of course get a contravariant induced map on spectral sequences in cohomology, $f^* : E'^{p,q}_r \to E^{p,q}_r$. The corresponding properties hold.

Actually, even if the map $f : E \to E'$ gives a map $g : B \to B'$ on the base which is not cellular, we can still construct an induced map $f_*$ on spectral sequences from the second page onwards. The first step is to cellulary approximate $g$, but the full argument is more involved for a Serre fibration. We provide a full proof since the result is important, although it may freely be skipped on first reading.

PROPOSITION 5.5. *Let $f : E \to E'$ be a morphism of homologically simple fibrations (i.e. inducing a diagram like (5.2)) with CW bases. Then starting*

*from the $E^2$ page, there is a well defined and canonical induced morphism of Serre spectral sequences $f_*$.*

PROOF. We define a homotopy $G : B \times [0,1] \to G'$ between $G_0 = g$ and a cellular map $G_1 = g'$. We obtain a pullback fibration $\mathcal{E} = G^* E'$ via the following diagram,

$$\begin{array}{ccc} \mathcal{E} & \xrightarrow{\widetilde{G}} & E' \\ \downarrow{\scriptstyle F'} & & \downarrow{\scriptstyle F'} \\ B \times [0,1] & \xrightarrow{G} & B'. \end{array}$$

We denote by $\mathcal{E}_t$ the inverse image of $B \times \{t\}$ under the above fibration. By restriction, we have the following diagram,

$$\begin{array}{ccc} \mathcal{E}_1 & \xrightarrow{\widetilde{g}'} & E' \\ \downarrow{\scriptstyle F'} & & \downarrow{\scriptstyle F'} \\ B \times \{1\} & \xrightarrow{g'} & B'. \end{array} \tag{5.3}$$

By the universal property of pullback fibrations, we get a map $f' : E \to \mathcal{E}_0$ completing the following diagram,

$$\begin{array}{ccccc} E & \overset{f'}{\dashrightarrow} & \mathcal{E}_0 & \xrightarrow{\widetilde{g}} & E' \\ \downarrow{\scriptstyle F} & & \downarrow{\scriptstyle F'} & & \downarrow{\scriptstyle F'} \\ B & \xrightarrow{\mathrm{id}} & B \times \{0\} & \xrightarrow{g} & B'. \end{array} \qquad \overset{f}{\frown} \tag{5.4}$$

We have the following two diagrams coming from inclusions,

$$\begin{array}{ccc} \mathcal{E}_0 & \xhookrightarrow{i_0} & \mathcal{E} \\ \downarrow{\scriptstyle F'} & & \downarrow{\scriptstyle F'} \\ B \times \{0\} & \hookrightarrow & B \times [0,1] \end{array} \qquad \begin{array}{ccc} \mathcal{E}_1 & \xhookrightarrow{i_1} & \mathcal{E} \\ \downarrow{\scriptstyle F'} & & \downarrow{\scriptstyle F'} \\ B \times \{1\} & \hookrightarrow & B \times [0,1]. \end{array} \tag{5.5}$$

Note the morphisms of fibrations $\widetilde{g}', f', i_0$ and $i_1$ depicted in (5.3)–(5.5) have cellular bases. Hence they define morphisms of spectral sequences which we now analyze starting from the $E^2$ page. We know that on the $E^2$ page, the induced morphism is the one coming from the induced map on homology of the base and fibre. So for diagram (5.3), we have,

$$(\widetilde{g}')^2_* : H_*(B; H_*(F')) \to H_*(B'; H_*(F')),$$

induced by $g'_* : H_*(B) \to H_*(B')$ and the identity on the fibres. For diagram (5.4), we have

$$(f')^2_* : H_*(B; H_*(F)) \to H_*(B; H_*(F')),$$

induced by $h_* H_*(F) \to H_*(F')$ and the identity on the base. For the diagrams in (5.5), we see that the map on the fibres is the identity and the maps on the base are homotopy equivalences. Hence $(i_0)_*$ and $(i_1)_*$ give isomorphisms of the $E^2$ pages of the spectral sequences for the three fibrations $\mathcal{E} \to B \times [0,1]$, $\mathcal{E}_0 \to B$, and $\mathcal{E}_1 \to B$.

Composing these morphisms gives the following,

$$E^2_{p,q}(E) = H_p(B; H_q(F)) \xrightarrow{(f')^2_*} E^2_{p,q}(\mathcal{E}_0) \xrightarrow{\cong} E^2_{p,q}(\mathcal{E}_1) \xrightarrow{(\widetilde{g}')^2_*} H_p(B'; H_q(F')) = E^2_{p,q}(E').$$

It is clear from our definition of these morphisms above, that the total morphism is given by $h_* : H_q(F) \to H_q(F')$ and $g'_* : H_p(B) \to H_p(B')$. Since $g$ and $g'$ are homotopic, they induce the same maps in homology. Thus the map of $E^2$ pages, $E^2_{p,q}(E) \to E^2_{p,q}(E')$, is precisely the one we would expect for the map $f : E \to E'$ of fibrations.

To conclude, we just need to show this morphism of spectral sequences extends to all higher pages. If a morphism of spectral sequences is an isomorphism on a certain page it is an isomorphism on all later pages including $\infty$ (since the induced map on homology will always be an isomorphism). Thus (5.5) gives

isomorphisms of spectral sequences for $\mathcal{E}_0, \mathcal{E}$ and $\mathcal{E}_1$ on all pages 2 and beyond. Composing with the other morphisms, we obtain for all $p, q$ and all $k > 2$,

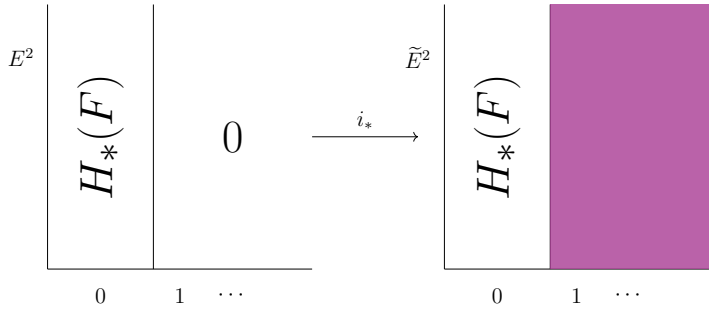
$$E^k_{p,q}(E) \xrightarrow{(f')^k_*} E^k_{p,q}(\mathcal{E}_0) \xrightarrow{\cong} E^k_{p,q}(\mathcal{E}_1) \xrightarrow{(\widetilde{g}')^k_*} E^k_{p,q}(E').$$

Thus we see that $g$ gives a morphism of Serre spectral sequences on all pages 2 and greater. □

From this we conclude that from the $E^2$ page onwards the Serre spectral sequence is independent of the cellular structure on the base. Moreover, this means if we have a homologically simple fibration $E \to B$, even if $B$ is not a CW complex, we may consider a CW approximation $B^{CW} \to B$. We can define the Serre–spectral sequence for $B^{CW}$ with respect to the pullback fibration. The long exact sequence of a fibration and the five lemma implies this gives a weak homotopy equivalence of total spaces. Thus we may extend the Serre spectral sequence to our non-CW fibration and there is no ambiguity on the $E^2$ page and beyond due to the above result on morphisms.

Remark 5.6. The original construction of Serre does not require a CW structure. In that approach, we work with homology defined in terms of cubical singular chains $\sigma : I^{p+q} \to E$. We may filter $C_{p+q}(E)$ so that $F_pC_{p+q}(E)$ is generated by cubical singular chains $\sigma : I^{p+q} \to E$ for which $\pi \circ \sigma : I^{p+q} \to B$ is independent of the last $q$ coordinates. The resulting spectral sequence of this filtration replicates the Serre spectral sequence as we have constructed it. A full proof (as well as a very good general presentation of spectral sequences) is given here. △

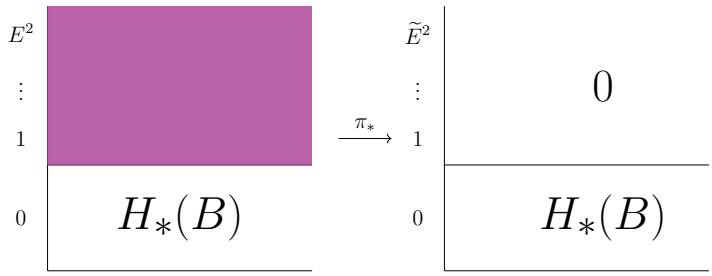
*Edge Morphisms.* There are two simple but important examples of morphisms of spectral sequences associated to any fibration $\pi : E \to B$ that we wish to study, both coming from the edges of the spectral sequence. For the first one, let $F_{x_0} \to x_0$ be the fibration $\pi : E \to B$ restricted to a fibre over a single point $x_0 \in B$. Then we have a morphism of fibrations coming from the embedding $i : F_{x_0} \hookrightarrow E$. The corresponding of morphism of spectral sequences on the $E^2$ page is represented by the embedding of the zeroth column as shown below, where the magenta region is unknown.



On the left side, the sequence degenerates: $E^2 = E^\infty$. On the right, the sequence may continue and the zeroth column $\widetilde{E}^\infty_{0,*}$ will end up as some quotient of $H_*(F)$ (since all differentials out of $E_{0,*}$ are trival). On $E^\infty$, the morphism should just be the map on homology induced by embedding $i_* : H_*(F) \to H_*(E)$. Hence, using (5.1), this map factors as follows:

$$\begin{array}{ccc} H_n(F) & \xrightarrow{i_*} & H_n(E) \\ \downarrow & \nearrow & \\ \widetilde{E}^\infty_{0,n}. & & \end{array}$$

The second morphism of fibrations we want to consider is the projection $\pi$ from $\pi : E \to B$ to the identity fibration $B \to B$. On the $E^2$ page, the corresponding morphism of spectral sequences is the projection killing all the rows above the zeroth row, as shown below.




On the right side, the sequence degenerates: $\widetilde{E}^2 = \widetilde{E}^\infty$. On the left, the sequence may continue and the zeroth row $E^\infty_{*,0}$ will be some subgroup of $H_*(B)$ (since all differentials into $E_{*,0}$ are trival). One $E^\infty$, the morphism must correspond to the map on homology induced by projection $\pi_* : H_*(E) \to H_*(B)$. Hence, again using (5.1), this map factors as follows:

$$\begin{array}{ccc} H_n(E) & \xrightarrow{\pi_*} & H_n(B) \\ \downarrow & \nearrow & \\ E^\infty_{n,0}. & & \end{array}$$

We get analogous diagrams applying these morphisms to the cohomological spectral sequence:

$$\begin{array}{ccc} H^n(B) & \xrightarrow{\pi^*} & H^n(E) \\ \downarrow & \nearrow & \\ E^{n,0}_\infty & & \end{array} \qquad \begin{array}{ccc} H^n(E) & \xrightarrow{i^*} & H^n(F) \\ \downarrow & \nearrow & \\ E^{0,n}_\infty. & & \end{array}$$

*Transgression.* Given a first quadrant spectral sequence (homological or cohomological), there are a series of maps that connect the zeroth row and column:

$$\mathrm{d}^r_{r0} : E^r_{r,0} \to E^r_{0,r-1} \quad \text{and} \quad \mathrm{d}^{0,r-1}_r : E^{0,r-1}_r \to E^{r,0}_r.$$

These maps are called *edge maps*. They are the last non-trivial map in the spectral sequence that these squares interact with. In light of (5.1), we can think of these as partially defined maps (i.e. maps from a subgroup to a quotient group),

$$\tau : E^2_{r,0} \to E^2_{0,r-1} \quad \text{and} \quad \tau : E_2^{0,r-1} \to E_2^{r,0}.$$

These maps are called *transgressions* and elements of the domain of $\tau$ are called *transgressive*. Understanding the transgressions is important for understanding the differentials of a spectral sequence. They can be thought of as generalizations of the connecting homomorphisms in the snake lemma.

In the Serre spectral sequence case, the transgressions are partially defined maps,

$$H_r(B) \to H_{r-1}(F) \quad \text{and} \quad H^{r-1}(F) \to H^r(B).$$

Thankfully these have a nice geometric interpretation.

**Theorem 5.7: Transgressions in the Serre Spectral Sequence**

Let $\pi : E \to B$ be a homologically simple fibration. The transgression maps in the homological and cohomological Serre spectral sequence of $\pi$ coincide respectively with the following partially defined maps:

$$H_n(B) = H_n(B, *) \xrightarrow{(\pi_*)^{-1}} H_n(E, F) \xrightarrow{\partial_*} H_{n-1}(F)$$

$$H^{n-1}(F) \xrightarrow{\delta^*} H^n(E, F) \xrightarrow{(\pi^*)^{-1}} H^n(B, *) = H^n(B).$$

Here $\partial_*, \delta^*$ come from the long exact sequences of a pair, and $\pi_*, \pi^*$ may not be invertible, hence these are only partially defined.

Proof. We prove just the homological version, the other being analogous. By restricting to a path component of the base, we may assume $B$ has only a single 0-cell $*$. The group $E^n_{n,0}$ is

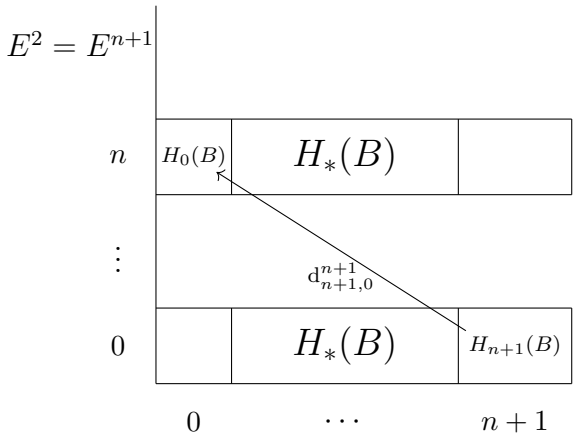
$$\frac{F_n C_n(E) \cap \mathrm{d}^{-1} F_0 C_{n-1}(E)}{F_{n-1} C_n(E) \cap \mathrm{d}^{-1} F_0 C_{n-1}(E) + F_n C_n(E) \cap \mathrm{d} F_{2n-1} C_{n+1}}.$$

Representatives of this group consist of chains $c \in C_n(\pi^{-1}(B_n))$ with boundaries in $C_{n-1}(\pi^{-1}(*)) = C_{m-1}(F)$. I.e. by relative cycles of the pair $(\pi^{-1}(B_n), F)$. We identify $E^n_{n,0}$ with a subgroup of $H_n(B)$ by sending $c$ to $\pi_*(c)$. The differential $\mathrm{d}^n_{n,0}$ acts by sending $c$ to $\partial c \in C_{n-1}(F)$. This completes the proof. □

*Two Exact Sequences.* We discuss now a couple of exact sequences that can be derived from the Serre spectral sequence. They both deal with fibrations involving spheres.

The first is the Gysin sequence for sphere bundles. Suppose $\pi : E \to B$ is a homologically simple fibration with spherical fibre $S^n$. If this is a smooth sphere bundle, the homological simplicity assumption is the same as asking that the bundle be orientable.

We know the homology of $S^n$ and so we deduce that the $E^2$ page of the spectral sequence is two copies of the homology of $B$ in rows 0 and $n$. The only non-trivial differential can occur on the page $E^{n+1}$ when we get a family of maps $\mathrm{d}^{n+1}_{r,0} : E^{n+1}_{r,0} \to E^{n+1}_{r-n-1,n}$ for each $r \geq n+1$. See the figure below.




We see these differentials give rise to maps $H_r(B) \to H_{r-n-1}(B)$ for each $r$.

The cohomology of this page is $E^{n+2} = E^\infty$ We conclude that each diagonal $E^\infty_{p+q=m}$ consists of at most two non-zero terms, namely $E^\infty_{r,0} = \ker(\mathrm{d}^{n+1}_{r,0})$ and $E^\infty_{r-n,n} = \mathrm{coker}(\mathrm{d}^{n+1}_{r+1,0})$. We know that because it lies in the last non-trivial row, $E^\infty_{r-n,n}$ is a subgroup of $H_r(E)$, while because it lies in the first non-trivial row, $E^\infty_{r,0}$ is a quotient of $H_r(E)$ (see (5.1)). So we have short exact sequences,

$$0 \to \mathrm{coker}(\mathrm{d}^{n+1}_{r+1,0}) \to H_r(E) \to \ker(\mathrm{d}^{n+1}_{r,0}) \to 0.$$

This is equivalent to the exact sequence,

$$H_{r+1}(B) \xrightarrow{\mathrm{d}^{n+1}_{r+1,0}} H_{r-n}(B) \to H_r(E) \to H_r(B) \xrightarrow{\mathrm{d}^{n+1}_{r,0}} H_{r-n-1}(B).$$

Because the differential squares to zero, these can be merged to a long exact sequence,

$$\cdots \xrightarrow{\pi_*} H_{r+1}(B) \xrightarrow{\mathrm{d}} H_{r-n}(B) \xrightarrow{\pi^!} H_r(E) \xrightarrow{\pi_*} H_r(B) \to \cdots .$$

This is called the *Gysin sequence*. The fact the map $H_r(E) \to H_r(B)$ coincides with $\pi_*$ follows directly from our earlier discussion of edge morphisms of spectral sequences. The map $\pi^!$ maps a homology class in $B$ to its inverse image under $\pi$. If we are dealing with a smooth bundle, then $\pi^!$ is the "wrong-way" map on homology given by the pullback on cohomology composed on both sides with the Poincaré isomorphism:

$$\pi^! : H_{r-n}(B) \xrightarrow{PD} H^{b-r+n}(B) \xrightarrow{\pi^*} H^{b-r+n}(E) \xrightarrow{PD} H_r(E),$$

where $\dim(B) = b$. In this context, this map $\pi^!$ is usually called the *Gysin homomorphism*. The map d, which came from the differential in the spectral sequence, can be interpreted as the cap product with the Euler class of our sphere bundle $e(E) \in H^{n+1}(B)$.

This last fact is a consequence of the following proposition.

PROPOSITION 5.8. *Let $F \hookrightarrow E \xrightarrow{\pi} B$ be a homologically simple fibration for which $F$ is $(n-1)$-connected. Define a tautological cohomology class $c_F \in H^n(F; \pi_n(F))$ so that $c_F(h(\alpha)) = \alpha$ for any $\alpha \in \pi_n(F)$, where $h$ is the Hurewicz homomorphism. Then under the transgression $\tau$ in the Serre spectral sequence of $\pi$, $c_F$ is sent to the first obstruction of our fibration:*

$$\tau(c_F) = C(E) \in H^{n+1}(B; \pi_n(F)).$$

PROOF. By Theorem 5.7, we want to show that $\delta^*(c_F) = \pi^*(C(E)) \in H^{n+1}(E, F; \pi_n(F))$.

Since $C(E)$ is the first obstruction to lifting a section of $E$, $\sigma : B \to E$ to the $n+1$ skeleton of $B$, we conclude $\pi^*(C(E))$ is the obstruction to lifting a fibrewise map $\widetilde{\sigma} : E_n \to E$ to the $(n+1)$-skeleton of $E$. Equivalently, this is the first obstruction to lifting the identity map $F \to F$ to a map $E \to F$ (refer to example 4.43).

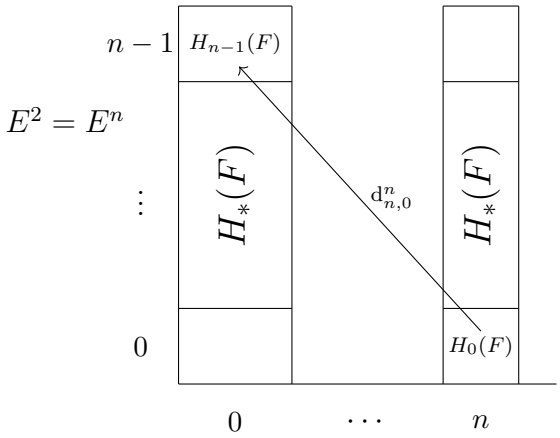
Now note $\delta^*(c_F)$ evaluated on some relative $(n+1)$-cell of $(E, F)$ is $c_F$ evaluated on its boundary, an $n$-cell of $F$. Since $c_F$ is tautological, this is just the homotopy class of the boundary. But this precisely describes $\delta^*(c_F)$ as the first obstruction to lifting the identity map $F \to F$ to a map $E \to F$ just like $\pi^*(C(E))$. □

We have an analogous Gysin sequence in cohomology:

$$\cdots \xrightarrow{\mathrm{d}} H^r(B) \xrightarrow{\pi^*} H^r(E) \xrightarrow{\pi_!} H^{r-n}(B) \xrightarrow{\mathrm{d}} H^{r+1}(B) \to \cdots .$$

Now d is the cup product with $e(E)$ and in the smooth case $\pi_!$ is the wrong-way map in cohomology.

Now we consider the Wang sequence. Suppose $\pi : E \to S^n$ is a fibration over a sphere with fibre $F$. If $n = 1$, we should assume homological simplicity, otherwise it is automatic. The $E^2$ page of the corresponding specral sequence consists of two copies of the homology of $F$ on the zeroth and $n$ columns. The only non-trivial differential can occur on the page $E^n$ when we get a family of maps $\mathrm{d}^n_{n,r-n+1} : E^n_{n,r-n+1} \to E^n_{0,r}$ for each $r \geq n-1$. See the figure below.




We see the differential gives rise to maps $H_r(B) \to H_{r-n+1}(B)$ for each $r$. We have $E^{n+1} = E^\infty$ and as with the Gysin sequence we obtain short exact sequences,

$$0 \to \mathrm{coker}(\mathrm{d}^n_{n,r-n+}) \to H_r(E) \to \ker(\mathrm{d}^n_{n,r-n}) \to 0,$$

which patch together into a long exact sequence,

$$\cdots \xrightarrow{i_!} H_{r-n+1}(F) \xrightarrow{\mathrm{d}} H_r(F) \xrightarrow{i_*} H_r(E) \xrightarrow{i_!} H_{r-n}(F) \to \cdots .$$

This is called the *Wang sequence*. The map $i_*$ is induced from inclusion, as we know from our discussion on morphisms. The map $i_!$ is given by restricting a cycle to a fibre; in the smooth situation this is the wrong way map corresponding to $i$. The map d coming from the differential can be dscribed as follows. Consider $h : D^n \to D^n/\partial D^n = S^n$ and let $h^*E$ be the pullback of our fibration; being over $D^n$ this is necessarily trivial. Thus the map $h^*E \to E$ gives a map $D^n \times F \to E$ which restricts to a map $\widetilde{h} : S^{n-1} \times F \to F$. The

map d is given by $\mathrm{d}(\alpha) = \widetilde{h}([S^{n-1} \times \alpha)$. A proof is left to the reader. We also have an analogous cohomological Wang sequence,

$$\cdots \xrightarrow{\mathrm{d}} H^{r-n}(F) \xrightarrow{i^!} H^r(E) \xrightarrow{i^*} H^r(F) \xrightarrow{\mathrm{d}} H^{r-n+1}(F) \to \cdots .$$

*Multiplication.* As for cohomology vs. homology, the cohomological Serre spectral sequence has a major advantage over the homological version in that it has the structure of a graded ring.

We can use the multiplication to understand the cup product structure of the total space of a fibration, or the cup product on the fibre and base. Additionally, the product structure sometimes tells us a priori what the differential maps of the spectral sequence need to be without having to work through their quite involved definition.

We begin in full generality. Suppose $F = F^0 \supset F^1 \supset \cdots$ is a filtered differential algebra equipped with a multiplication $F^p \times F^q \to F^{p+q}$ compatible with the differential in the sense that there is a Leibniz rule,

$$\delta(ab) = (\delta a)b + (-1)^{|a|}a(\delta b).$$

We are assuming here that $F$ is $\mathbb{Z}_2$ or $\mathbb{Z}$-graded. If not then just we can set $|a| = 0$ for all $a$.

This multiplication descends to a map $F \otimes F \to F$. Note that this tensor product has an induced filtration and differential given by,

$$(F \otimes F)^n = \sum_{p+q \leq n} F^p \otimes F^q \quad \text{and} \quad \delta(a \otimes b) = \delta a \otimes b + (-1)^{|a|} a \otimes \delta b.$$

This makes $F \otimes F \to F$ a map of filtered differential algebras and so we get a corresponding map of spectral sequences. This begins with the induced map on the associated graded groups $E_0 \otimes E_0 \to E_0$. By passing to the cohomology, we get a map,

$$H^*(E_0) \otimes H^*(E_0) \to H(E_0 \otimes E_0) \to H(E_0).$$

Here the first map is the embedding from the Künneth formula and the second is the induced map on $E_1$ pages from $F \otimes F \to F$. Thus, we obtain a map $E_1 \otimes E_1 \to E_1$. In this exact way, we pass to all later pages. If $F$ possesses a grading, then this whole structure is compatible with the bigrading of the spectral sequence, as we expand on in the following central result.

**Theorem 5.9: Multiplication in the Serre Spectral Sequence**

Let $\pi : E \to B$ be a homologically simple fibration with fibre $F$. Then the cohomological Serre spectral sequence of $\pi$ obtains a multiplication on each page beginning at $E_2$ possessing the following properties:

(i) The multiplication is compatible with the bigrading in the sense that if $\alpha \in E_r^{p,q}$ and $\beta \in E_r^{s,t}$, then $\alpha\beta \in E_r^{p+s,q+t}$.

(ii) The differentials satisfy a bigraded Leibniz rule: if $\alpha \in E_r^{pq}$ and $\beta \in E_r^{st}$ then,

$$\mathrm{d}_r^{p+s,q+t}(\alpha\beta) = (\mathrm{d}_r^{p,q}\alpha)\beta + (-1)^{p+q}\alpha(\mathrm{d}_r^{s,t}\beta).$$

(iii) The multiplication on the page $E_{r+1}$ is induced from the multiplication on $E_r$, so that if $\alpha, \beta \in E_{r+1}$ are represented by $\mathrm{d}_r$-cycles $a, b$, then $\alpha\beta$ is represented by $ab$.

(iv) On the $E_2$ page, the multiplication is just the multiplication in $H^*(B)$ with coefficients in the ring $H^*(F)$.

(v) The multiplication on $E_\infty$ is inherited from the multiplication on $H^*(E)$ in the following sense. If $a \in F^pH^m(E)$ and $b \in F^pH^n(E)$ represent $\alpha \in E_\infty^{p,m-p}$ and $\beta \in E_\infty^{q,n-q}$, then $ab \in F^{p+q}H^{m+n}(E)$ represents $\alpha\beta \in E_\infty^{p+q,m+n-p-q}$.

Proof. We need only construct the multiplication on $E_2$. The facts (i)–(iii) are just general consequences of our set up described above.

By CW approximation, we can restrict to the case where $B$ is cellular. Consider the diagonal embedding of fibrations factored through the pullback bundle.

$$\begin{array}{ccccc} E & \xrightarrow{\quad\Delta_E\quad} & \Delta_E^*E \times E & \longrightarrow & E \times E \\ \downarrow{\scriptstyle F} & & \downarrow{\scriptstyle F\times F} & & \downarrow{\scriptstyle F\times F} \\ B & \xrightarrow{\mathrm{id}} & B & \xrightarrow{\Delta_B} & B \times B \end{array}$$

By Proposition 5.5, we may CW approximate $\Delta_B$ by $\Delta_B^{CW}$ and obtain a morphism of Serre spectral sequences beginning at the page $E_2$. Note the cup product is usually defined as the pullback under $\Delta_B$ of the cross product of cohomology classes. We may equally well use $\Delta_B^{CW}$ since they are homotopic and hence induce the same maps on cohomology.

From the proof of Proposition 5.5, the induced map on $E_2$ factors as in the commutative diagram above into maps,

$$H^*(B \times B; H^*(F \times F)) \to H^*(B; H^*(F \times F)) \tag{5.6}$$

$$H^*(B; H^*(F \times F)) \to H^*(B; H^*(F)). \tag{5.7}$$

We also have a morphism given by the Künneth formula, which on the $E_2$ page looks like,

$$H^*(B; H^*(F)) \otimes H^*(B; H^*(F)) \to H^*(B \times B; H^*(F \times F)) \tag{5.8}$$

The composition of (5.6) with (5.8) is the cup product on $H^*(B)$ and the cross product on $H^*(F)$. Composing with (5.7), the total construction is the cup product on $H^*(B)$ and $H^*(F)$. This is our desired map of spectral sequences.

From the usual properties of morphisms of spectral sequences, the map on $E_\infty$ is the associated graded of the cup product on $H^*(E)$, proving (v). □

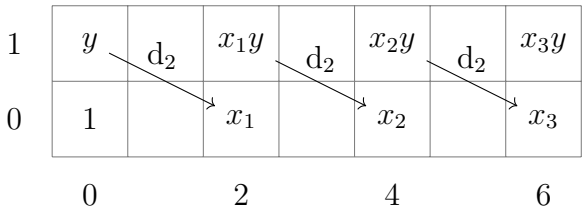
Remark 5.10. The induced multiplication on $E_\infty = \mathrm{Gr}(H^*(E))$ may be much less rich than that on $H^*(E)$. For example, even if the multiplication on $E_\infty$ is trivial, we only know that $F^pH^*(E) \cdot F^qH^*(E) \subset F^{p+q+1}H^*(E)$. △

**5.4. Computations with the Serre Spectral Sequence.** We will now take a look at several applications of computing homology (and homotopy) from the Serre spectral sequence. The sequence is useful for computing the stable homotopy groups of spheres and the homology of Eilenberg-MacLane spaces and we will briefly touch on these subjects. Our main interest will be in finding the homology rings of some Lie groups and homogeneous spaces; the Serre spectral sequence is very useful for this purpose since Lie groups and homogeneous spaces come equipped with many geometrically motivated fibrations, some of which we touched on earlier. We will also give a proof of the Thom isomorphism theorem.

*Some Miscellaneous Examples.* We begin with a few examples to become acquainted with explicit calculations using spectral sequences.

To ease into things, let us re-compute the ring structure of $H^*(\mathbb{C}P^\infty)$ which we previously found from intersection theory. We will use the cohomological Serre spectral sequence associated to the infinite complex Hopf fibration $S^1 \hookrightarrow * = S^\infty \to \mathbb{C}P^\infty$, which is homologically simple. We know the cohomology ring of $S^1$ is $H^*(S^1) = \mathbb{Z}[y]/(y^2)$ where $|y| = 1$. From cellular homology we know that $H^i(\mathbb{C}P^\infty)$ is zero in odd degrees and $\mathbb{Z}$ in even degrees. Let $x_i$ denote a generator of $H^{2i}(\mathbb{C}P^\infty)$. For ease of computation, it is customary when drawing a page of a spectral sequence to write the generator of each square in place of the group it generates. If a square is the trivial group, we

usually leave it empty. Thus our $E_2$ page will look like the following.



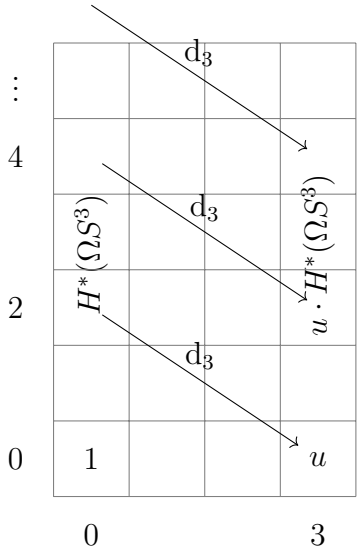

This is the last page with non-trivial differentials and so $E_3 = E_\infty$. Since $S^\infty$ is contractible, the $E_\infty$ page should be empty apart from $E_\infty^{0,0}$. Hence all the diagonal maps $\mathrm{d}_2 : E_2^{2k,1} \to E_2^{2k+2,0}$ must be isomorphisms. Without loss of generality, we can inductively modify the generators $x_i$ up to a sign so that $\mathrm{d}_2(y) = x$ and $\mathrm{d}_2(x_i y) = x_{i+1}$ for $i \geq 1$. Since the differential leaving $x_i$ is trivial, the graded derivation structure implies that,

$$\mathrm{d}_2(x_i y) = \cancelto{0}{\mathrm{d}_2(x_i)y} + x_i \mathrm{d}_2 y = x_i x_1.$$

We conclude, $x_2 = x_1^2$, $x_3 = x_2 x_1 = x_1^3$, and so on. So that $x_i = x_1^i$. We conclude all the powers of $x_1$ are non-zero and that the $2n$th cohomology group of $\mathbb{C}P^\infty$ is generated by $x_1^n$. Thus we rediscover

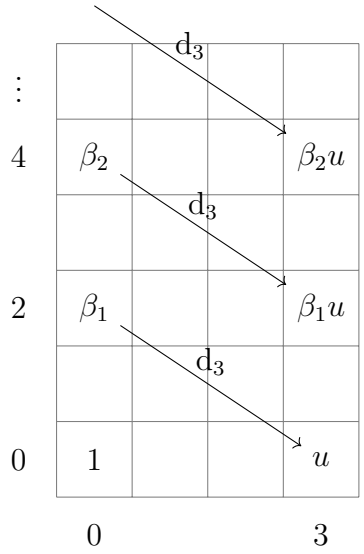
$$H^*(\mathbb{C}P^\infty) = \mathbb{Z}[x], \quad \text{where} \quad |x| = 2.$$

Let us compute the cohomology of the loop space $\Omega S^3$ of $S^3$. Note there is a path space fibration $\Omega S^3 \hookrightarrow * = ES^3 \to S^3$. This is homologically simple and so we can use the Serre spectral sequence. We know the cohomology ring of $S^3$ is $H^*(S^3) = \mathbb{Z}[u]/(u^2)$, where $|u| = 3$. The $E_2 = E_3$ page of the cohomological spectral sequence of our fibration looks like the following.



This is the last page on which non-trivial differentials occur. Since the spectral sequence must converge to the trivial homology of $ES^3$, all the maps $\mathrm{d}_3 : E_3^{0,n} \to E_3^{3,n-2}$ for $n \geq 1$ must be isomorphisms. We conclude inductively that $H^n(\Omega S^3) = \mathbb{Z}$ if $n$ is even and 0 if $n$ is odd. This could have been deduced from (really it is the same computation as) the Wang sequence.

We can now look at the $E_3$ page again to deduce the multiplicative structure Let $\beta_n$ denote a generator for $H^{2n}(\Omega S^3)$. The $E_3$ page looks as follows.




Inductively, we may choose $\beta_i$ so that $\mathrm{d}_2(\beta_1) = u$ and $\mathrm{d}_3(\beta_i) = \beta_{i-1}u$. for $i > 1$. We would like to determine the relationship between the products of the generators. Suppose that $\beta_k \cdot \beta_\ell = \lambda_{k,\ell}\beta_{k+\ell}$ for some coefficient $\lambda_{k,\ell} \in \mathbb{Z}$. We determine from the derivation structure,

$$\mathrm{d}_3(\beta_k\beta_\ell) = \mathrm{d}_3(\beta_k)\beta_\ell + \beta_k\mathrm{d}_3(\beta_\ell) = \beta_{k-1}\beta_\ell u + \beta_k\beta_{l-1}u = (\lambda_{k-1,\ell} + \lambda_{k,\ell-1})\beta_{k+\ell-1}u.$$

On the other hand,

$$\mathrm{d}_3(\beta_k\beta_\ell) = \mathrm{d}_3(\lambda_{k,\ell}\beta_{k+\ell}) = \lambda_{k,\ell}\beta_{k+\ell-1}u.$$

We conclude that the coefficients satisfy the coupled integer equations,

$$\lambda_{k,\ell} = \lambda_{k-1,\ell} + \lambda_{k,\ell-1} \quad \text{for} \quad k, \ell \in \mathbb{N}.$$

But note these are exactly the relations for $\lambda_{k,\ell}$ the $k$th entry of the $(k+\ell)$th row of Pascal's triangle. So our coefficients are binomial coefficients:

$$\beta_k\beta_\ell = \binom{k+\ell}{k}\beta_{k+\ell}.$$

We conclude the ring structure is,

$$H^*(\Omega S^3) = \mathbb{Z}[x^k/k! : k = 1, 2, \ldots] \quad \text{where} \quad |x| = 2.$$

EXERCISE 5.11. Show that if $n$ is odd, then $H^*(\Omega S^n)$ has the same ring structure but where $|x| = n-1$. Find the ring structure of $H^*(\Omega S^n)$ for $n$ even. ▲

*Homotopy Groups of Spheres.* As a first application to stable homotopy theory, we can finally compute the first stable homotopy group of the sphere, namely $\pi_4(S^3)$. Consider the killing space $S^3|_4$. We have by Hurewicz,

$$\pi_4(S^3) = \pi_4(S^3|_4) = H_4(S^3|_4).$$

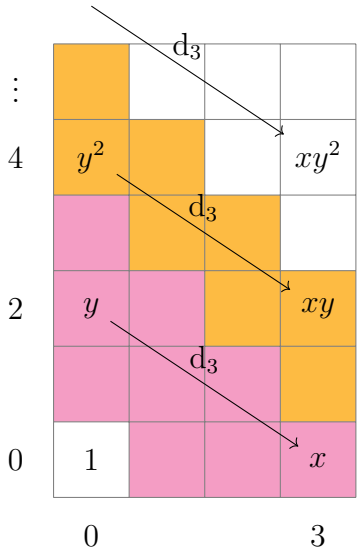
Recall from our construction of killing spaces, $S^3|_4$ is the total space of a Serre fibration,

$$K(\mathbb{Z}, 2) \hookrightarrow S^3|_4 \to S^3.$$

This fibration is homologically simple and so we can study the associated Serre spectral sequence. We know the cohomology of $S^3$ and $K(\mathbb{Z}, 2) = \mathbb{C}P^\infty$ as rings:

$$H^*(S^3) = \mathbb{Z}[x]/(x^2) \quad \text{and} \quad H^*(K(\mathbb{Z}, 2)). = \mathbb{Z}[y],$$

where $|x| = 3$ and $|y| = 2$. The $E_2 = E_3$ page of the spectral sequence looks like the following.



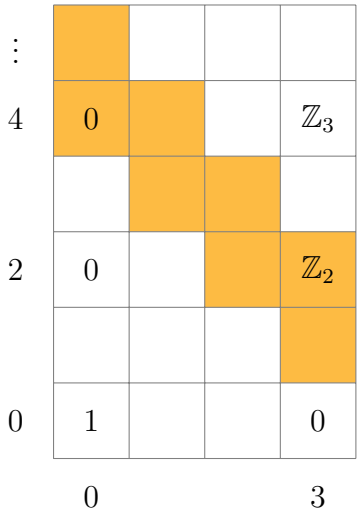

This will be the last page on which non-trivial differentials occur so that $E_4 = E_\infty$. We know by the Hurewicz theorem that $H^i(S^3|_4) = \pi_i(S^3|_4) = 0$ for $i = 1, 2$. Hence all contents of the pink squares must be killed by the $E_\infty$ page. By the universal coefficient theorem, $H_4(S^3|_4)$ is the free part of $H^4(S^3|_4)$ direct sum the torsion part of $H^5(S^3|_4)$, so to compute our homotopy group we need to understand the yellow squares of the spectral sequence on the $E_\infty$ page.

In order that the contents of the pink squares are killed, we need the transgressive map $\mathrm{d}_3^{0,2}$ to be an isomorphism. Without loss of generality, changing

the sign of $x$ if necessary, we can set $\mathrm{d}_3 y = x$. Applying this to the next map $\mathrm{d}_3^{0,4}$ and using the graded derivation structure,

$$\mathrm{d}_3(y^2) = 2y\,\mathrm{d}_3 y = 2xy.$$

We conclude the $E_4 = E_\infty$ page looks like,




Since each diagonal has at most one non-trivial entry we have,

$$H^4(S^3|_4) = 0 \quad \text{and} \quad H^5(S^3|_4) = \mathbb{Z}_2.$$

Thus, $H_4(S^3|_4) = Z_2$. In fact we easily compute all higher homotopy groups:

$$H_n(S^3|_4) = \begin{cases} \mathbb{Z}_m & n = 2m \\ 0 & \text{else.} \end{cases}$$

We conclude $\pi_4(S^3) = H_4(S^3|_4) = \mathbb{Z}_2$. And so the first stable homotopy group of the sphere is $\mathbb{Z}_2$. By the Freudenthal suspension theorem, we know this homotopy class of map $S^4 \to S^3$ is given by suspending the Hopf fibration $S^3 \to S^2$. The preceding computation demonstrates that if we precompose the Hopf fibration with a degree two map $S^3 \to S^3$ and then suspend (or suspend the Hopf fibration and precompose with a degree two map $S^4 \to S^4$) the resulting map is null-homotopic.

This looks like a promising method for computing homotopy groups of spheres, and indeed it is. But further computations will not be as easy because understanding the cohomology of the capping spaces becomes harder as its relation to the sphere in the associated Whitehead tower becomes more complicated.

We now give another application to homotopy theory by proving Serre's finiteness theorem (Proposition 4.24). We begin with a lemma.

LEMMA 5.12. *Let $X$ be a simply connected space with finitely generated homology which is a rational odd homology sphere. I.e. $H^*(X;\mathbb{Q}) = H^*(S^m;\mathbb{Q})$ for $m$ odd. Then $\pi_q(X)$ is finite for $q \neq m$ and has rank one if $q = m$.*

PROOF. We have $H^m(X;\mathbb{Q}) = \mathbb{Q}$ and so there is an element of infinite order $\alpha \in H^m(X;\mathbb{Z})$. Let $F_m \in H^m(K(\mathbb{Z},m);\mathbb{Z})$ be the tautological class. By obstruction theory, we can construct $f : X \to K(\mathbb{Z},m)$ so that $f^*(F_m) = \alpha$.

EXERCISE 5.13. Inductively apply the Serre spectral sequence to the fibrations $K(\mathbb{Z},n-1) \hookrightarrow * \to K(\mathbb{Z},n)$ to conclude the rational homology of these Eilenberg–MacLane spaces is,

$$H^*(K(\mathbb{Z},n);\mathbb{Q}) = \begin{cases} \Lambda_{\mathbb{Q}}[x] & n \text{ odd} \\ \mathbb{Q}[x] & n \text{ even} \end{cases} \quad \text{where} \quad |x| = n.$$

Apply the spectral sequence of the more general fibration $K(G,n-1) \hookrightarrow * \to K(G,n)$, to show the homology $H^q(K(G,n)$ is finitely generated for any $q$ and any $G$ finitely generated abelian. ▲

Having proved this exercise, we conclude $f^*$ is an isomorphism on rational homology.

We may homotope $f$ to be a fibration with homotopy fibre $F$ and consider the corresponding Serre spectral sequence. Recall from our discussion of edge morphisms that the induced pullback projection map $f^* : H^n(K(\mathbb{Z},m);\mathbb{Q}) \to H^*(X,\mathbb{Q})$ has to factor through the quotient map $H^n(K(\mathbb{Z},m;\mathbb{Q})) \to E_\infty^{*,0}$ onto the first row of the infinity page. Because $f^*$ is an isomorphism, this quotient must be an isomorphism and so the zeroth row of the spectral sequence can never participate in any differentials.

On the other hand, because $H^n(K(\mathbb{Z},m);\mathbb{Q}) \cong H^*(X,\mathbb{Q})$, the only contents of the $E_\infty$ page of the spectral sequence can be in the first row, since additional non-zero squares would complicate the homology of the total space. Suppose $H^m(F;\mathbb{Q})$ is the non-trivial rational homology group of $F$ with minimal $m > 0$. Then as we just noted, $E_2^{0,m} = H^m(F;\mathbb{Q})$ must be killed on some page of the sequence. But the only row below $m$ containing non-zero squares is the zeroth row. So $E_2^{0,m}$ can only be killed by a transgression. But this would force a non-trivial map to hit the zeroth row, which we said could not happen. We obtain a contradiction and so the homology $H^m(F;\mathbb{Q}) = 0$ for all $m > 0$.

Hence by the universal coefficient theorem, all the homology groups of $F$ are torsion. Because the integral homology of $X$ and $K(\mathbb{Z},m)$ are finitely generated (by the above exercise) the same must be true of $F$ (using the integral

Serre spectral sequence). Hence $F$ has finite homology groups. Now inductively studying the fibrations,

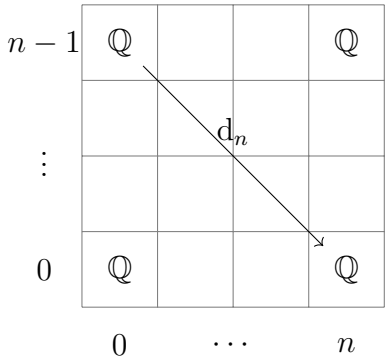
$$K(H_n(F|_n); n-1) \hookrightarrow F|_{n+1} \to F|_n,$$

we conclude from the corresponding spectral sequences that the killing spaces $F|_n$ all have finite integral homology groups. Hence $\pi_n(F) = H_n(F|_n)$ is finite for each $n$.

Now we can look at the homotopy long exact sequence of the fibration $F \hookrightarrow X \to K(\mathbb{Z}, m)$. We conclude the maps $f_* : \pi_n(X) \to \pi_n(K(\mathbb{Z}, m))$ all have finite kernel and cokernel. Hence using the homology of $K(\mathbb{Z}, m)$ when $m$ is odd, we obtain the result. □

PROOF. (of Serre's finiteness theorem) We wish to show $\pi_q(S^n)$ is finite except for $q = n$ or in the case $\pi_{4m-1}(S^{2m})$ where the rank is one. If $n$ is odd, then we are done directly by the above lemma.

Let $n = 2m$ be even and consider the fibration $K(\mathbb{Z}, n-1) \hookrightarrow S^n|_{n+1} \to S^n$. The rational cohomological spectral sequence will look like the following with the first (and last) non-trivial differential on the $n$th page.




Because $H^{n-1}(S^n|_{n+1}) = H^n(S^n|_{n+1}) = 0$, the illustrated map $\mathrm{d}_n^{0,n-1}$ has to be an isomorphism. Hence, $H^*(S^n|_{n+1}; \mathbb{Q}) \cong H^*(S^{2n-1}; \mathbb{Q})$.

We conclude from the previous lemma that for $q > 2m$, $\pi_q(S^{2m}|_{2m+1}) = \pi_q(S^{2m})$ is finite unless $q = 4m-1$, in which case it has rank one. □

In fact we can give a nice construction of a free generator in $\pi_{4m-1}(S^{2m})$. There is a graded Lie algebra structure on the direct sum of the homotopy groups of a space called the *Whitehead bracket*,

$$[\cdot, \cdot] : \pi_k(X) \times \pi_\ell(X) \to \pi_{k+\ell-1}(X).$$

This defines a graded Lie algebra in the sense it is bilinear, supersymmetric, and satisfies a super/graded Jacobi identity.

To construct this bracket, note $S^k \times S^\ell$ is formed from $S^k \vee S^\ell$ by attaching a $(k+\ell)$-cell. Hence there is a cell attaching map $\psi : S^{k+\ell-1} \to S^k \vee S^\ell$. Then the Whitehead bracket is defined by,

$$[f, g] = \psi^*(f \vee g).$$

PROPOSITION 5.14. *Let* $\mathrm{id}_{2n} \in \pi_{2n}(S^{2n})$ *be the homotopy class of the identity map. Then* $[\mathrm{id}_{2n}, \mathrm{id}_{2n}] \in \pi_{4n-1}(S^{2n})$ *has infinite order.*

PROOF. (Sketch) Given $f : S^{4n-1} \to S^{2n}$, form its cone $C_f = S^{2n} \cup_f D^{4n}$ which up to homotopy equivalence depends only on the class $[f] \in \pi_{4n-1}(S^{2n})$. The cohomology ring of $C_f$ is easily seen to be,

$$H^*(C_f) = \mathbb{Z}[\alpha, \beta]/(\alpha\beta, \beta^2, \alpha^2 - h(f)\beta), \quad \text{where} \quad |\alpha| = 2n, |\beta| = 4n.$$

Here $h(f)$ is some integer depending on $f$, called the *Hopf invariant*.

It can be shown that $h : \pi_{4n-1}(S^{2n}) \to \mathbb{Z}$ is a group homomorphism. To see this, one studies the space $Y_{f,g} = S^{2n} \cup_{f \vee g} (D^{4n} \vee D^{4n})$ and the effect on cohomology of the natural maps $C_f, C_g, C_{f+g} \to Y_{f,g}$.

One can also show $h([\mathrm{id}_{2n}, \mathrm{id}_{2n}]) = 2$. For this, one can form $C_{[\mathrm{id}_{2n}, \mathrm{id}_{2n}]}$ as the quotient of $S^{2n} \times S^{2n}$ by identifying points $(x, x_0)$ with $(x_0, x)$, where $x_0$ is the basepoint. An explicit study of the cohomology of this space gives the Hopf invariant.

We conclude that non-zero multiples of $[\mathrm{id}_{2n}, \mathrm{id}_{2n}]$ in $\pi_{4n-1}(S^{2n})$ all have non-zero Hopf invariants and hence can never be null-homotopic. □

Lastly for this section, let us compute the second stable homotopy group of the sphere. Consider the identification $\Sigma S^2 \to S^3$. By the loop-suspension adjunction, we obtain a map $f : S^2 \to \Omega S^3$. We may homotope $f$ to a fibration with homotopy fibre $F$. From the homotopy long exact sequence of this fibration, $F$ is 2-connected, and so by Hurewicz $\pi_3(F) = H_3(F)$. We computed earlier the ring structure of $H^*(\Omega S^3)$. We deduce from this that $H_4(\Omega S^3) = \mathbb{Z}$. Considering the spectral sequence of our fibration, since $H_3(S^2) = H_4(S^2) = 0$, the trangressive map on the $E^4$ page, $H_4(\Omega S^3) \to H_3(F)$ must be an isomorphism (it is the first and last non-trivial map to hit the squares $E_{4,0}$ and $E_{0,3}$). Hence, $\mathbb{Z} = H_3(F) = \pi_2(F)$.

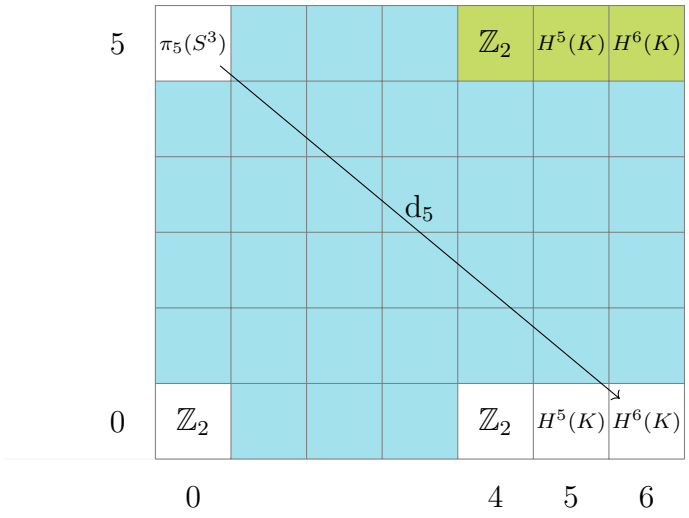
We consider the following segment of the homotopy long exact sequence of our fibration.

$$\begin{array}{ccccccccccccc}
\cdots & \longrightarrow & \pi_4(S^2) & \xrightarrow{f_*} & \pi_4(\Omega S^3) & \longrightarrow & \pi_3(F) & \longrightarrow & \pi_3(S^2) & \xrightarrow{f_*} & \pi_3(\Omega S^3) & \longrightarrow & \cdots \\
 & & \| & & \| & & \| & & \| & & \| & & \\
\cdots & \longrightarrow & \mathbb{Z}_2 & \longrightarrow & \pi_5(S^3) & \longrightarrow & \mathbb{Z} & \longrightarrow & \mathbb{Z} & \longrightarrow & \mathbb{Z}_2 & \longrightarrow & \cdots
\end{array}$$

We have used $\pi_4(S^2) = \pi_4(S^3) = \mathbb{Z}_2$ from the long exact sequence of the Hopf fibration and the fact $\pi_n(\Omega S^3) = \pi_{n+1}(S^3)$. The map $\mathbb{Z} \to \mathbb{Z}_2$ is either the quotient map or trivial. Hence the map $\mathbb{Z} \to \mathbb{Z}$ is multiplication by $\pm 2$ or $\pm 1$, in particular it is injective. Thus the map $\pi_5(S^3) \to \mathbb{Z}$ is trivial and so $\mathbb{Z}_2 \to \pi_5(S^3)$ is a surjection. We now want to show it is non-trivial and hence equals $\mathbb{Z}_2$.

Now consider the fibration $K(\mathbb{Z}, 2) \to S^3|_4 \to S^3$. We computed $\pi_4(S^3|_4) = \mathbb{Z}_2$. So construct a map $g : S^3|_4 \to K(\mathbb{Z}_2, 4)$ inducing an isomorphism on $\pi_4$ (this is clearly doable by obstruction theory). We can homotope $g$ to a fibration with homotopy fibre $Y$. From this fibration, we conclude that $Y$ is 4-connected and so by Hurewicz and the fibration exact sequence we have $H^5(Y) = \pi_5(S^3)$. Since $\pi_5(S^3)$ is 0 or $\mathbb{Z}_2$, we may reduce mod 2: $H^5(Y; \mathbb{Z}_2) = \pi_5(S^3)$.

Now consider the cohomological spectral sequence of the fibration $Y \hookrightarrow S^3|_4 \to K(\mathbb{Z}_2, 4)$ over $\mathbb{Z}_2$. On the $E_2$ page the sequence looks like the following.




Here the blue squares are all zero and we write $H^n(K)$ to denote $H^n(K(\mathbb{Z}_2, 4); \mathbb{Z}_2)$. The green squares are not relevant to our computation. Clearly no non-trivial differentials affect the squares we are interested in up until the $E_5$ page and none will affect them again afterwards. On the $E_5$ page the only important differential is the pictured transgression,

$$\mathrm{d}_5^{0,5} : H^5(Y; \mathbb{Z}_2) = \pi_5(S^3) \to H^6(K(\mathbb{Z}_2, 4); \mathbb{Z}_2).$$

By Serre's theorem on cohomology operations, we know that $H^5(K(\mathbb{Z}_2,4);\mathbb{Z}_2)$ and $H^6(K(\mathbb{Z}_2,4);\mathbb{Z}_2)$ are generated by $Sq^1F_4$ and $Sq^2F_4$ respectively, where $F_4 \in H^4(K(\mathbb{Z}_2,4);\mathbb{Z}_2)$ is the tautological class. (We will shortly give some indication of how to prove Serre's theorem for those concerned by this gap in our argument). Hence we know the entries of our spectral sequence $E_5^{0,5} = E_5^{0,6} = \mathbb{Z}_2$. Now recall in finding the first stable homotopy group of the sphere we determined the homology of $S^3|_4$, in particular $H^5(S^3|_4;\mathbb{Z}_2) = \mathbb{Z}_2$ and $H^6(S^3|_4) = 0$. Since the above transgression $\mathrm{d}_5^{0,5}$ is the last map impacting the relevant region of our spectral sequence, we must have in order to converge to the correct cohomology of $S^3|_4$ that $E_6^{0,5} = E_6^{6,0} = 0$. Hence, the map $\mathrm{d}_5^{0,5}$ is an isomorphism. In particular,

$$\pi_5(S^3) = H^6(K(\mathbb{Z}_2,4);\mathbb{Z}_2) = \mathbb{Z}_2.$$

We are almost done. Consider the quaternionic Hopf fibration $S^3 \hookrightarrow S^7 \to S^4$. From this it is clear that $\pi_n(S^4) = \pi_{n-1}(S^3)$ for $n \leq 6$. Hence, $\pi_6(S^4) = \pi_5(S^3) = \mathbb{Z}_2$. This is the second stable homotopy group. Note the standard Hopf fibration also tells us $\pi_5(S^2) = \mathbb{Z}_2$. At this stage we know the following chart of homotopy groups depicted in Table 3. The coloured squares are the stable groups. Referring back to the Table 1, we see that $\pi_6(S^2) = \pi_6(S^3) = \mathbb{Z}_{12}$, although we haven't computed that.

| | $S^1$ | $S^2$ | $S^3$ | $S^4$ |
|---|---|---|---|---|
| $\pi_1$ | $\mathbb{Z}$ | 0 | 0 | 0 |
| $\pi_2$ | 0 | $\mathbb{Z}$ | 0 | 0 |
| $\pi_3$ | 0 | $\mathbb{Z}$ | $\mathbb{Z}$ | 0 |
| $\pi_4$ | 0 | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}$ |
| $\pi_5$ | 0 | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ |
| $\pi_6$ | 0 | ? | ? | $\mathbb{Z}_2$ |

TABLE 3. The Computed Homotopy Groups of Spheres

It is possible to go a lot further in building out this chart than we have done, but it requires progressively more work. See Chapters 3–5 of [**FF**] for a comprehensive discussion of the theory and a computation of the first thirteen stable homotopy groups. The low dimensional unstable groups can also be computed using similar methods involving Steenrod squares; see [**HatcherSS**, Thm 5.4.1] for a computation of $\pi_6(S^2)$.

*Constructing Steenrod Squares.* We now use the homology of Eilenberg–MacLane spaces to give a quick construction of Steenrod squares. Recall the Steenrod squares were stable (mod 2) cohomology operations,

$$Sq^i : H^n(X;\mathbb{Z}_2) \to H^{i+n}(X;\mathbb{Z}_2).$$

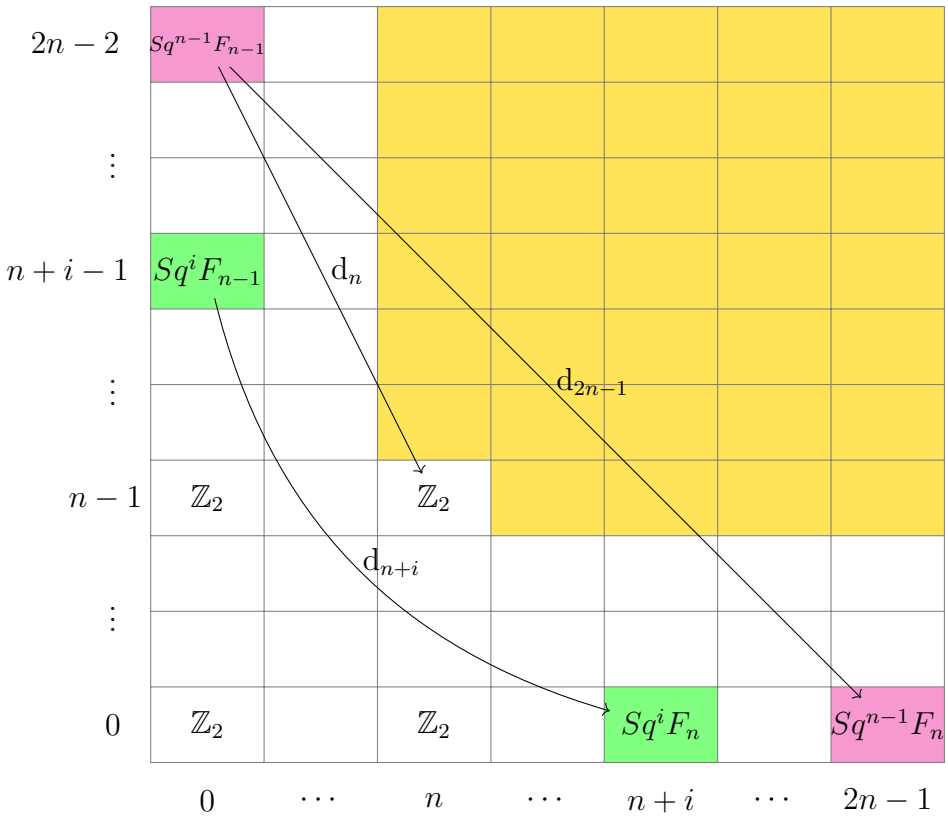
By the Yoneda lemma, these are determined by knowing

$$Sq^i(F_n) \in H^{i+n}(K(\mathbb{Z}_2, n); \mathbb{Z}_2),$$

where $F_n \in H^n(K(\mathbb{Z}_2; n); \mathbb{Z}_2)$ is the canonical cohomology class.

The axiomatic properties of Steenrod squares force $Sq^0 F_1 = F_1, Sq^1 F_1 = F_1^2$ and $Sq^i F_1 = 0$ for $i > 1$. We now work inductively, supposing we know $Sq^i F_{n-1}$ for all $i$ and determining $Sq^i F_n$.

Consider the fibration $K(\mathbb{Z}_2, n-1) \hookrightarrow * \to K(\mathbb{Z}_2, n)$ and the corresponding cohomological spectral sequence over $\mathbb{Z}_2$. Suppose $i < n-1$. We then have the following schematic picture of several pages of our sequence.




The yellow squares contain potentially non-zero entries while all the clear empty squares are zero from $E_2$ onwards. Since $i < n-1$, the only non-trivial differential impacting the squares $E^{0,n+i-1}$ and $E^{0,n+i}$ is the transgression,

$$\mathrm{d}_{n+i}^{0,n+i-1} : H^{n+i-1}(K(\mathbb{Z}_2, n-1); \mathbb{Z}_2) \to H^{n+i}(K(\mathbb{Z}_2, n); \mathbb{Z}_2).$$

Because this spectral sequence has to converge to the homology of a point, this transgression must be an isomorphism. Hence we can define,

$$Sq^i F_n := \mathrm{d}_{n+i} Sq^i F_{n-1}.$$

Now consider the case of $Sq^n$. The square $E^{0,2n-2}$ is involved in two potentially non-trivial differentials: the map $\mathrm{d}_n^{0,2n-2}$ and the map $\mathrm{d}_{2n-1}^{0,2n-2}$, both shown above. Note since the top Steenrod square is the actual square,

$$\mathrm{d}_n(Sq^{n-1}F_{n-1}) = \mathrm{d}_n(F_{n-1}^2) = 2F_{n-1}\mathrm{d}_n(F_{n-1}) = 0 \pmod 2.$$

Hence $Sq^{n-1}F_{n-1}$ persists to the $E_{2n-1}$ page and is in the domain of the transgression,

$$\mathrm{d}_{2n-1}^{0,2n-2} : \ker(\mathrm{d}_n^{0,2n-2}) \to H^{2n-1}(K(\mathbb{Z}_2, n); \mathbb{Z}_2).$$

Since this has to again be an isomorphism, we may define,

$$Sq^{n-1}F_n := \mathrm{d}_{2n-1} Sq^{n-1} F_{n-1}.$$

We know axiomatically that we must have $Sq^n F_n = F_n^2$ and $Sq^i F_n = 0$ for $i > n$. Thus we have inductively defined the Steenrod squares for all $i, n$.

The stability of these operations can be expressed as saying that the sequence of elements $Sq^r F_n$ should be obtained by applying the suspension maps,

$$H^{r+n-1}(K(\mathbb{Z}, n-1); \mathbb{Z}) \xrightarrow{\Sigma} H^{r+n}(\Sigma K(\mathbb{Z}, n-1); \mathbb{Z}) \xrightarrow{(i_n^*)^{-1}} H^{r+n}(K(\mathbb{Z}, n); \mathbb{Z}).$$

But one can show these maps agree with the transgression maps in our above spectral sequence. Hence we find the Steenrod squares are stable. Conversely, we see that to be stable, the Steenrod squares need to commute with transgressions. This proves the uniqueness of our axiomatic definition of the squares since we inductively defined them just in terms of transgressions and the other axioms. The proof of Cartan's formula proceeds from some geometric manipulations with the Eilenberg–MacLane spaces. See [**FF**, §29–30] for the details of these proofs as well as a proof of the Adem relations.

Lastly, let us sketch an argument for Serre's theorem on the fact the Steenrod squares multiplicatively generate all $\mathbb{Z}_2$ cohomology operations. The key ingredient is a theorem of Borel.

PROPOSITION 5.15 (Borel's Theorem). *Suppose $F \hookrightarrow E \to B$ is a fibration with simply connected base and weakly $\mathbb{Z}_2$ contractible total space. Further, suppose $H^*(F; \mathbb{Z}_2)$ is multiplicatively generated by elements $a_i$ which are transgressive in the Serre spectral sequence and so that the monomials $a_{i_1} a_{i_2} \cdots a_{i_k}$ with $i_1 < i_2 < \cdots < i_k$ form an additive basis for $H^*(F; \mathbb{Z}_2)$. Then $H^*(B; \mathbb{Z}_2)$ is the polynomial algebra generated by the images $b_i = \tau(a_i)$ under transgressions.*

PROOF. (Sketch) We can construct an abstract multiplicative spectral sequence whose $\widetilde{E}_2$ page is $H^*(F;\mathbb{Z}_2) \otimes \mathbb{Z}_2[b_1, b_2, \ldots]$ and compare to the Serre spectral sequence $E_2$ of the fibration. We let the differentials act trivially on each $b_i$ and act trivially on the $a_i$'s until the appropriate page $|a_i| + 1$ where $\mathrm{d}a_i = b_i$. It is routine to check this is all well defined and gives a multiplicative spectral sequence. Each generator of the $\widetilde{E}_2$ page either eventually has non-trivial image under a differential or lies in the image of a differential. Hence the $\widetilde{E}_\infty$ page of this sequence is trivial.

There is a map between our two spectral sequence by sending the corresponding elements $a_i, b_i$ to each other. Now suppose this spectral sequence $\widetilde{E}_2$ differs from the Serre spectral sequence of $F \hookrightarrow E \to F$ on the $E_2$ page, so that the multiplicative map between the first rows is not an isomorphism. Then the first row of $E_2$ either contains an element $c$ (of minimal dimension) which is not generated by the image of the transgressions or contains a non-trivial relation $f(b_i) = 0$ (of minimal dimension) between the image of the transgressions. But comparing with the spectral sequence we constructed, there is nothing on $E_2$ to kill $c$ and so $E_\infty$ can't be trivial, or there is nothing on $\widetilde{E}_2$ to kill $f(b_i)$ and so $\widetilde{E}_\infty$ can't be trivial. □

PROOF. (of Serre's theorem) We work inductively on $n$ in $K(\mathbb{Z}_2, n)$. The algebra $H^*(K(\mathbb{Z}_2, 1); \mathbb{Z}_2) = H^*(\mathbb{R}P^\infty; \mathbb{Z}_2) = \mathbb{Z}_2[e_1]$ is the polynomial algebra on $Sq^0$ as desired.

Suppose we have obtained the desired description of $H^*(K(\mathbb{Z}_2, n-1); \mathbb{Z}_2)$. We may pick a suitable indexed set of Steenrod squares $Sq^I F_{n-1}$ and their powers which multiplicatively generate and whose monomials with increasing indices additively generate $H^*(K(\mathbb{Z}_2, n-1); \mathbb{Z}_2)$ (this can be done for any polynomial algebra, for example $\mathbb{Z}[x]$ is generated by $x^{2^i}$ in the sense we desire). We can then apply Borel's theorem to the fibration $K(\mathbb{Z}_2, n-1) \to * \to K(\mathbb{Z}_2, n)$. We conclude that $H^*(K(\mathbb{Z}_2, n); \mathbb{Z}_2)$ is also a polynomial algebra generated by images of transgressions of the generators. By construction of the Steenrod squares, the image of the transgressions will be Steenrod squares $Sq^I F_n$ and their powers. Fiddling around with the indexing for a while, one can show the set of generators matches the description given in the statement of Serre's theorem. □

In addition to helping compute homotopy groups of spheres, Steenrod squares also have several other applications throughout algebraic topology. They can be used as a homotopy invariant to distinguish spaces with the same cohomology/homotopy, like lens spaces or the following example.

EXERCISE 5.16. Prove with Steenrod squares that $\Sigma\mathbb{C}P^2$ and $\Sigma(S^2 \vee S^4)$ are not homotopy equivalent. ▲

Steenrod squares also express an important property of Stiefel–Whitney classes, which we present in the next chapter.

*Thom Isomorphism Theorem.* We now prove an extremely important result for understanding the topology of vector bundles that we will need for our chapter on characteristic classes.

**Theorem 5.17: Thom Isomorphism Theorem**

Let $\pi : E \to B$ be a rank $n$ oriented vector bundle. The cup and cap product with the Thom class $t_E \in H^n(E, E \setminus B; \mathbb{Z})$ determine isomorphisms,

$$\smile t_E : H^i(B; \mathbb{Z}) \xrightarrow{\cong} H^{i+n}(E, E \setminus B; \mathbb{Z})$$
$$\frown t_E : H_{n+i}(E, E \setminus B; \mathbb{Z}) \xrightarrow{\cong} H_i(B; \mathbb{Z}).$$

In the unoriented case, the same holds over $\mathbb{Z}_2$.

PROOF. Pick a metric on our vector bundle so that we can deal with the unit disk and sphere bundles $DE$ and $SE$ which are deformation retracts of $E$ and $E \setminus B$ respectively.

It is a simple extension of the Serre spectral sequence to work for relative fibrations. So consider the fibration of pairs,

$$(D^n, S^{n-1}) \hookrightarrow (DE, SE) \to B.$$

We will ignore coefficients to treat $\mathbb{Z}$ and $\mathbb{Z}_2$ simultaneously. The important point is that in the oriented case, the coefficients coming from the homology of the fibres are untwisted and we can freely work over $\mathbb{Z}$. The $E_2$ page of the relative cohomological spectral sequence is given, using the universal coefficient theorem, by

$$E_2^{p,q} = H^p(B; H^q(D^n, S^{n-1})) \cong H^p(B) \otimes H^q(D^n, S^{n-1}) = \begin{cases} 0 & q \neq n \\ H^p(B) & q = n. \end{cases}$$

This isomorphism is given by tensoring with the fundamental class of the fibre $[D^n, S^{n-1}]$. Since there is only one non-trivial row of this spectral sequence, there are no non-trivial differentials and $E_2^{p,q} = E_\infty^{p,q}$. This spectral sequence must converge to the homology of the pair $(DE, SE)$. Hence,

$$H^p(B) \cong H^{p+n}(DE, SE).$$

Let $t_E \in H^n(DE, SE) \cong H^0(B)$ be the multiplicative unit under the isomorphism. Note that by our discussion of edge morphisms, pulling back $t_E$ by inclusion of a fibre can be seen by looking at the $E_2$ page to give the fundamental class of the fibre. Hence $t_E$ agrees with our earlier definition of the Thom class.

Note because the spectral sequence was degenerate and respects multiplication, that since the isomorphism sends $1 \in H^0(B)$ to $t_E$, it must send $\alpha \in H^0(B)$ to $t_E \smile \alpha$. So our isomorphism is as described.

For the second isomorphism, the universal coefficient theorem implies that $H_{i+n}(E, E \setminus B) \cong H_i(B)$. By duality, this isomorphism is given by $x \mapsto z$ where $\beta(z) = (\beta \smile t_E)(x)$ for all $\beta \in H^i(B)$. But,

$$(\beta \smile t_E)(x) = \beta \frown (t_E \frown x).$$

Hence $z = t_E \frown x$ and the isomorphism on homology comes from capping with the Thom class. □

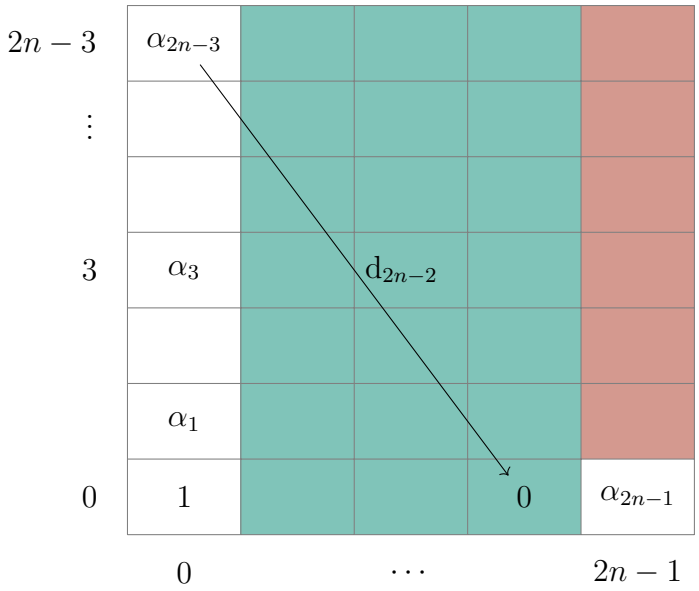
*Cohomology of Lie Groups.* As an essential application of spectral sequences, we find the cohomology rings of the classical compact Lie groups. Some of these results were stated at the end of the chapter on homology.

PROPOSITION 5.18. *The cohomology ring of* $\mathrm{U}(n)$ *is,*

$$H^*(\mathrm{U}(n)) = \Lambda_{\mathbb{Z}}[\alpha_1, \alpha_3, \ldots, \alpha_{2n-1}] \quad \textit{where} \quad |\alpha_i| = i.$$

PROOF. We work by induction, the case $H^*(\mathrm{U}(1)) = H^*(S^1) = \Lambda_{\mathbb{Z}}[\alpha_1]$ is known to us. Suppose the cohomology ring of $\mathrm{U}(n-1)$ is as claimed. Consider the fibration $\mathrm{U}(n-1) \hookrightarrow \mathrm{U}(n) \to S^{2n-1}$ induced by sending a unitary matrix $A$ to the vector in $\mathbb{R}^{2n}$ represented by the first column of $A$. The last $n-1$ columns represent a unitary basis of the orthogonal complement of the first column (hence an element of $\mathrm{U}(n-1)$).

Let $\alpha_{2n-1}$ denote a generator of $H^{2n-1}(S^{2n-1})$. The $E_2$ page of our Serre spectral sequence looks like the following, where we only write down the multiplicative generators.




The green region consists entirely of zeroes. Note all the differentials leaving the elements $\alpha_i$ are trivial. Since these multiplicatively generate the whole page and the differential satisfies a Leibniz rule, we conclude all differentials are zero and thus $E_2 = E_\infty$. We conclude that additively $H^*(\mathrm{U}(n))$ has the claimed description. It remains to show the multiplicative structure is correct (since the fact $E_\infty$ is associated graded to $H^*(\mathrm{U}(n))$ may distort the multiplication). We do know that $\alpha_1, \ldots, \alpha_{2n-1}$ have representatives in $H^*(\mathrm{U}(n))$ (which we'll refer to by the same name) that multiplicatively generate the homology of $\mathrm{U}(n)$. These elements must all anti-commute in the cohomology ring by the super-commutativity of the cup product. Because there is no 2-torsion, we also know $\alpha_i^2 = 0$. There can be no other relations between the $\alpha_i$, otherwise the rank of $H^*(\mathrm{U}(n))$ would be less than the rank of $\Lambda_{\mathbb{Z}_2}[\alpha_1, \ldots, \alpha_{2n-1}]$ contradicting the additive isomorphism. Thus the ring structure of $H^*(\mathrm{U}(n))$ is as claimed. □

PROPOSITION 5.19. *The cohomology ring of* $\mathrm{SU}(n)$ *is,*

$$H^*(\mathrm{SU}(n)) = \Lambda_{\mathbb{Z}}[\alpha_3, \alpha_5, \ldots, \alpha_{2n-1}] \quad \textit{where} \quad |\alpha_i| = i.$$

PROOF. This is the same argument as the $\mathrm{U}(n)$ case. We know $H^*(\mathrm{SU}(2)) = H^*(S^3) = \Lambda_{\mathbb{Z}}[\alpha_3]$ as claimed. We can then work by induction using the fibration $\mathrm{SU}(n-1) \hookrightarrow \mathrm{SU}(n) \to S^{2n-1}$. The spectral sequence will look like the picture above just without the $\alpha_1$ square. The differentials are again all trivial and hence $E_2 = E_\infty$ is additively the correct group. The same reasoning shows a multiplicative isomorphism holds as well. □

PROPOSITION 5.20. *The cohomology ring of the complex Stiefel manifold* $\mathbb{C}V(n,k)$ *is,*

$$H^*(\mathbb{C}V(n,k)) = \Lambda_{\mathbb{Z}}[\beta_{2(n-k)+1}, \beta_{2(n-k)+3}, \ldots, \beta_{2n-1}], \quad \textit{where} \quad |\beta_i| = i.$$

PROOF. Again the argument is the same. We work inductively on $n$ and $k$ simultaneously. In the base case, $\mathbb{C}V(n,1) = S^{2n-1}$ and we have the correct description.

We can then use the Serre spectral sequence of the fibration $\mathbb{C}V(n-1,k-1) \hookrightarrow \mathbb{C}V(n,k) \to \mathbb{C}V(n,1) = S^{2n-1}$. All the differentials will be trivial giving an additive isomorphism and by the same argument a multiplicative isomorphism. □

PROPOSITION 5.21. *The cohomology ring of* $\mathrm{Sp}(n)$ *is,*

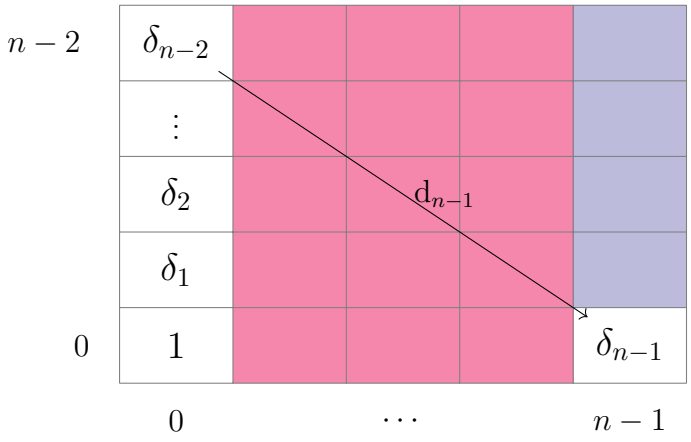
$$H^*(\mathrm{Sp}(n)) = \Lambda_{\mathbb{Z}}[\gamma_3, \gamma_7, \ldots, \gamma_{4n-1}] \quad \textit{where} \quad |\gamma_i| = i.$$

PROOF. Proceed as above, inductively, applying the Serre spectral sequence to the fibration $\mathrm{Sp}(n-1) \hookrightarrow \mathrm{Sp}(n) \to S^{4n-1}$. □

PROPOSITION 5.22. *The mod 2 cohomology of* $\mathrm{SO}(n)$ *is* ***additively*** *given by,*

$$H^*(\mathrm{SO}(n); \mathbb{Z}_2) = \Lambda_{\mathbb{Z}_2}[\delta_1, \delta_2, \ldots, \delta_{n-1}] \quad \textit{where} \quad |\alpha_i| = i.$$

PROOF. We work inductively. The case $n = 2$ is clear. Suppose we know the cohomology of $\mathrm{SO}(n-1)$. Consider the fibration $\mathrm{SO}(n-1) \hookrightarrow \mathrm{SO}(n) \to S^{n-1}$. Our spectral sequence looks like the following, where we let $\delta_{n-1}$ denote the generator of $H^{n-1}(S^{n-1}; \mathbb{Z}_2)$.



Here the pink squares are all zero. The differentials acting on $\delta_i$ are all necessarily zero, except maybe on $\delta_{n-2}$. We need to check the transgression $\mathrm{d}_{n-1}^{0,n-1}$ sends $\delta_{n-2}$ to zero. Consider the following morphism of fibrations given on total spaces by sending a matrix to its first two columns,

$$\begin{array}{ccccc} S^{n-2} & \hookrightarrow & V(n,2) & \longrightarrow & S^{n-1} \\ \uparrow & & \uparrow & & \uparrow{\scriptstyle \mathrm{id}} \\ \mathrm{SO}(n-1) & \hookrightarrow & \mathrm{SO}(n) & \longrightarrow & S^{n-1}. \end{array}$$

This induces a corresponding morphism in the other direction of cohomological Serre spectral sequences. On the square $E_2^{0,n-2}$, this must send the fundamental class of $S^{n-2}$ to $\delta_{n-2}$ (since we inductively know the fibration $\mathrm{SO}(n-1) \to S^{n-2}$ induces an injection in cohomology). So it suffices to show $\mathrm{d}_{n-1}^{0,n-2}$ is zero (mod 2) in the spectral sequence of $V(n,2) \to S^{n-1}$. We know from the Gysin sequence that this differential takes the fundamental class of $S^{n-2}$ to the Euler class of the sphere bundle $V(n,2) \to S^{n-1}$. Note this bundle $V(n,2) \to S^{n-1}$ is actually the unit tangent bundle of $S^{n-1}$ (consider the tangent bundle as a sub-bundle of $T\mathbb{R}^n$ under $S^{n-1} \hookrightarrow \mathbb{R}^n$ using the standard metric). Hence,

$$\mathrm{d}_{n-1}^{0,n-2}[S^{n-2}] = \chi(S^{n-1})[S^{n-1}].$$

But spheres all have even Euler characteristic, and so the differential is zero mod 2. □

**5.5. Generalized Cohomologies and the Atiyah–Hirzebruch Spectral Sequence.** We briefly[1] describe a generalization of the Serre spectral sequence to generalized homology, amely the Atiyah–Hirzebruch spectral sequence.

*K-theory.* To apply the Atiyah–Hirzebruch spectral sequence, we should know a little more about generalized cohomology. Recall a generalized cohomology theory is a contravariant functor from spaces to abelian groups satisfying all the Eilenberg–Steenrod axioms except possibly the dimension axiom. We know that singular homology and stable (co)homotopy are two such theories. We introduce a couple more, beginning with K-theory.

Let $X$ be a space and consider the commutative monoid of equivalence classes of finite dimensional vector bundles on $X$ with addition given by direct sum. There is a classic construction to turn a commutative monoid into an abelian group, called its *Grothendieck group*. What we do is simply consider formal differences $a - b$ of elements of the monoid with the natural addition and the equivalence $c - c = 0$. In our case, we want to study *virtual vector bundles*

[1] Whoops.

$(V_+, V_-)$ on $X$ with the equivalence $(V, V) \sim (0, 0)$. The Grothendieck group of vector bundles on $X$ is called the *$K$-theory* of $X$. If we are considering complex vector bundles, this is denoted $K^0(X)$. For real vector bundles, we write $KO^0(X)$ and for quaternionic vector bundles we write $KSp^0(X)$.

It turns out, for suitably nice spaces, that for every vector bundle $V$, there is a bundle $W$ so that $V \oplus W$ is trivial. This implies that all the elements of $K^0(X)$ have the form $(V, \underline{\mathbb{C}}^n)$, where $\underline{\mathbb{C}}^n$ is a trivial bundle, and this form is unique up to stabilization by direct summing both with another trivial bundle (there are analogous descriptions for real and quaternionic cases).

As we will see in the next chapter, there are classifying spaces $BO(n)$, $BU(n)$, and $BSp(n)$ so that $n$-dimensional real, complex, or quaternionic vector bundles on a finite CW complex $X$ are in bijection with homotopy classes of maps from $X$ to the corresponding classifying space. These spaces can be described as the Grassmannian of $n$-dimensional subspaces of infinite dimensional real, complex, or quaternionic space.

Taking a colimit of the inclusions $\mathrm{U}(1) \hookrightarrow \mathrm{U}(2) \hookrightarrow \cdots$ gives a space U with a classifying space $BU$. Homotopy classes of maps $[X, BU]$ classify stable isomorphism classes of complex vector bundles on $X$. Similarly to define $BO$ and $BSp$. Our $K$-theory group records pairs $(V, n)$ where $V$ is an stable isomorphism class of bundle on $X$ and $n$ is an integer, the dimension of our virtual vector bundle. Thus we have the following description of our $K$-theory groups for suitably nice spaces $X$,

$$\begin{aligned} K^0(X) &= [X_+, BU \times \mathbb{Z}], \\ KO^0(X) &= [X_+, BO \times \mathbb{Z}], \\ KSp^0(X) &= [X_+, BSp \times \mathbb{Z}]. \end{aligned}$$

Here $X_+$ denotes $X$ with an adjoined basepoint, which we need to do in keeping with the convention that $[-, -]$ denotes based homotopy classes.

It turns out these define the zeroth groups of generalized cohomology theories. Using the suspension axiom for cohomology and the loop-suspension adjunction, we are forced into the following definition for the higher $K$ groups of $X$.

**Definition 5.23: $K$-theory**

The *complex $K$-theory* of a space $X$ is the graded group,

$$K^{-i}(X) := [X_+, \Omega^i(BU \times \mathbb{Z})].$$

Similarly, the *real $K$-theory* of $X$ is,

$$KO^{-i}(X) := [X_+, \Omega^i(BO \times \mathbb{Z})],$$

and the *quaternionic $K$-theory* is,

$$KSp^{-i}(X) := [X_+, \Omega^i(BSp \times \mathbb{Z})].$$

This only defines the negative graded part of $K$-theory but we will shortly see from Bott periodicity that this extends to all $i \in \mathbb{Z}$.

These have a group structure coming from the fact loop spaces are $H$-spaces (we'll see in a second $BU, BO, BSp$ are themselves loop spaces too). In fact in the case of $K$ and $KO$, these generalized cohomology theories have a graded multiplication, like the cup product in usual cohomology, given by taking the tensor product of vector bundles. Formally, this means some $K$-theories are ring spectra.

Because $K$-theory is a generalized cohomology theory, the $K$-theory of a point should be non-trival. We see that, $K^{-i}(\mathrm{pt})$ is the group of homotopy classes of maps from a point to $\Omega^i(BU \times \mathbb{Z})$, or equivalently, $\pi_i(BU \times \mathbb{Z})$. Analogously for real and quaternionic theories. We see immediately that $K^0(\mathrm{pt}) = \mathbb{Z}$; this is just the fact that virtual vector spaces (i.e. virtual bundles over a point) are classified by virtual dimension. The same holds in the other cases.

PROPOSITION 5.24. *If $G$ is a topological group, then $G$ is weak homotopy equivalent to $\Omega BG$.*

PROOF. As we will explain in the next chapter, the *classifying space* $BG$ is defined in general as the base of a fibration $EG \to BG$ with fibre $G$ and a contractible total space $EG$ called the *universal bundle*. Consider the path space fibration $\Omega BG \hookrightarrow EBG \to BG$. Since $EBG$ is contractible, we may also think of this as the universal bundle of $\Omega BG$. Let $\mathcal{E}$ be the pullback by $EBG \to BG$ of the universal bundle; by universal properties it is also the pullback by $EG \to BG$ of the universal bundle of $\Omega BG$. We have the following commutative diagram.

$$\begin{array}{ccccc}
 & & G & & G \\
 & & \downarrow & & \downarrow \\
\Omega BG & \hookrightarrow & \mathcal{E} & \longrightarrow & EG \sim \mathrm{pt} \\
 & & \downarrow & & \downarrow \\
\Omega BG & \hookrightarrow & EBG \sim \mathrm{pt} & \longrightarrow & BG.
\end{array}$$

By the long exact sequence of a fibration applied to the horizontal and vertical pullbacks, we see that $\mathcal{E}$ is weak homotopy equivalent to $G$ and to $\Omega BG$. By transitivity, $G \cong_{WHE} \Omega BG$. □

Since $\Omega BG \times \mathbb{Z} = \Omega BG$, we conclude that for $i \geq 1$ the $K$-theories are given by homotopy classes of maps into iterated loop spaces of $U, O$, and $Sp$. In particular for $i \geq 1$,

$$K^{-i}(\text{pt}) = \pi_{i-1}(U), \quad KO^{-i}(\text{pt}) = \pi_{i-1}(O), \quad \text{and} \quad KSp^{-i}(\text{pt}) = \pi_{i-1}(Sp).$$

One of the magical results of algebraic topology is that these groups are periodic. Moreover, iterated loopings of the groups $U, O, Sp$ return back to the group themselves. This implies the $K$-theory of all spaces is periodic. We will not prove this wonderful result; the classical methods are by Morse theory although there are many other approaches.

**Theorem 5.25: Bott Periodicity**

Up to weak homotopy equivalence,

$$\Omega^2 U \cong U, \quad \Omega^8 O \cong O, \quad \text{and} \quad \Omega^8 Sp \cong Sp.$$

Explicitly, there are periodic sequences of loop spaces and their corresponding periodic $\pi$ groups, as depicted in Tables 4 and 5. In particular,the complex $K$-theory of any space is 2-periodic, while the real and quaternionic $K$-theories are 8-periodic. Further, quaternionic $K$-theory is the same as real $K$-theory up to a shift in degree by 4.

| $i \pmod 2$ | 0 | 1 |
|---|---|---|
| $\Omega^i U$ | $U$ | $BU \times \mathbb{Z}$ |
| $K^{i+1}(\text{pt}) = \pi_i(U)$ | 0 | $\mathbb{Z}$ |

TABLE 4. Bott Periodicity for $U$

| $i \pmod 8$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| $\Omega^i O$ | $O$ | $O/U$ | $U/Sp$ | $BSp \times \mathbb{Z}$ | $Sp$ | $Sp/U$ | $U/O$ | $BO \times \mathbb{Z}$ |
| $KO^{-i+1}(\text{pt}) = \pi_i(O)$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | 0 | $\mathbb{Z}$ | 0 | 0 | 0 | $\mathbb{Z}$ |

TABLE 5. Bott Periodicity for $O$ and $Sp$

As rings, these have the following descriptions [Str],

$$\begin{aligned} K^*(\text{pt}) &= \mathbb{Z}[\nu, \nu^{-1}], \\ KO^*(\text{pt}) &= \mathbb{Z}[\eta, \alpha, \beta, \beta^{-1}]/(2\eta, \eta^3, \eta\alpha, \alpha^2 - 4\beta) \\ KSp^*(\text{pt}) &= KO^*(\text{pt}) \cdot \theta, \end{aligned}$$

where $|\nu| = -2$, $|\eta| = -1$, $|\alpha| = -4$, $|\beta| = -8$ and $|\theta| = -4$. Here, $\nu \in \pi_2(BU)$ represents the tautological complex line bundle on $\mathbb{C}P^1$, $\eta \in \pi_1(BO)$ represents the tautological line bundle on $\mathbb{R}P^1$, $\theta \in \pi_4(BSp)$ is the tautological quaternionic line bundle on $\mathbb{H}P^1$ with $\alpha \in \pi_4(BO)$ the underlying real bundle, and $\beta \in \pi_8(BO)$ is the tautological octonionic line bundle on $\mathbb{O}P^1 = S^8$.

EXAMPLE 5.26. Let us find the complex $K$-theory of a sphere. We have,

$$K^0(S^n) = [S^n_+, BU\times\mathbb{Z}] = \mathbb{Z}\oplus\pi_n(BU) = \mathbb{Z}\oplus\pi_{n-1}(U) = \begin{cases}\mathbb{Z} & n \cong 1 \pmod 2\\ \mathbb{Z}^2 & n \cong 0 \pmod 2.\end{cases}$$

And,

$$K^1(S^n) = [S^n_+, U] = \mathbb{Z} \oplus \pi_n(U) = \begin{cases}\mathbb{Z}^2 & n \cong 1 \pmod 2\\ \mathbb{Z} & n \cong 0 \pmod 2.\end{cases}$$

In the case of $n = 1$, we will see in the next chapter that all complex line bundles on $S^1$ are trivial and hence $K^0(S^1) = \mathbb{Z}$ just tracks the dimension of a trivial bundle.

In the case $n = 2$, the complex line bundles on $S^2$ are tensor products of copies of the tautological/Hopf line bundle $\xi$ and its dual. Since $c_1(\xi^{\otimes n}) = nc_1(\xi)$ are all different for different $n$, the integers index a family of stably distinct bundle. Since $K^0(S^2) = \mathbb{Z}^2$, we conclude the stable classes of bundles on $S^2$ are given by $n\underline{\mathbb{C}} \oplus \xi^{\otimes m}$ for $m, n \in \mathbb{Z}$. It turns out that $(\xi \otimes \xi) \oplus \underline{\mathbb{C}} \cong \xi \oplus \xi$. Hence the $K$-theory of $S^2$ may be described as a ring,

$$K^0(S^2) \cong \mathbb{Z}[\xi]/(\xi - 1)^2.$$

▲

*Bordism Theory.* Now we move on to a brief description of bordism. We begin with some important intuition. In homotopy theory, we understand a space $X$ in terms of its holes; we consider all the ways to map spheres $S^n$ into $X$ up to identifying maps $S^n \to X$ connected by a map of cylinders $S^n \times I \to X$. This may seem awfully restrictive in what types of holes we can capture. Why don't we consider maps of arbitrary $n$-manifolds $M$ into $X$? Because we now are dealing with many distinct spaces mapping into $X$ we should define identify maps not just through cylinders but through arbitrary $(n+1)$-manifolds connecting maps of $n$-manifolds. This is essentially the content of bordism theory (although we are often interested in adding extra structure to these manifolds). As an extra note, one could see homology theory as taking this one step further: we consider maps not just from manifolds into $X$ but from arbitrary "closed" CW complexes (i.e. cycles) into $X$ up to boundaries

(this analogy is not quite right because there is the extra structure of a chain complex in the homology case).

Consider a space $X$ and "singular closed $n$-manifolds" in $X$, i.e. continuous maps $f : M^n \to X$ from a closed smooth $n$-manifold $M$ to $X$. Two such singular $n$-manifolds $f : M \to X$, $g : N \to X$ are called *bordant* (or *cobordant*) if there is a compact $(n+1)$-manifold $W$ and a continuous map $F : W \to X$ so that $\partial W = M \sqcup N$ and $F|_{\partial W} = f \sqcup g$. In this case, we call $W$ a *bordism* (or *cobordism*) between $f$ and $g$. If $f : M \to X$ has a bordism to the empty map $\varnothing \to X$ (i.e. there is $g : W \to X$ with $M = \partial W$ and $g|_{\partial W} = f$) we say $f$ is *null-bordant*.

**Definition 5.27: Bordism Groups**

Given a space $X$, its *nth bordism group* is the set,

$$\Omega_n^O(X) := \{\text{singular } n\text{-manifolds in } X\}/\{\text{null-bordant elements}\},$$

equipped with the group operation of disjoint union. That is, we identify bordant singular $n$-manifolds.

If we require that our singular $n$-manifolds be oriented and that our bordisms are oriented compatibly (so that $\partial W = M \sqcup -N$), then we obtain the *nth oriented bordism group*,

$$\Omega_n^{SO}(X) := \{\text{oriented singular } n\text{-manifolds}\}/\{\text{oriented bordism}\}.$$

A *stable almost complex structure* on a smooth manifold $M$ is a complex vector bundle structure on the normal bundle $\nu$ of some embedding $M \hookrightarrow \mathbb{R}^N$. We can study singular $n$-manifolds up to bordisms $\partial W = M \sqcup N$ with stable almost complex structures inducing[a] the given stable almost complex structures on $M$ and on $-N$:

$$\Omega_n^U(X) := \{\text{stable a.c. singular } n\text{-manifolds}\}/\{\text{stable a.c. bordism}\}.$$

There are several more versions of this construction, $\Omega_n^G(X)$ given by asking the stable normal bundles to have a $G$-structure. For example, we could take $G = \text{Spin}, \text{SU}, \text{Sp}$, etc.. More exotically, we could consider $\Omega_n^{\text{fr}}(X)$, cobordisms of manifolds with stable normal framings (i.e. stable trivializations of the normal bundle). Also there are $\Omega_n^{\text{PL}}(X), \Omega_n^{\text{Top}}(X)$, where we extend from cobordisms of smooth to topological or piecewise linear manifolds.

[a] Given a stable almost complex structure on a bordism $W$ coming from an embedding $i : W \to \mathbb{R}^n$ with a normal bundle $\nu_W$ with a complex structure, it will restrict to a bundle $\nu_M \oplus \mathbb{R}$ on a boundary component $M$ with a complex structure. Composing $i|_M$ with the standard embedding $\mathbb{R}^n \to \mathbb{R}^{n+1}$ gives an embedding of $M$ with normal

bundle $\nu_M \oplus \mathbb{R}$, which we know has a complex structure and so gives a stable almost complex structure on $M$.

These all define generalized homology theories; the corresponding cohomology theories are called the *cobordism groups*. Since this is a generalized theory, the homology of a point is non-trivial. The graded groups $\Omega_*^G := \Omega_*^G(\text{pt})$ are called the *cobordism rings*. These are the groups of all closed smooth $n$-manifolds up to abstract cobordism (with the appropriate stable $G$-structure). They have a graded ring structure coming from taking the Cartesian product of spaces.

When $X$ is a manifold, these bordism groups also possess a multiplication, which can be described as follows. Suppose $f : M \to X$ and $g : N \to X$ are two singular manifolds of dimension $n$ and $m$ and $X$ is a closed manifold of dimension $k$. We can consider the product, $f \times g : M \times N \to X \times X$. Let $X_\Delta \subset X \times X$ be the diagonal. We can homotope $f \times g$ so that $P = (f \times g)^{-1}(X_\Delta)$ is a smooth submanifold of $M \times N$ of dimension $n + m - k$. Thus we get a singular manifold $(f \times g)|_P : P \to X$ which up to cobordism is independent of our choice of homotopy and hence defines a multiplication $\Omega_n^G(X) \times \Omega_m^G(X) \to \Omega_{n+m-k}^G(X)$.

Note that every element of $\Omega_*^O$ is 2-torsion. This is because a cylinder $M \times I$ defines a null bordism of $M \sqcup M$ for any closed manifold $M$. The same is not true in $\Omega_*^{SO}$ or $\Omega_*^U$ because $M \times I$ defines the wrong orientation on one of the boundary components, instead this shows us that $-M$ is the additive inverse to $M$ in the bordism group.

Let us compute these groups $\Omega_n^G$ for $G = O, SO$ in low dimensions $n = 0, 1, 2$. Zero dimensional manifolds are points and bordisms between 0-manifolds are line segments. An orientation on a point is given by a choice of sign. We conclude $\Omega_0^O = \mathbb{Z}_2$ (by counting points mod 2) and $\Omega_0^{SO} = \mathbb{Z}$ (by counting signed points).

One dimensional closed manifolds are all circles. These are always null-bordant via the disk. Hence $\Omega_1^O = \Omega_1^{SO} = 0$.

In the two dimensional case, oriented two manifolds are surfaces $\Sigma_g$, which are all null-bordant via handlebodies. Hence $\Omega_2^{SO} = 0$.

To find $\Omega_2^O = 0$ we thus need only study the non-orientable surfaces, which are connected sums of projective planes. The projective plane $\mathbb{R}P^2$ is not null-bordant. This can be seen from the fact $\chi(\mathbb{R}P^2) = 1$ and the following proposition.

PROPOSITION 5.28. *Any closed null-bordant manifold has even Euler characteristic.*

PROOF. If $M$ is odd dimensional, by Poincaré duality it has even Euler characteristic. If $M = \partial N$ is even dimensional, then $X = N \sqcup_M N$ has Euler characteristic $\chi(X) = 2\chi(N) - \chi(M)$. But $X$ is closed and odd dimensional, hence has even Euler characteristic, which implies $M$ does too. □

The Klein bottle $K = \mathbb{R}P^2 \# \mathbb{R}P^2$ is null-bordant. To see this, note the Klein bottle can be formed by gluing the ends of a cylinder with opposite orientations. A null-bordism is just given by filling in the cylinder: see Figure 5 where the pink Klein bottle is filled in by the green interior.

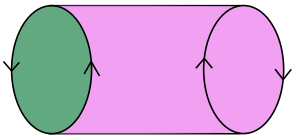


FIGURE 5. Null Bordism of Klein Bottle

Given compact manifolds with boundary $M, N$, we can define a boundary connect sum $M \#_\partial N$ by cutting out small hemispheres in boundary charts of $M$ and $N$, with the flat faces of the hemispheres sitting in the boundaries of $M$ and $N$, and gluing together along a cylinder with boundary gluing the boundaries together. It is clear that $\partial(M \#_\partial N) = \partial M \# \partial N$. In this way, we see that we can boundary connect sum Klein bottles to show that connect sums of Klein bottles are null-bordant. This shows that connect sums of even numbers of copies of $\mathbb{R}P^2$ are null-bordant, and it remains to deal with the odd number cases. For that, we have the following proposition which generalizes the results we just stated on Klein bottles.

PROPOSITION 5.29. *The group operation of disjoint union in the cobordism ring is equivalent to the operation of connect sum. That is, given closed manifolds $M$ and $N$, the connect sum $M \# N$ is cobordant to $M \sqcup N$.*

PROOF. Take a look at the picture proof in figure 6. □

From this we conclude that $\#^{2n+1}\mathbb{R}P^2 = (\#^n T^2)\#\mathbb{R}P^2 = \Sigma_n \#\mathbb{R}P^2$ is bordant to $\mathbb{R}P^2$, since $\Sigma_n$ is null-bordant. Thus, every surface is either null-bordant or bordant to $\mathbb{R}P^2$ and we conclude $\Omega_2^O = \mathbb{Z}_2$.

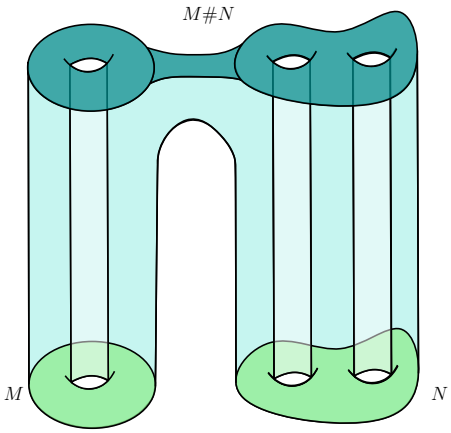



FIGURE 6. A bordism between $M \sqcup N$ and $M\#N$

Analyzing the case $n = 3$ is possible using handle decompositions and one finds $\Omega_3^{SO} = 0$. For $n \geq 4$ the classification of manifolds is too intractable to directly compute bordism groups. It turns out that,

$$\begin{aligned}
\Omega_4^{SO} &= \mathbb{Z} \text{ generated by } \mathbb{C}P^2, \\
\Omega_5^{SO} &= \mathbb{Z}_2 \text{ generated by } SU(3)/SO(3), \\
\Omega_6^{SO} &= 0, \\
\Omega_7^{SO} &= 0, \\
\Omega_8^{SO} &= \mathbb{Z}^2 \text{ generated by } \mathbb{C}P^4 \text{ and } \mathbb{C}P^2 \times \mathbb{C}P^2.
\end{aligned}$$

All higher oriented bordism groups turn out to be non-zero. In spite of the geometric complexity, the methods of characteristic classes, stable homotopy, and spectra allow for computations of the bordism rings, as first undertaken by Thom [**Thom**].

**Theorem 5.30: (Dold, Milnor, Novikov, Quillen, Thom, Wall)**

The unoriented bordism ring is,

$$\Omega_*^O \cong \mathbb{Z}_2[x_i : 1 \leq i \neq 2^k - 1] \cong \mathbb{Z}_2[x_2, x_4, x_5, x_6, x_8, \dots],$$

where $|x_i| = i$. For even $i$ we may take $x_i$ to be the cobordism class of $\mathbb{R}P^i$. For odd $i$ we may take $x_i$ to be the cobordism class of $P(2^r - 1, s2^r)$ where $i = 2^r(2s + 1)$ and $P(m, n)$ is the *Dold manifold* defined as the quotient of $S^m \times \mathbb{C}P^n$ under a $\mathbb{Z}_2$ action given by swapping antipodes on $S^m$ and complex conjugation on $\mathbb{C}P^n$.

Rationally, the oriented and complex bordism rings are,

$$\Omega_*^{SO} \otimes \mathbb{Q} = \mathbb{Q}[y_2, y_4, y_6 \cdots]$$
$$\Omega_*^{U} \otimes \mathbb{Q} = \mathbb{Q}[y_1, y_2, y_3, \cdots],$$

where $|y_i| = 2i$ and we may take $y_i$ to be the cobordism class of $\mathbb{C}P^i$ in both cases. These complex projective spaces do not generate integrally, although it turns out that,

$$\Omega_*^{U} = \mathbb{Z}[a_1, a_2, a_3, \ldots] \quad \text{where} \quad |a_i| = 2i,$$

and modulo 2-torsion, $\Omega_*^{SO}$ is integrally generated by $b_i$ of degree $4i$.

REMARK 5.31. There is a known family of manifolds called the *Milnor hypersurfaces* $H_{ij}$ for $i \leq j$ which multiplicatively generate $\Omega_*^U$ over $\mathbb{Z}$. They are defined as subvarieties of $\mathbb{C}P^i \times \mathbb{C}P^j$ given by,

$$H_{ij} = \{([z_0 : \cdots : z_i], [w_0 : \cdots : w_j]) \in \mathbb{C}P^i \times \mathbb{C}P^j |\ z_0 w_0 + \cdots + z_i w_i = 0\}$$

These are $\mathbb{C}P^{j-1}$-bundles over $\mathbb{C}P^i$. However, there are algebraic relations in the bordism ring between these surfaces so that they are not a free set of generators; there is not a known nice general description for the $a_i$ which form a multiplicative basis.

*The Atiyah–Hirzebruch Spectral Sequence.* Now we can study the Atiyah–Hirzebruch spectral sequence for the computation of generalized homology. We will prove its construction and give a few simple applications.

Recall back to our construction of the Serre spectral sequence that given a homologically simple Serre fibration $p : E \to B$ over a CW base, there was a filtration of the chain complex of $E$ by singular chains contained in the preimages of successive skeleta of $B$, $E_p = p^{-1}(B_p)$. This gave the zeroth page of a spectral sequence: $E_0^{p,q} = C_{p+q}(E_p, E_{p-1})$. The first page of this spectral sequence then had entries $E_1^{p,q} = H_{p+q}(E_p, E_{p-1})$, which we were able to show agreed with $\mathcal{C}_p(B; H_q(f))$.

Now suppose $h_*$ is a generalized homology theory. The first page of a spectral sequence still makes sense,

$$E_{p,q}^1 = h_{p+q}(E_p, E_{p-1}).$$

We follow the argument we used for the Serre case, but diverge to avoid the use of Künneth (which does not hold for generalized homology). By excision and Feldbau's lemma,

$$h_{p+q}(E_p, E_{p-1}) = \bigoplus h_{p+q}(F \times D_\alpha^p, F \times \partial D_\alpha^p),$$

where $F$ is the fibre, and $D^p_\alpha$ are small balls contained in the $p$-cell $e_\alpha$ of $B$. The definition of reduced homology and the relation between relative homology and reduced homology of a cofibre still make sense for a generalized theory,

$$= \bigoplus_\alpha \widetilde{h}_{p+q}(F \times D^p / F \times \partial D^p).$$

With some thought, one sees this quotient as repeated suspensions,

$$= \bigoplus_\alpha \widetilde{h}_{p+q}(\Sigma^p F).$$

It follows from excision that the suspension isomorphism extends to generalized theories,

$$= \bigoplus_\alpha h_q(F)$$
$$= \mathcal{C}_p(B; h_q(F)).$$

As before, the differential will be the usual cellular differential and hence we obtain,

$$E^2_{p,q} = H_p(B; h_q(F)).$$

Going to the $E^\infty$ page, we should obtain the (associated graded) generalized homology of the total space $E$. The question of convergence is a bit subtle, as $h_*(F)$ and $H_*(B)$ may both be non-trivial in infinitely many degrees. But if $B$ is a finite CW then we are definitely okay. Of course an analogous idea holds in cohomology, and the same proof as for Serre gives a multiplication whenever $h^*$ is multiplicative (i.e. represented by a ring spectrum).

**Theorem 5.32: Atiyah–Hirzebuch Spectral Sequence**

Let $p : E \to B$ be a homologically simple Serre fibration with a finite CW base and a fibre $F$. Let $h_*$ be an extraordinary homology theory. Then there is a spectral sequence,

$$E^2_{p,q} = H_p(B; h_q(F)) \implies h_{p+q}(E).$$

And analogously in cohomology,

$$E_2^{p,q} = H^p(B; h^q(F)) \implies h^{p+q}(E).$$

In particular, considering the fibration $* \hookrightarrow X \xrightarrow{\text{id}} X$, one has a spectral sequence,

$$E^2_{p,q} = H^p(X; h^q(*)) \implies h^{p+q}(X).$$

If $h^*$ is a multiplicative cohomology theory, then the cohomological spectral sequences possess a multiplication with the usual properties.

Remark 5.33. The general spectral sequence is sometimes called the Atiyah–Hirzebruch–Serre spectral sequence, with the name Atiyah–Hirzebruch spectral sequence being reserved for the special case of a trivial fibration. The case of a trivial fibration is particularly useful because if we know the cohomology of a space and the stable homotopy groups $h^*(*)$ of a generalized cohomology theory (like in the case of $K$-theory) then we can directly compute the generalized cohomology using our spectral sequence. △

Let us explore a couple properties of this sequence and apply it to a few examples, beginning with $K$-theory.

In the case of complex $K$ theory of a space $X$, remembering $K^0(*) = \mathbb{Z}$ and $K^1(*) = 0$, the Atiyah–Hirzebruch sequence yields the following $E_2$ page.

| | | | | | |
|---|---|---|---|---|---|
| 6 | $H^0(X)$ | $H^1(X)$ | $H^2(X)$ | $H^3(X)$ | $H^4(X)$ |
| | | | | | |
| 4 | $H^0(X)$ | $H^1(X)$ | $H^2(X)$ | $H^3(X)$ | $H^4(X)$ |
| | | | | | |
| 2 | $H^0(X)$ | $H^1(X)$ | $H^2(X)$ | $H^3(X)$ | $H^4(X)$ |
| | | | | | |
| 0 | $H^0(X)$ | $H^1(X)$ | $H^2(X)$ | $H^3(X)$ | $H^4(X)$ |
| | | | | | |
| −2 | $H^0(X)$ | $H^1(X)$ | $H^2(X)$ | $H^3(X)$ | $H^4(X)$ |
| | 0 | | | | 4 |

Clearly the differentials on this page are trivial, and so the first possible action happens on the $E_3$ page.

One can also construct the Atiyah–Hirzebruch spectral sequence using ideas related to Postnikov towers and it turns out the first differential comes from a Postnikov $k$-invariant. In particular, the first differential in an Atiyah–Hirzebruch spectral sequence must be a stable cohomology operation.

Remark 5.34. One interesting corollary of this is the following. It is not hard to show that every stable cohomology operation over $\mathbb{Q}$ is trivial (i.e. the cohomology $H^*(K(\mathbb{Q}, n); \mathbb{Q})$ is concentrated in degree $n$). Hence over $\mathbb{Q}$ there can be no non-trivial first differentials. Hence we have,

$$h^q(X) \otimes \mathbb{Q} = \bigoplus_{r+s=q} H^r(X; \mathbb{Q}) \otimes h^s,$$

for any homology theory $h$.

For the case of complex $K$-theory, we are looking for an operation $H^*(X;\mathbb{Z}) \to H^{*+3}(X;\mathbb{Z})$. It turns out there is only one such non-trivial operation, and one can contruct an example where this differential is non trivial. Hence we have the following.

PROPOSITION 5.35. *The first non-trivial differential in the complex $K$-theory Atiyah–Hirzebruch spectral sequence is,*

$$\mathrm{d}_3 = \beta \circ Sq^2 \circ r : H^*(X;\mathbb{Z}) \to H^{*+3}(X;\mathbb{Z}),$$

*where $\beta : H^*(X;\mathbb{Z}_2) \to H^{*+1}(X;\mathbb{Z})$ is the Bockstein operation coming from $\mathbb{Z} \xrightarrow{\times 2} \mathbb{Z} \to \mathbb{Z}_2$ and $r : H^*(X;\mathbb{Z}) \to H^*(X;\mathbb{Z}_2)$ is reduction mod 2.*

In the case of real $K$-theory, the $E_2$ page looks as follows.

In this case, there can be differentials on the $E_2$ page leaving from the $8k$th and $(8k-1)$th rows and they are given by,

$$\begin{aligned} \mathrm{d}_2^{*,8k} &= Sq^2 \circ r : E_2^{*,8k} &&= H^*(X;\mathbb{Z}) \to H^{*+2}(X;\mathbb{Z}_2) = E_2^{*+2,8k-1} \\ \mathrm{d}_2^{*,8k-1} &= Sq^2 : E_2^{*,8k-1} &&= H^*(X;\mathbb{Z}_2) \to H^{*+2}(X;\mathbb{Z}_2) = E_2^{*+2,8k-2}. \end{aligned}$$

On $E_3$, the first non-trivial differential leaving the $(8k+6)$th rows is,

$$\mathrm{d}_3^{*,8k+6} = \beta \circ Sq^2 : E_3^{*,8k+6} = H^*(X;\mathbb{Z}_2) \to H^{*+3}(X;\mathbb{Z}) = E_3^{*+3,8k+4}.$$

And finally on $E_5$ the differential leaving the $(8k+4)$th rows is,

$$\mathrm{d}_5^{*,8k+4} = \beta \circ Sq^4 \circ r : E_5^{*,8k+4} = H^*(X;\mathbb{Z}) \to H^{*+5}(X;\mathbb{Z}) = E_5^{*+5,8k}.$$

EXAMPLE 5.36. Note since every non-trivial differential of Atiyah–Hirzebruch for complex $K$-theory is on an odd page, the differentials always move an odd

horizontal distance. Hence if the cohomology is concentrated in odd or even degrees, the sequence has to degenerate on the $E_2$ page.

For example, we know the cohomology of $\mathbb{C}P^n$ and the complex Grassmannians is contained in purely in even grades. For the case of $\mathbb{C}P^n$, we have,

$$E_2^{p,q} = E_\infty^{p,q} = \begin{cases} \mathbb{Z} & p,q \text{ even and } 0 \leq p \leq 2n, \\ 0 & \text{otherwise.} \end{cases}$$

We have the associated graded to $K^i(\mathbb{C}P^n)$ is $\bigoplus_p E_\infty^{p,i-p}$ which is $\mathbb{Z}^{n+1}$ if $i$ is even and zero otherwise. Recall that a short exact sequence of abelian groups with final group $\mathbb{Z}$ always splits (i.e. $\mathbb{Z}$ is projective). Hence there are no extension problems and the associated graded is isomorphic to the actual group:

$$K^0(\mathbb{C}P^n) = \mathbb{Z}^{n+1} \quad \text{and} \quad K^1(\mathbb{C}P^n) = 0.$$

Actually since nothing happened in this sequence, and everything was free, this isomorphism is multiplicative,

$$K^*(\mathbb{C}P^n) = \mathbb{Z}[\nu, \nu^{-1}, x]/(x^{n+1}), \quad \text{where} \quad |\nu| = |x| = 1.$$

The element $\nu$ is the complex Hopf line bundle generating $K^*(\text{pt})$ and $x$ is the usual generator of $H^2(\mathbb{C}P^n)$. So we see that $K^0(\mathbb{C}P^n)$ has an explicit basis given by $x^i\nu^{-i}$ for $0 \leq i \leq n$. We already computed the $K$ theory of $S^2 = \mathbb{C}P^1$ in example 5.26. By the naturality of the embedding $\mathbb{C}P^1 \hookrightarrow \mathbb{C}P^n$, we conclude that $x\nu^{-1}$ can be represented by $(1-\xi)$, where $\xi$ is the tautological line bundle on $\mathbb{C}P^n$. Hence,

$$K^0(\mathbb{C}P^n) = \mathbb{Z}[\xi]/(1-\xi)^{n+1}.$$

Since the first non-trivial differential in our sequence was on the $E_3$ page, another setting in which everything immediately degenerates is if the homology is concentrated in degrees $0, 1$, and $2$. For example consider the genus $g$ orientable surface $\Sigma_g$. We see that,

$$E_2^{p,q} = E_\infty^{p,q} = \begin{cases} \mathbb{Z} & q \text{ even and } p = 0, 2, \\ \mathbb{Z}^{2g} & q \text{ even, } p = 1, \\ 0 & \text{otherwise.} \end{cases}$$

The associated graded to $K^i(\Sigma_{2g})$ is $\mathbb{Z}^{2g+2}$ when $i$ is even and zero otherwise. Again everything is free and so this gives the $K$-theory:

$$K^0(\Sigma_g) = \mathbb{Z}^{2g+2} \quad \text{and} \quad K^1(\Sigma_g) = 0.$$

Other examples can be done similarly.

EXERCISE 5.37. Compute the complex $K$-theory of $\mathbb{R}P^n$ and of the complex Grassmannians $\mathbb{C}G(n,k)$. ▲

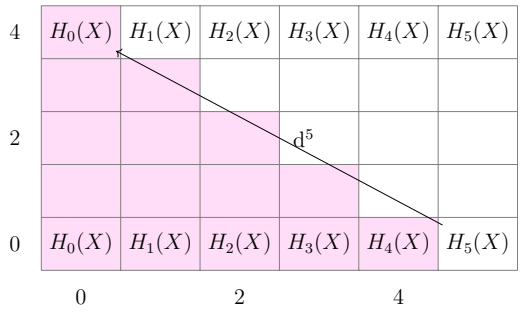
In the case of complex bordism $\Omega_*^U$, the groups for a point are restricted to even dimensions like with complex $K$-theory. We conclude that for spaces with only non-trivial even homology groups the Atiyah–Hirzebruch spectral sequence degenerates on the $E^2$ page. Hence we find for example, that additively,

$$\Omega_*^U(\mathbb{C}P^n) = \Omega_*^U(\mathrm{pt})[x]/(x^{n+1}), \quad \text{where} \quad |x| = 1.$$

If we switch from bordism to cobordism, we can make this isomorphism multiplicative.

REMARK 5.38. This is really expressing that $KU$ and $\Omega_U$ are *complex oriented cohomology theories*. These are multiplicative generalized cohomology theories $E^*$ so that the map $i^* : E^2(\mathbb{C}P^\infty) \to E^2(S^2)$ induced by inclusion is surjective, along with a choice of $t \in \widetilde{E}^2(\mathbb{C}P^\infty)$ so that $i^*(t) \in \widetilde{E}^2(S^2) = E^0(*)$ is the unit element (this is something like a first Chern class). Such generalized cohomology theories are well studied and form the basis for chromatic homotopy theory. △

For oriented bordism, we know the groups in low dimensions are $\Omega_i^{SO}(*) = 0$ for $i = 1, 2, 3$ and $\mathbb{Z}$ for $i = 0, 4$. So for a space $X$, the beginning of the $E^2$ page of the homological Atiyah–Hirzebruch spectral sequence looks like the following.




Recall from our discussion of edge morphisms in the Serre spectral sequence (and it extends easily to Atiyah–Hirzebruch), that the map $\Omega_n(\mathrm{pt}) \to \Omega_n(X)$ should factor through $E^\infty_{n,0}$ by taking the quotient from $E^2$ to $E^\infty$ and then including $E^\infty_{n,0}$ into $\bigoplus_{p+q=n} E^\infty_{p,q}$. But we know that $\Omega_n(\mathrm{pt}) \to \Omega_n(X)$ must be an injection, because there is a right inverse induced from the constant map $X \to \mathrm{pt}$. Hence the quotient map $E^2_{n,0} \to E^\infty_{n,0}$ must be the identity, and so the maps $\mathrm{d}^r : E^r_{r,n-r+1} \to E^r_{0,n}$ in the spectral sequence have to be trivial. In particular, the transgressive map $\mathrm{d}^5$ shown above must be trivial. Hence the

contents of the pink squares never participate in any non-trivial differentials and persist to $E^\infty$.

This tells us the zeroth through third oriented bordism groups of $X$. Assume $X$ is connected so that $H_0(X) = \mathbb{Z}$. To determine $\Omega_4^{SO}(X)$, we need to solve the extension problem,

$$0 \to \mathbb{Z} \to \Omega_4^{SO}(X) \to H_4(X) \to 0.$$

As just discussed, the first map can be identified via the edge morphism with $\Omega_4^{SO}(\text{pt}) \to \Omega_4^{SO}(X)$. But we know this map has a right inverse, and hence by the splitting lemma the above short exact sequence splits. Actually, the inverse $\sigma : \Omega_4^{SO}(X) \to \mathbb{Z}$ can be taken to be the signature map, which sends $f : M \to X$ to the signature $\sigma(M)$ (the signature of the quadratic intersection form $H_2(M) \times H_2(M) \to \mathbb{Z}$); Thom showed this is a cobordism invariant.

Hence we conclude,

$$\Omega_i^{SO}(X) = H_i(X) \quad \text{for} \quad i = 0, 1, 2, 3 \quad \text{and} \quad \Omega_4^{SO}(X) = \mathbb{Z} \oplus H_4(X).$$

We could try to extend this analysis further, since we know some higher bordism groups of a point, but we would have to deal with non-trivial differentials.

This has beautiful applications to the problem of geometric realization, also known as *Steenrod's problem*, which asks if homology classes of a general space can be represented by the fundamental class of singular manifolds. These results are due to Thom [**Thom**].

One can show that the edge morphism $\Omega_n^{SO}(X) \to H_n(X)$ is the map sending a singular manifold $f : M \to X$ to the pushforward homology class $f_*([M]) \in H_n(X)$. Since cobordant singular manifolds define homologous cycles, this is well defined. Since this edge map is a surjection for $n \leq 4$, we conclude that any homology class in $H_n(X)$ is represented by a manifold. For $n > 4$, there may be non-trivial differentials leaving the square $E^*_{n,0}$ and so we cannot say if the morphism $\Omega_n^{SO}(X) \to H_n(X)$ will continue to be surjective. In fact it will be surjective as long as $n \leq 6$. For $n \geq 7$ surjectivity fails, for example Thom discovered a non-geometrically realizable class in $L^7(3) \times L^7(3)$, where $L^7(3)$ is the lens space $S^7/\mathbb{Z}_3$.

However notice, as we remarked earlier, if we work over $\mathbb{Q}$, the Atiyah–Hirzeburch spectral sequence has to immediately degenerate. Hence our rational edge morphism is always surjective, which has the following amazing upshot.

PROPOSITION 5.39 (Thom). *Given a space $X$, and an integral homology class $\alpha \in H_n(X)$. There is an integer $k$ so that $k\alpha$ is represented by $f_*([M])$ for some oriented singular manifold $f : M \to X$ of dimension $n$.*

In fact this can be strengthened slightly. It happens that the spectral sequence immediately degenerates modulo odd torsion (for the same reason as unoriented bordism below). The proof can be found in [**CF**, §14] or [**Rud**, IV §7]. This implies that the coefficient $k$ in the preceding proposition may always be taken to be odd.

If we instead consider unoriented bordism, it can be shown that the unoriented cobordism spectrum is just a direct sum of Eilenberg–MacLane spectra. Since the Atiyah–Hirzerbuch spectral sequence for ordinary homology always immediately degenerates, the same must be true for unoriented cobordism; see [**CF**, §8]. Since $\mathbb{Z}_2$ is a vector space, there can be no extension problems and so one has,

$$\Omega_n^O(X) = \bigoplus_{p+q=n} H_p(X; \Omega_q^O(\mathrm{pt})).$$

This gives a complete description of oriented bordism groups. It also means the edge morphism we just considered is always surjective for $\Omega_n^O$. This of course has the following corollary.

PROPOSITION 5.40 (Thom). *Given a space $X$, and a mod 2 homology class $\alpha \in H_n(X; \mathbb{Z}_2)$, $\alpha$ is represented by $f_*([M])$ for some singular manifold $f : M \to X$ of dimension $n$.*

One final possible application for the Atiyah–Hirzebruch spectral sequence is to stable homotopy. Suppose there were a space $X$ whose homology groups and stable homotopy groups we knew or were not too hard to determine. Then we could apply Atiyah–Hirzebruch:

$$H_p(X; \pi_q^{\mathrm{st}}) \implies \pi_{p+q}^{\mathrm{st}}(X)$$

to possibly determine the stable homotopy groups $\pi_q^{\mathrm{st}}$ of the sphere.

This is actually a promising approach. Kochman is able to compute the first 64 stable homotopy groups of the sphere using this idea and some computer assistance [**Koch**]. The space $X$ he uses is actually not a space but a spectrum, namely the Brown–Peterson spectrum $BP$. We will later see that complex cobordism $\Omega_U^*$ is represented by the complex Thom spectra $MU$. It happens that $MU$ when localized at a prime $p$ becomes homotopy equivalent to a wedge sum of suspensions of spectra; these wedge summands are precisely the $BP$. Another way to construct $BP$ is by formally killing off the odd

homotopy groups of spheres. It can be computed that the homotopy of $BP$ are $\mathbb{Z}_p$ polynomial algebras and the homology can be computed explicitly as a module over the Steenrod algebra. This makes these spaces perfect for the above application.

Let's say a few more words about Atiyah–Hirzebruch for computing stable homotopy of some space $X$. It is important to note that stable homotopy is a reduced homology theory (since we don't adjoin a base point), which means we should use reduced homology in our spectral sequence. Recalling the first few homotopy groups of the spheres, the $E^2$ page of the Atiyah–Hirzebruch spectral sequence will look as follows (assuming $X$ is path-connected).

| | | | | | | |
|---|---|---|---|---|---|---|
| 6 | | $H_1(X;\mathbb{Z}_2)$ | $H_2(X;\mathbb{Z}_2)$ | $H_3(X;\mathbb{Z}_2)$ | $H_4(X;\mathbb{Z}_2)$ | $H_5(X;\mathbb{Z}_2)$ |
| | | | | | | |
| 4 | | | | | | |
| | | $H_1(X;\mathbb{Z}_{24})$ | $H_2(X;\mathbb{Z}_{24})$ | $H_3(X;\mathbb{Z}_{24})$ | $H_4(X;\mathbb{Z}_{24})$ | $H_5(X;\mathbb{Z}_2)$ |
| 2 | | $H_1(X;\mathbb{Z}_2)$ | $H_2(X;\mathbb{Z}_2)$ | $H_3(X;\mathbb{Z}_2)$ | $H_4(X;\mathbb{Z}_2)$ | $H_5(X;\mathbb{Z}_2)$ |
| | | $H_1(X;\mathbb{Z}_2)$ | $H_2(X;\mathbb{Z}_2)$ | $H_3(X;\mathbb{Z}_2)$ | $H_4(X;\mathbb{Z}_2)$ | $H_5(X;\mathbb{Z}_2)$ |
| 0 | | $H_1(X)$ | $H_2(X)$ | $H_3(X)$ | $H_4(X)$ | $H_5(X)$ |
| | 0 | | 2 | | 4 | |

In general, there will be non-trivial differentials which need to be analyzed. Following from the discussion for cohomology, the first non-trivial differential leaving a row must be dual to a stable cohomology operation. For example in the zeroth row we have,

$$\mathrm{d}^2 : H_*(X) \to H_{*-2}(X;\mathbb{Z}_2).$$

One can show this is in general non-trivial. There is only one non-trival operation of degree two, so the differential must be reduction mod 2 composed with the dual of the Steenrod square $\mathrm{Sq}^2$.

In particular, we conclude $\mathrm{d}^2_{3,0} : H_3(X) \to H_1(X;\mathbb{Z}_2)$ is trivial, since the Steenrod square $H_1(X;\mathbb{Z}_2) \to H_3(X;\mathbb{Z}_2)$ is always zero. There is nothing else affecting the second diagonal and so we obtain a short exact sequence,

$$0 \to H_1(X;\mathbb{Z}_2) \to \pi_2^{\mathrm{st}}(X) \to H_2(X) \to 0.$$

In fact this extension problem is always trivial, as explained here. Thus we find for any path-connected space $X$,

$$\pi_1^{\mathrm{st}}(X) = H_1(X) \quad \text{and} \quad \pi_2^{\mathrm{st}}(X) = H_1(X;\mathbb{Z}_2) \oplus H_2(X).$$

Depending on the homology of $X$, one may be able to say more in certain cases.

One further piece of information from the Atiyah–Hirzebruch spectral sequence is the edge homomorphism,

$$h : \pi_i^{\mathrm{st}}(X) \to H_i(X).$$

It is not so surprising that this map is the (stabilization) of the usual Hurewicz homomorphism.

## 6. Characteristic Classes

### 6.1. Vector Bundles and Classifying Spaces.

**Definition 6.1: Principal $G$-Bundles and Vector Bundles**

A locally trivial fibration (we will call this a fibre bundle) $p : E \to B$ is a *principal $G$-bundle* for a topological group $G$ if there is a free right $G$ action on $E$ and $B = E/G$. That is, $G$ acts freely and transitively on the fibres of $p$.

A *rank $n$ vector bundle* is a fibre bundle $p : V \to B$ with fibre $\mathbb{R}^n$ so that the bundle has a covering of local trivializations so that overlapping trivializations differ by continuously varying linear isomorphisms of $\mathbb{R}^n$. Similarly for complex or quaternionic vector bundles but with $\mathbb{C}^n$ or $\mathbb{H}^n$ instead of $\mathbb{R}^n$.

A principal $G$-bundle can be expressed by local data as follows. Suppose $B$ is covered by open sets $U_\alpha$ over which the bundle trivializes as $p : U_\alpha \times G \to U_\alpha$ (we will see taking the $U_\alpha$ contractible suffices). Given two overlapping trivializations $U_\alpha, U_\beta$, there should be a transition map $\varphi_{\alpha\beta} : U_\alpha \cap U_\beta \to G$ so that if $(x, g)$ and $(x, h)$ represent $\widetilde{x} \in E$ in the two trivializations, then $\varphi_{\alpha\beta}(x) = g^{-1}h$. For these transition functions to define a prinicipal $G$-bundle, it is necessary and sufficient that they satisfy the following cocyle conditions,

$$\varphi_{\beta\alpha} = \varphi_{\alpha\beta}^{-1} \quad \text{and} \quad \varphi_{\alpha\beta}\varphi_{\beta\gamma}\varphi_{\gamma\alpha} = \mathrm{id}.$$

Vector bundles can be similarly defined in terms of transition functions $\varphi_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{GL}(\mathbb{R}^n)$.

Remark 6.2. Note that the fibres of $G$-bundles are $G$ torsors, they are homeomorphic to $G$ with an action of $G$ acting like left multiplication but there is no preferred identity element. One way to think of this condition is like a non-linear version of an affine vector space. Complementarily, the fibres of a vector space have a preferred zero element, but there is no way to define

an addition action on the vector space. One way to think of this is like an abstract vector space with no reference to a standard basis or an addition.

If we give either bundle its missing structure: i.e. give the fibres of the principal $G$-bundle an identity element or the fibres of a vector bundle an addition action by $\mathbb{R}^n$, then they become trivial. △

**Definition 6.3**

A morphism between principal $G$-bundles $P, Q$ over $B$ is a $G$-equivariant map $P \to Q$. A morphism between vector bundles $E, E'$ over $B$ is a map commuting with the projection to the base which is linear on the fibres in a local trivialization.

PROPOSITION 6.4. *Principal $G$-bundles have the following properties.*

*(i) A morphism of principal $G$-bundles is an isomorphism.*

*(ii) A principal $G$-bundle with a section is trivial.*

*(iii) Every principal $G$-bundle and every vector bundle over a contractible paracompact base is trivial.*

PROOF. (i) Consider a morphism $\sigma : P \to Q$. If $P$ and $Q$ are trivially equal to $B \times G$, then the morphism has the form $\sigma(b, g) = (b, gf(b))$ for a map $f : B \to G$. But then right multiplication by $f(b)^{-1}$ gives an inverse to $\sigma$. If $P, Q$ non-trivial, we know they are locally trivial and hence $\sigma$ is still an isomorphism since it is locally.

(ii) Given a section $s : B \to P$ of a principal $G$-bundle $P \to B$, there is a morphism $\varphi : B \times G \to P$ given by $\varphi(b, g) = gs(b)$. By (i) this gives an isomorphism of $P$ to a trivial bundle.

(iii) By Feldbau's lemma, a principal $G$-bundle $P \to B$ with $B$ contractible is trivial as a fibration. We need to show this trivialization is $G$-equivariant. Subdivide $B$ into contractible pieces over which $P$ trivializes. On each such piece $P$ possesses a section. Inductively, we can extend our section to each new local trivialization since each additional trivializable neighbourhood is contractible (we need paracompactness for this induction). Thus we obtain a section on all of $B$ and hence $P$ is trivial. For the vector bundle case, look at the associated frame bundle (defined shortly) which must be trivial. □

We now discuss a couple way to get new bundles from old ones .

**Definition 6.5: Pullback Bundles**

Let $p : E \to B$ be a principal $G$-bundle or a vector bundle (or any fibre bundle). And $f : B' \to B$ be a continuous map. The *pullback bundle* $f^*E$ is defined to make the following diagram a pullback square in the categorical sense,

$$\begin{array}{ccc} f^*E & \xrightarrow{\tilde{f}} & E \\ \downarrow{\scriptstyle f^*p} & & \downarrow{\scriptstyle p} \\ B' & \xrightarrow{f} & B \end{array}$$

In terms of universal properties, $f^*E$ is defined so that any bundle $F \to B''$ with a bundle map $F \to E$ that factors as $B'' \to B' \xrightarrow{f} B$ on the base, must factor through $\widetilde{F}$ on the total space. Explicitly, the pullback is given as $f^*E = \{(b', e) | p(e) = f(b')\} \subset B' \times E$. This bundle will be a principal $G$-bundle or a vector bundle when $p$ is.

Recall a $G$-space $F$ is a topological space with a continuous left $G$-action.

**Definition 6.6: Associated Bundle**

If $F$ is a $G$-space and $p : E \to B$ is a principal $G$-bundle, we can form the *associated bundle* $E \times_G F$ defined as the quotient of $E \times F$ by the $G$-action $g \cdot (e, f) = (xg^{-1}, gf)$. Categorically, this is defined to make the following diagram a pushout square,

$$\begin{array}{ccc} E \times_G F & \longleftarrow & E \times F \\ \uparrow & & \uparrow{\scriptstyle \rho_F} \\ E \times F & \xleftarrow[\rho_E]{} & E \times G \times F \end{array}$$

Here the maps $\rho_{E,F}$ denote the action of $G$ on $E, F$ with the identity on the other factor, and the two maps $E \times F \to E \times_G F$ are the quotient map. Note $E \times_G F$ will be a fibre bundle with fibre $F$. Furthermore it will be a bundle with structure group $G$, meaning the transition maps $\varphi_{\alpha\beta} : U_\alpha \cap U_\beta \to \text{Homeo}(F)$ factor through the $G$-action $G \to \text{Homeo}(G)$.

Conversely, suppose $X \to B$ is an $F$-bundle with structure group $G$. Then we obtain a principal $G$-bundle using the transition data $U_\alpha \cap U_\beta \to G$ to determine the local data of the $G$ bundle. If the $G$-action on $F$ is faithful, i.e. $G \to \text{Homeo}(F)$ has no kernel, then this will be inverse to the associated bundle construction.

In our settings of interest $F$ will be a vector space with a faithful $G$-representation. Then the associated bundle $E \times_G F$ will be a vector bundle. Conversely, given a vector bundle $V$ with a faithful $G$-representation on the fibre vector space, we will get a principal $G$-bundle, which we call the *associated frame bundle*. Because the associated constructions are inverse to each other, classifying vector bundles with a given faithful $G$-representation is the same as classifying principal $G$-bundles. We will see some examples shortly.

PROPOSITION 6.7. *Sections of an associated bundle $E \times_G F$ are the same thing as $G$-equivariant maps $E \to F$.*

PROOF. A good exercise. □

Consider a bundle $E \to B$ with structure group $G$. Suppose $G \subset H$. It is always possible to extend our bundle to have structure group $H$ just by considering the transition functions as living in $H$. If $H \subset G$, then the question of *reduction of the structure group* to $H$ is more subtle.

More generally, suppose we have some homomorphism $\varphi : H \to G$. We can ask whether the bundle $E \to B$ with structure group $G$ reduces[2] to an $H$-structure so that the transition maps factor through $\varphi$. In terms of associated principal bundles, a principal $G$-bundle $P$ has a reduction to a principal $H$-bundle $Q$ if there is an (iso)morphism, $Q \times_H G \to P$, where we use $\varphi$ to think of $G$ as an $H$-space.

If $H \subset G$, then the reducibility of structure group of $P$ from $G$ to $H$ is the same as the fibre bundle $P/H$ being trivial. And a choice of reduction is a choice of trivialization, or equivalently a choice of section. This is because a section $\sigma : X \to P/H$ gives a pullback bundle $\sigma^*(P \to P/H)$ which is a reduction of $P$ to structure group $H$. Conversely, a choice $Q \times_H G \cong P$ gives,

$$P/H \cong P \times_G (G/H) \cong Q \times_H G \times_G (G/H) \cong Q \times_H (G/H) \cong B \times G/H.$$

EXAMPLE 6.8. Let us see some examples of the relationship between vector bundles with different structures.

(i) Consider a rank $n$ real vector bundle $E$. Using its standard representation, this has structure group $GL(\mathbb{R}^n)$. It has an associated frame bundle $F_E$ which is a principal $GL(\mathbb{R}^n)$-bundle. Explicitly, the fibre of $F_E$ over $x$ consists of linear isomorphisms of $E_x$. Hence classifying

[2] If $\varphi$ is a surjection rather than injection this is sometimes called a lift instead of reduction of structure group.

real vector bundles on $X$ is the same as classifying principal $GL(\mathbb{R}^n)$-bundles on $X$. The same analysis works for complex and quaternionic bundles with $GL(\mathbb{C}^n)$ and $GL(\mathbb{H}^n)$.

(ii) Consider $O(n)$ with its standard representation. A rank $n$ vector bundle with structure group $O(n)$ is the same thing as a vector bundle with a Riemannian metric, since the fact transition maps are isometries means we can extend a metric across trivializations. The associated frame bundle has fibres consisting of orthogonal transformations of the fibre.

Any $GL(\mathbb{R}^n)$ bundle has a reduction to an $O(n)$ bundle. We can see this in two ways. Geometrically, assuming $X$ is paracompact, a partition of unity argument show a vector bundle on $E$ always inherits a metric. Topologically, $O(n)$ is a deformation retract of $GL(\mathbb{R}^n)$ by Gram–Schmidt; polar decomposition for matrices implies $GL(\mathbb{R}^n)/O(n)$ is the space of positive definite matrices, hence a trivialization is a choice of continuously varying positive definite matrix (this is just a non-degenerate bilinear form). Since $GL(\mathbb{R}^n)/O(n)$ is contractible, a section of $P/O(n)$, i.e. a metric, always exists. This means classifying real vector bundles is the same as classifying principal $O(n)$-bundles.

Analogously, having structure group $U(n)$ or $Sp(n)$ means the bundle is complex with a Hermitian metric or quaternionic with a quaternionic Hermitian metric respectively. By the same argument as above, $GL(\mathbb{C}^n)$ bundles and $GL(\mathbb{H}^n)$ bundles always have a reduction to $U(n)$ and $Sp(n)$ bundles. Hence classifying complex (resp. quaternionic) vector bundles is the same as classifying principal $U(n)$-bundles (resp. $Sp(n)$-bundles).

(iii) A reduction of structure group $GL(\mathbb{C}^n) \subset GL(\mathbb{R}^{2n})$ is a realization of a $2n$-dimensional real bundle as a complex bundle. Such a reduction is a choice of section of $GL(\mathbb{R}^{2n})/GL(\mathbb{C}^n)$, or equivalently an *almost-complex structure*, i.e. a section $J$ of the frame bundle with $J^2 = -I$.

(iv) A reduction of structure group $Sp(2n;\mathbb{R}) \subset GL(\mathbb{R}^{2n})$ is a realization of a $2n$-dimensional real bundle as a symplectic bundle. Such a reduction is a choice of section of $GL(\mathbb{R}^{2n})/Sp(2n;\mathbb{R})$, or equivalently an *almost-symplectic structure*, i.e. a non-degenerate two-form. Since $Sp(2n;\mathbb{R})$ deformation retracts to $U(n)$, every symplectic bundle can be given a complex structure.

(v) Consider a reduction of structure group to $GL^+(\mathbb{R}^n)$; this is the same as orientability of the bundle and a section of the quotient of frame bundles is an orientation. We can always reduce further to $SL(\mathbb{R}^n)$ or even $SO(n)$. This is the same as a volume element (in the case of $SO(n)$ given by the metric). To see this note if $F$ is the associated $O(n)$ frame bundle, then $F/SO(N)$ is a $\mathbb{Z}_2$-bundle measuring the sign of the determinant of a choice of orthonormal basis for each fibre of the bundle. A section of this quotient bundle must be constant, i.e. we need a consistent orientation for the fibres of our space. So classifying oriented vector bundles is the same as classifying principal $SO(n)$-bundles. Note $U(n), Sp(2n;\mathbb{R}) \subset SO(2n)$, hence every complex or symplectic vector bundle determines an underlying oriented real bundle.

The reduction $SU(n) \subset U(n)$ does not have as important a geometric interpretation. One way to express it is that a complex vector bundle $E$ reduces to structure group $SU(n)$ if and only if its complex determinant line bundle $\Lambda^{\text{top}}_{\mathbb{C}} E$ is trivial, and a choice of reduction is a choice of trivialization. This is analogous to the fact a real orientation is a trivialization of the real determinant line bundle. Note $Sp(n) \subset SU(n)$ so every quaternionic vector bundle relaxes to a complex vector with a trivial determinant bundle, and even more generally to an oriented real bundle.

(vi) Consider a reduction of structure group $GL(\mathbb{R}^p, \mathbb{R}^{n-p}) \subset GL(\mathbb{R}^n)$, where $GL(\mathbb{R}^p, \mathbb{R}^{n-p})$ denotes block matrices $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$, for $A$ a $p \times p$ matrix and $C$ an $(n-p) \times (n-p)$ matrix. Such a reduction is the same as a continuously varying choice of $p$-dimensional subspace of the fibres, i.e. a rank $p$ regular distribution.

(vii) A reduction of structure group to the trivial group $\{e\} \subset GL(\mathbb{R}^n)$ means the bundle is parallelizable. A framing of the bundle is determined by a section of the frame bundle, i.e. a trivialization.

(viii) Consider an oriented real bundle $E$ with structure group $SO(n)$. Recall that for $n \geq 3$, $SO(n)$ had a simply connected double cover $Spin(n)$. We can ask about a lift of the structure group via the covering $Spin(n) \to SO(n)$. If such a lift of the associated frame bundle exists, we say $E$ is *spin* and a choice of lift is a *spin structure*. If the tangent bundle of an oriented manifold $M$ is spin, we say $M$ is spin. Suppose $P$ is a principal $Spin(n)$-bundle representing a spin structure on some vector bundle or manifold. Given some representation

of $Spin(n)$ on a vector space $V$ (there are several such representations we might care about), the associated bundle $P \times_{Spin(n)} V$ is called a *spinor bundle*. These are of much interest in mathematical physics because sections of spinor bundles correspond to classical fields representing fermionic particles.

(ix) The group $Spin(n)$ has a central extension to the group $Spin^{\mathbf{c}}(n) = Spin(n) \times_{\mathbb{Z}_2} \mathrm{U}(1)$, which can be thought of as a complex Spin group. Projecting to $Spin(n)$ and applying the double cover gives a map $Spin^{\mathbf{c}}(n) \to SO(n)$. If the associated frame bundle of an oriented vector bundle $E$ admits a lift with respect to this map, a choice of such lift is called a *spin$^{\mathbf{c}}$ structure* on $E$. These are mostly of interest because they allow for many of the same constructions as spin structures, and many more spaces admit spin$^{\mathbf{c}}$ structures on their tangent bundle than spin structures.

(x) A fundamental theorem in the theory of Lie groups asserts that all compact simple Lie groups come in four families of "classical Lie groups": $SU(n), Sp(n), SO(2n+1)$, and $SO(2n)$ except for five "exceptional Lie groups." The smallest of these exceptional groups is $G_2$, which can be realized as the automorphism group of the octonions $\mathbb{O}$ thought of as a normed algebra. The action of $G_2$ on the hyperplane of purely imaginary octonions gives an inclusion $G_2 \hookrightarrow GL(\mathbb{R}^7)$.

Given a seven dimensional vector bundle, a reduction of structure group from $GL(\mathbb{R}^7)$ to $G_2$ is called a *$G_2$-structure*. The inclusion $G_2 \hookrightarrow GL(\mathbb{R}^7)$ factors through $O(7)$, so a $G_2$ structure naturally induces a metric on the bundle. We say a manifold $M$ has a $G_2$-structure if there is a $G_2$-structure on its tangent bundle. We say $M$ is a *$G_2$-manifold* if it has a $G_2$-structure whose associated metric has special holonomy contained in $G_2$. ▲

REMARK 6.9. Suppose $M$ is a smooth manifold with a $G$-structure, i.e. a reduction of the structure group of its tangent bundle to $G$. There are a sequence of integrability conditions one can impose on the $G$-structure. The first integrability condition asks that the $G$-structure (which can always be realized as some tensor on the manifold) be torsion free.

- In the case $G = O(n)$, first integrability means the associated metric on $M$ has a torsion free connection. By the Levi–Civita theorem, such a connection always exists and so every manifold $O(n)$-structure is first integrable.

- In the case $G = GL(\mathbb{C}^n)$, we say our manifold is almost-complex. First integrability is the condition that the Nijenhuis tensor $N_J$ of the complex structure $J$ vanishes (or equivalently that the holomorphic distribution $T^{1,0}_{\mathbb{C}}M$ is involutive/integrable).
- In the case $G = Sp(2n;\mathbb{R})$, we say our manifold is almost-symplectic. First integrability is the condition that the almost-symplectic form is closed.
- In the case $G = G_2$, first integrability is the condition mentioned above that the holonomy of the metric be contained in $G_2$, or equivalently that the metric be Ricci flat.
- In the case $G = U(n)$, one obtains a compatible trio of a metric, almost-complex structure, and almost-symplectic form $(g, J, \omega)$; this is called an almost-Hermitian structure. If the symplectic structure is first-integrable ($\mathrm{d}\omega = 0$), then we say the manifold is almost-Kähler. The $U(n)$ structure is first integrable if it is almost Kähler and its complex structure is first integrable ($N_J = 0$); in this case we call it a *Kähler manifold*.
- In the case $G = GL(\mathbb{R}^p, \mathbb{R}^{n-p})$, first integrability is the condition that the rank $p$-distribution is *involutive*, meaning closed under the Lie bracket.

If all the integrability conditions are satisfied, we say the $G$-structure is *(totally) integrable*. This is equivalent to saying that the manifold has local charts adapted to a trivial $G$-structure.

- The Darboux theorem in symplectic geometry says a first integrable almost-symplectic manifold is automatically integrable (hence locally symplectomorphic to $(\mathbb{R}^{2n}, \omega_{\mathrm{std}})$); in this case we call $M$ a *symplectic manifold*.
- The Newlander–Nirenberg theorem says a first integrable almost-complex manifold is also automatically integrable (hence locally holomorphic to $\mathbb{C}^n$); in this case we call $M$ a *complex manifold*. Note this implies Kähler manifolds are both complex and symplectic.
- An $Sp(n)$ structure which is first integrable is also automatically integrable and in this case we say $M$ is a *hyper-Kähler manifold*.
- The Frobenius theorem says a first integrable rank $p$ regular distribution is always totally integrable and thus defines a rank $p$ foliation of $M$.
- In contrast, a first integrable Riemannian manifold need not be totally integrable. Total integrability of an $O(n)$-structure is equivalent to the manifold being flat, and there are many local curvature obstructions to this being the case.
- A $G_2$ manifold is also usually not totally integrable. △

Associated to any topological group $G$, there is an important space that will help us classify principal $G$-bundles and hence vector bundles with structure group $G$.

**Definition 6.10: Classifying Space**

Given a topological group $G$, the *classifying space* $BG$ of $G$ is the space uniquely defined up to weak homotopy equivalence, so that there is a weakly contractible space $EG$ and a principal $G$-bundle $EG \to BG$ called the *universal bundle* of $G$.

The fact that classifying spaces are unique will follow shortly from our central theorem. But we can prove they always exist in a fairly explicit way.

**Theorem 6.11: Milnor's Construction**

For any topological group $G$, there is a classifying space $BG$.

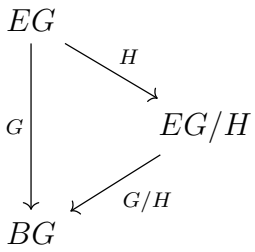
Proof. Take $EG$ to be the infinite join of copies of $G$ (i.e. the colimit of the join of $n$ copies of $G$),

$$EG = G * G * G * \cdots .$$

A more intuitive way of writing this is that $EG$ consists of infinite vectors $(t_1 g_1, t_2 g_2, \ldots)$ where $t_i \in [0,1]$, $g_i \in G$, only finitely many of the $t_i$ are non-zero, and the $t_i$ sum to one (note we are also identifying entries $0g_i = 0\widetilde{g}_i$). The group $G$ has a continuous free action on $EG$ by right multiplication on each entry of our infinite vector. Set $BG = EG/G$. By construction, this is a principal $G$-bundle (e.g. $(t_1 e, t_2 g_2, \cdots)$ gives a local section/trivialization over the region $t_1 \neq 0$ ).

Note $EG$ is contractible. Given a spheroid $S^n \to EG$, it must land in some finite join $G * \cdots * G$. But then it must be contract inside the next join since $(A, 0)$ is always contractible inside $A * B$. □

If $EG \to BG$ is a universal bundle and $H \subset G$ is a subgroup, then we can consider first the quotient of $EG$ by $H$ and then by $G/H$:



We conclude that $BH = EG/H$ and we obtain a fibration,

$$G/H \hookrightarrow BH \to BG.$$

It is also clear, using products of universal bundles that,

$$B(G_1 \times G_2) = BG_1 \times BG_2.$$

EXAMPLE 6.12. We list several classifying spaces of interest.

(1) If $G = \mathbb{Z}_2$, then $S^\infty \to \mathbb{R}P^\infty$ is a principal $\mathbb{Z}_2$ bundle with contractible total space. So $B\mathbb{Z}_2 = \mathbb{R}P^\infty$.

(2) If $G = \mathbb{Z}$, then $\exp(2\pi i -) : \mathbb{R} \to S^1$ is a principal $\mathbb{Z}$-bundle. So $B\mathbb{Z} = S^1$.

(3) More generally, whenever $G$ is discrete $BG = K(G, 1)$.

(4) For $G = O(n)$, the universal bundle is $V(\infty, n) \to G(\infty, n)$. This bundle is given by sending an orthonormal $n$-frame in $\mathbb{R}^\infty$ to the subspace it spans. Note that $V(\infty, 1) = S^\infty$ is contractible, and the fibration,

$$V(\infty, n-1) \hookrightarrow V(\infty, n) \to S^\infty,$$

shows inductively by the homotopy long exact sequence that $V(\infty, n)$ is always weakly contractible (and contractible by Whitehead). Hence $BO(n) = G(\infty, n)$ and $EO(n) = V(\infty, n)$.

For analogous reasons, $BU(n) = \mathbb{C}G(\infty, n)$ and $BSp(n) = \mathbb{H}G(\infty, n)$, and $EU(n) = \mathbb{C}V(\infty, n)$ and $ESp(n) = \mathbb{H}V(\infty, n)$. Of particular interest is that $BU(1) = \mathbb{C}P^\infty$. Using the fact about classifying spaces of products, $BT^n = (\mathbb{C}P^\infty)^n$.

(5) For $G = SO(n)$, note this is a subgroup of $O(n)$. Hence,

$$BSO(n) = V(\infty, n)/SO(n) = G_+(\infty, n).$$

We also obtain a fibration $O(n)/SO(n) \to G_+(\infty, n) \to G(\infty, n)$ which is just the usual double cover. Analogously one has $BSU(n) = \mathbb{C}V(\infty, n)/SU(n)$.

Consider a discrete group $G$ and a nice enough space $X$, say a CW complex. A path connected principal $G$-bundle $P$ on $X$ is a connected covering space of $X$. The $G$-action on the fibres are the deck transformations of the covering and the fact the $G$-action is transitive on fibres, means that $P$ is a normal covering. Conversely, any normal covering of $X$ with deck transformation group $G$ is a connected principal $G$-bundle on $X$ and an isomorphism of normal coverings gives an isomorphism of bundles and vice versa. Since the covering is normal,

isomorphism classes of coverings of $X$ with deck transformation group $G$ are in bijection with groups $H$ so that $\pi_1(X)/G \cong H$. This is the same as an exact sequence,

$$0 \to H \to \pi_1(X) \to G \to 0.$$

Up to isomorphism class of $H$, it is enough that we just know the surjective map $\pi_1(X) \to G$. We conclude connected principal $G$-bundles are classified by surjective homomorphisms $\pi_1(X) \to G$.

If we generalize to consider principal $G$-bundles $P$ that are not connected, then we know each path component is a regular covering of $X$ for a subgroup $G' \subset G$. Moreover, the set of path components will have the structure of a $G/G'$ torsor. Thus, we find principal $G$-bundles are classified by arbitrary homomorphisms $\varphi : \pi_1(X) \to G$; the map $\pi_1(X) \to \mathrm{im}(\varphi)$ determines a connected bundle and the inclusion $\mathrm{im}(\varphi) \to G$ describes the torsor structure of the set of path components.

To reiterate, we have found for $G$ discrete,

$$\{\text{Principal } G\text{-bundles on } X\}/\sim \quad \longleftrightarrow \quad \mathrm{Hom}(\pi_1(X), G).$$

We can make this characterization purely in terms of $X$ instead of its fundamental group. Using the representability of cohomology and the universal coefficient theorem (these will both work fine for $G$ non-abelian in rank one) along with Hurewicz,

$$[X, K(G,1)] = H^1(X; G) = \mathrm{Hom}(H_1(X), G) = \mathrm{Hom}(\pi_1(X), G).$$

Since $K(G,1) = BG$, we conclude,

$$\{\text{Principal } G\text{-bundles on } X\}/\sim \quad \stackrel{\cong}{\to} \quad [X, BG].$$

We can actually make this isomorphic explicit. We see that tracing through the definitions, a principal bundle $P$ on $X$ is determined from the map $f : X \to BG$ by pulling back the universal bundle: $P = f^*EG$.

This wonderful result, that principal $G$ bundles for $G$ discrete are represented by maps to $BG$, generalizes verbatim to arbitrary topological groups. The classifying spaces $BG$ are so named because they classify isomorphism classes of bundles. That is, $BG$ is a representing object for the functor,

$$\mathrm{PBund}_G : \mathbf{CWhTop}^{\mathrm{op}} \to \mathbf{Set},$$

taking a CW complex $X$ to the set of isomorphism classes of principal $G$-bundles on $X$.

**Theorem 6.13: The Classification of Principal Bundles**

Let $\mathrm{PBund}_G(X)$ denote the equivalence classes of principal $G$-bundles on a CW complex $X$. Then there is a correspondence,

$$\mathrm{PBund}_G(X) \cong [X, BG],$$

given by sending $f : X \to BG$ to the pullback bundle $f^*EG$.

LEMMA 6.14. *Suppose $p : E \to B$ is a principal $G$-bundle and $B'$ is a CW complex. If two maps $g, h : B' \to B$ are homotopic, then the pullback bundles $g^*E$ and $h^*E$ are equivalent.*

PROOF. Let $F : B' \times [0,1] \to B$ be a homotopy between $g$ and $h$. There is a pullback bundle $\mathcal{E} = F^*E$ defined over $B' \times [0,1]$. It suffices to show $\mathcal{E}$ is equivalent to $h^*E \times [0,1]$, or equivalently to construct a morphism $\mathcal{E} \to h^*E \times [0,1]$. This is the same thing as finding a section of $\mathcal{E} \times_G (h^*E \times [0,1]) \to B' \times [0,1]$. On the zero skeleton of $B$, this bundle is defined over a $[0,1]$, hence it is trivial and must possess a section. Now inductively, since this bundle is in particular a Serre fibration, we can extend this section over increasing skeleta of $B'$ to obtain a section and hence a morphism $\mathcal{E} \to h^*E \times [0,1]$. □

PROOF. (of Theorem 6.13) The above lemma shows the correspondence in the theorem statement is a well defined map from $[X, BG]$ to $\mathrm{PBund}_G(X)$. We wish to show it is a bijection.

We begin with surjectivity. Given a principal $G$-bundle $P \to B$, consider the associated bundle $P \times_G EG \to B$. The fibre of this bundle is $EG$, which is weakly contractible, and hence this bundle has a section (by obstruction theory). Such a section defines a $G$-equivariant map $P \to EG$. Since this is $G$-equivariant, we may pass to the quotient spaces to obtain a map $f : B \to BG$. By universal properties, $f^*EG = P$.

Now we show injectivity. Suppose $f_0, f_1 : B \to BG$ are two maps so that $f_0^*EG$ and $f_1^*EG$ are isomorphic. Let $P = (f_0^*EG) \times [0,1] \to B \times [0,1]$. There is an associated bundle $Q = P \times_G EG \to B \times [0,1]$. The $G$-equivariant bundle maps $f_0^*EG \to EG$ and $f_0^*EG \xrightarrow{\cong} f_1^*EG \to EG$ define two sections of $(f_0^*EG) \times_G EG \to B$ and hence a section of $Q$ over $B \times \{0\} \cup B \times \{1\}$. Since $Q$ has weakly contractible fibre, this section extends to one over all of $B \times [0,1]$ (by obstruction theory). This then gives a $G$-equivariant bundle map $P \to EG$. We may pass to the quotient to obtain a map $B \times [0,1] \to BG$ which is a homotopy between $f_0$ and $f_1$. □

COROLLARY 6.15. *The space $BG$ is unique up to weak homotopy equivalence.*

PROOF. Suppose we have another model for a universal bundle $\widetilde{EG} \to \widetilde{BG}$. By Theorem 6.13, there is a map $f : \widetilde{BG} \to BG$ so that $f^*EG = \widetilde{EG}$. Let $F : \widetilde{EG} \to EG$ be the induced map on total spaces. We thus have a morphism of fibrations,

$$\begin{array}{ccccc} G & \hookrightarrow & EG & \longrightarrow & BG \\ {\scriptstyle \mathrm{id}}\uparrow & & {\scriptstyle F}\uparrow & & {\scriptstyle f}\uparrow \\ G & \hookrightarrow & \widetilde{EG} & \longrightarrow & \widetilde{BG}. \end{array}$$

Note that $\mathrm{id} : G \to G$ clearly induces an isomorphism on homotopy groups. And $F$ does as well since $EG, \widetilde{EG}$ are contractible. But then by the five lemma applied to the homotopy long exact sequences of these fibration, $f$ must also induce isomorphisms of homotopy groups. Hence $f$ is a weak homotopy equivalence of $\widetilde{BG}$ and $BG$. □

One can also show on general abstract nonsense grounds that any space representing the functor $\mathrm{PBund}_G$ must be weak homotopy equivalent to $BG$, and hence is the base of a $G$-bundle with contractible total space.

EXAMPLE 6.16. This is a very powerful result and we can make preliminary classifications of some low dimensional bundles using it.

(1) Consider the set of rank one real vector bundles on $X$. By above, this is the same as classifying principal $O(1) = \mathbb{Z}_2$ bundles on $X$. Note $B\mathbb{Z}_2 = \mathbb{R}P^\infty = K(\mathbb{Z}_2, 1)$. So,

$$\mathrm{PBund}_{O(1)}(X) \cong [X, B\mathbb{Z}_2] \cong [X, \mathbb{R}P^\infty] \cong H^1(X; \mathbb{Z}_2).$$

So real line bundles are classified by elements of $H^1(X; \mathbb{Z}_2)$. We will see shortly that the cohomology class corresponding to a line bundle $E$ under this classification is called the first Stiefel-Whitney class $w_1(E)$.

In the case of $X = S^1$, $H^1(S^1; \mathbb{Z}_2) = \mathbb{Z}_2$. The two isomorphism classes of line bundles on $S^1$ are the trivial bundle and the Möbius strip (which can't be trivial since it is non-orientable). By naturality, we conclude that for any connected CW complex $X$ and a line bundle $E \to X$, the classifying cohomology class $w_1(E)$ evaluated on a 1-cell $e_1$ of $X$ gives 0 or $1 \in \mathbb{Z}_2$ if $E$ is the trivial bundle or the Möbius strip respectively when pulled back to $e_1 \sqcup e_0$.

(2) Now consider rank one complex vector bundles on $X$ or equivalently principal $U(1)$-bundles. Note $BU(1) = \mathbb{C}P^\infty = K(\mathbb{Z}, 2)$. So,

$$\mathrm{PBund}_{U(1)}(X) \cong [X, BU(1)] \cong [X, \mathbb{C}P^\infty] \cong H^2(X; \mathbb{Z}).$$

Thus complex line bundles are classified by elements of $H^2(X; \mathbb{Z})$. The cohomology class corresponding to a complex line bundle $\mathcal{E}$ under this classification is called the first Chern class $c_1(\mathcal{E})$.

For $X = S^2$, $H^2(S^2; \mathbb{Z}) = \mathbb{Z}$. These bundles are classified by their degree: $S^2 = \mathbb{C}P^1$ has tautological line bundle $\mathcal{O}(-1)$, its dual $\mathcal{O}(1)$, and a $\mathbb{Z}$'s worth of bundles $\mathcal{O}(\pm k) = \mathcal{O}(\pm 1)^{\otimes k}$.

(3)

EXERCISE 6.17. Classify principal $\mathrm{SL}(2, \mathbb{C})$-bundles on $\mathbb{C}P^2$. ▲

Suppose one has a map of topological groups $f : G \to H$. Consider the associated bundle $EG \times_G H$, where we let $G$ act on $H$ via $f$. This is a principal $H$-bundle over $BG$ and so by Theorem 6.13, it must be induced by a unique homotopy class of map $Bf : BG \to BH$. This map can also be recovered as follows. Given a principal $G$-bundle $P$ on $X$, we obtain an associated principal $H$-bundle $P \times_G H$ on $X$. This gives a map $\mathrm{PBund}_G(X) \to \mathrm{PBund}_H(X)$ for any $X$, i.e. a natural transformation $\mathrm{PBund}_G \to \mathrm{PBund}_H$. By the Yoneda lemma, this corresponds to a map $BG \to BH$ which must agree with $Bf$. It is clear this map is functorial in the appropriate senses. Hence the "classifying operation" $B(\cdot)$ defines a functor from topological groups **TopGp** to **hTop**.

EXAMPLE 6.18. Consider the usual short exact sequence of a subgroup $H \subset G \to G/H$. Functorially, we get a fibration,

$$BH \to BG \to B(G/H).$$

This fits in with a previously obtained map to give a fibration sequence,

$$G/H \to BH \to BG \to B(G/H).$$

We have proved that $\Omega B(G/H) \cong G/H$ and so this fibre sequence is nothing but a portion of the Puppe sequence for $BH \to BG$. △

**6.2. Abstract Characteristic Classes and the Cohomology of Classifying Spaces.** We saw in Example 6.16 that we could classify line bundles using cohomology. While this is too strong to hope for in general, we may

reasonably conclude that homotopy classes of maps $[X, BG]$ and hence $G$-bundles can be distinguished through invariants with values in cohomology classes. This is precisely the idea of characteristic classes.

**Definition 6.19: Characteristic Classes**

A *characteristic class* $c$ of $G$-bundles over a coefficient group $F$ is an assignment of a cohomology class $c(E) \in H^*(X; F)$ to every principal $G$-bundle $E$ on every space $X$ which is natural with respect to pullbacks. That is for $f : X \to Y$ and a bundle $E$ on $Y$,

$$c(f^*E) = f^*c(E).$$

More succinctly, a characteristic class is a natural transformation,

$$c : \mathrm{PBund}_G \to H^*(-; F).$$

We will usually consider characteristic classes for vector bundles which we take to be the characteristic class of their associated frame bundle.

Note that a characteristic class is uniquely determined by its value on the universal bundle $EG$, since every bundle is the pullback of $EG$ and using naturality. And moreover, any cohomology class of $BG$ defines a characteristic class $c(E)$ by looking at its pullback under the defining map $X \to BG$ of a bundle $E \to X$. Hence (this is just the Yoneda lemma) characteristic classes of $G$-bundles are the same thing as cohomology classes of $BG$.

So, to understand characteristic classes on real/complex/oriented vector bundles, we will want to understand the cohomology of $BO(n)$, $BU(n)$ and $BSO(n)$. We will obtain the cohomology of these spaces, and hence a list of important characteristic classes, from spectral sequences.

Remark 6.20. The preceding definition could be replaced with asking for a natural transformation $\mathrm{PBund}_G \to E^*$, where $E$ is some generalized cohomology theory. This will give us characteristic classes in a generalized cohomology theory and by Yoneda they are determined by $E^*(BG)$. These are of interest in many situations, particularly for $K$-theory. Also well studied are characteristic classes in the equivariant cohomology of an equivariant bundle. △

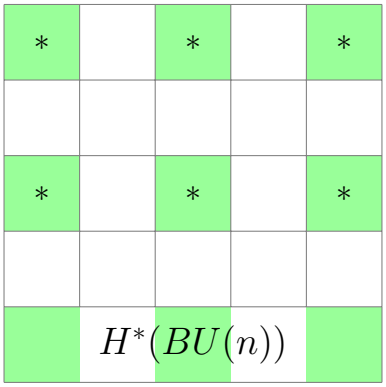
Also note that whenever we have a map $G \to H$, and hence a map $BG \to BH$, this gives a way to pullback characteristic classes of $H$-bundles to characteristic classes of $G$-bundle corresponding to a pullback of bundles. Hence whenever we have a geometric way to turn one bundle into another, we will get some relation between characteristic classes; we will use this idea several times.

Now we get to work computing the cohomology of our classifying spaces.

**Homology of $BU(n)$:** Note that the torus $T^n$ is a subgroup of $U(n)$ by considering unitary matrices which are diagonal. Thus we get a fibration,

$$U(n)/T^n \hookrightarrow BT^n \to BU(n).$$

Note that $U(n)/T^n$ is the manifold of complete flags in $\mathbb{C}^n$. This has a cell structure with only even dimensional cells by the standard Bruhat description. $BU(n)$ is the infinite dimensional complex Grassmannians and using the Schubert cell decomposition in the direct limit, again the cells are all even. Consider the associated cohomological spectral sequence on the $E_2$ page.



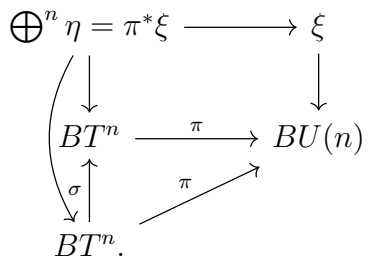

Note the only non-trivial entries are in the green squares, since the cellular chain groups and hence cohomology groups of $BU(n)$ and $U(n)/T^n$ are only non-trivial in even degrees. But then we see that all differentials are trivial and so $E_2 = E_\infty$ is the associated graded of the cohomology of $BT^n$. In particular, the natural edge morphism given by pullback under the fibration,

$$\pi^* : H^*(BU(n)) \to H^*(BT^n),$$

is an injection.

Note that $BT^n = (BS^1)^n = (\mathbb{C}P^\infty)^n$. Its cohomology is the free polynomial algebra generated by elements $x_1, \ldots, x_n$ each of degree two corresponding to each copy of $\mathbb{C}P^\infty$ in $(\mathbb{C}P^\infty)^n$. We claim that $\pi^*$ lands in the subalgebra of $\mathbb{Z}[x_1, \ldots, x_n]$ generated by symmetric polynomials. To see this let $\sigma \in S_n$ act on $(\mathbb{C}P^\infty)^n$ by permutation of factors. By naturality, the map $\pi : BT^n \to BU(n)$ must pull back $EU(n)$ to the $ET^n$. Let $\xi$ be the associated $\mathbb{C}^n$ bundle to $EU(n)$ and $\eta$ the complex line bundle associated to $S^\infty \to \mathbb{C}P^\infty$. We have a

diagram of the above form,




This should commute up to homotopy, because the permutation $\sigma$ induces an isomorphism of the vector bundles $\pi^*\xi$ lying above (it just shuffles the copies of $\eta$). Thus given $\alpha \in H^*(BU(n))$, we have $\pi^*\alpha = \sigma^*\pi^*\alpha$. Since clearly $\sigma^*(x_{\sigma(i)}) = x_i$, we conclude $\pi^*$ is an injection into the space of symmetric polynomials of the $x_i$. In fact it has to be an isomorphism onto this subalgebra. This just follows from cell counting. We know $\dim H^i(BU(n))$ is the number of Young Tableaux diagrams with $i$ squares in an "$\infty \times n$ grid". It is a basic combinatorially exercise to see this is also the dimension of degree $i$ symmetric polynomials in $n$ variables. Note the map $\pi^*$ is also a ring map. Thus we conclude as rings,

$$H^*(BU(n)) \cong \mathbb{Z}[\sigma_1(\vec{x}), \ldots, \sigma_n(\vec{x})] \quad \text{where} \quad |\sigma_i| = 2i.$$

Here $\sigma_i(\vec{x})$ is the $i$th elementary symmetric polynomial in $x_1, \ldots, x_n$. It is customary to write $c_i$ for $\sigma_i(\vec{x})$; this is called the *ith universal Chern class*. The corresponding characteristic class $c_i(E)$ for a complex vector bundle $E$ is called the *ith Chern class of* $E$. The elementary symmetric polynomials and thus the classes $c_i$ can be expressed in terms of the $x_i$ using the following simple formula,

$$(1 + x_1) \cdots (1 + x_n) = 1 + c_1 + \cdots + c_n := c \in H^*(BU(n)),$$

where $c_i$ is the homogeneous part of degree $2i$. This element $c$ in the cohomology ring (and its pullback via any classifying map) is called the *total Chern class*. By naturality, the inclusion $BU(n) \hookrightarrow BU(n+1)$ induces a pullback map which is the identity on the classes $c_i$ for $i \leq n$ and sends $c_{n+1}$ to zero. In particular, the first Chern class $c_1$ can clearly be seen to agree with the characteristic class for complex line bundles we defined before under the same name.

**Homology of $BSp(n)$:** Exactly the same analysis goes through for the quaternionic case. Let $\mathbb{H}^*(B(S^3)^n) = H^*((\mathbb{H}P^\infty)^n) = \mathbb{Z}[y_1, \ldots, y_n]$ where $|y_i| = 4$. We obtain an injection,

$$H^*(BSp(n)) \to H^*((\mathbb{H}P^\infty)^n),$$

which is an isomorphism onto symmetric polynomials of the $y_i$. So if we let,

$$(1+y_1)\cdots(1+y_n) = 1 + p_1 + \cdots + p_n,$$

where $p_i$ is the homogeneous part of degree $4i$, we find,

$$H^*(BSp(n)) \cong \mathbb{Z}[p_1, \ldots, p_n] \quad \text{where} \quad |p_i| = 4i.$$

The element $p_i$ is called the *ith universal Pontryagin class* and for a quaternionic vector bundle $E$ it corresponds to a characteristic class $p_i(E)$ called *ith Pontryagin class of $E$.*

**Homology of $BO(n)$ over $\mathbb{Z}_2$:** The integral homology of $O(n)$ is more complicated, but we can work modulo 2 and obtain a very similar analysis.

Let $H^*(B(\mathbb{Z}_2)^n; \mathbb{Z}_2) = H^*((\mathbb{R}P^\infty)^n) = \mathbb{Z}_2[t_1, \ldots, t_i]$ where $|t_i| = 1$. We can consider the usual map induced by inclusion,

$$H^*(BO(n); \mathbb{Z}_2) \to H^*(B\mathbb{Z}_2^n; \mathbb{Z}_2).$$

We may consider the fibration $O(n)/\mathbb{Z}_2^n \to B\mathbb{Z}_2^n \to BO(n)$ and the associated spectral sequence (it is slightly subtle to show this fibration is homologically simple, but we will omit the details). Over $\mathbb{Z}_2$ we know the dimension of the homology of all these spaces (by cell counts) and we deduce that the spectral sequence must immediately degenerate (otherwise $H^*(B\mathbb{Z}_2^n; \mathbb{Z}_2)$ would not have enough generators). Hence the map above is an injection and by the same argument as above, using the commutative diagram and a cell count, it defines an isomorphism onto the subalgebra of symmetric polynomials.

So letting,

$$(1+t_1)\cdots(1+t_n) = 1 + w_1 + \cdots + w_n,$$

where $w_i$ is the homogeneous part of degree $i$, we find,

$$H^*(BO(n); \mathbb{Z}_2) \cong \mathbb{Z}_2[w_1, \ldots, w_n] \quad \text{where} \quad |w_i| = i.$$

The element $w_i$ is called the *ith universal Stiefel–Whitney class* and for a real vector bundle $E$ it corresponds to a characteristic class $w_i(E)$ called *ith Stiefel–Whitney class of $E$.*

**Homology of $BSO(n)$ over $\mathbb{Q}$:** We need a certain lemma first. Recall a largest abelian subgroup of a compact connected Lie group $G$ is necessarily isomorphic to $T^n$ for some $n$ and is called the *maximal torus*. There is then an exact sequence,

$$1 \to T^n \to N_G(T^n) \to W \to 1,$$

defining a finite group $W$ known as the *Weyl group* of $G$.

PROPOSITION 6.21. *For a compact connected Lie group $G$ with maximal torus $T$ and Weyl group $W$,*

$$H^*(BG;\mathbb{Q}) \cong H^*(BT;\mathbb{Q})^W,$$

*where $H^*(\cdot)^W$ denotes the $W$-invariant subalgebra under $W$'s action induced by left multiplication on $T \subset G$.*

PROOF. (Sketch) We proceed in a few steps.

(1) Let $p : \widetilde{X} \to X$ be a normal covering map with the structure of a connected principal $\pi$-bundle for $\pi$ a finite discrete group. Then, by the map lifting lemma, there is an isomorphism of singular chain groups $C_*(\widetilde{X};\mathbb{Q})^\pi \cong C_*(X;\mathbb{Q})$. Passing to homology,

$$H^*(X;\mathbb{Q}) = H^*(\widetilde{X};\mathbb{Q})^\pi.$$

(2) The quotient $G/T$ is a "generalized flag manifold," and there is always a generalized Bruhat partition allowing us to realize $G/T$ as a cell complex with $|W|$ cells of only even dimension. We have, by (1),

$$\begin{aligned} H^*(G/N_G(T);\mathbb{Q}) &\cong H^*(G/T;\mathbb{Q})^W \\ \implies \chi(G/N_G(T)) &= \chi(G/T)/|W| \\ &= 1. \end{aligned}$$

But $G/N_G(T)$ has only homology in even degrees (since $G/T$ does). Thus $G/N_G(T)$ must have the rational homology of a point.

(3) We have a pair of fibrations,

$$BT = EG/T \xrightarrow{W} EG/N_G(T) = BN_G(T) \xrightarrow{G/N_G(T)} BG.$$

The second map has rationally contractible fibre, so the Serre spectral sequence implies $H^*(BG;\mathbb{Q}) = H^*(BN_G(T);\mathbb{Q})$. Applying (1) to the first map, gives $H^*(BN_G(T);\mathbb{Q}) = H^*(BT;\mathbb{Q})^W$.

□

We can apply this proposition to our special orthogonal groups. We need to split into cases based on parity. Let's begin with the odd case of $BSO(2n+1)$. Let $R_\theta$ be the $2 \times 2$ rotation matrix by angle $\theta$.

Then, the maximal torus of $SO(2n+1)$ is,

$$T = \left\{ \begin{pmatrix} R_{\theta_1} & & & \\ & \ddots & & \\ & & R_{\theta_n} & \\ & & & 1 \end{pmatrix} : \theta_1, \dots \theta_n \in S^1 \right\}.$$

The Weyl group consists of transformations permuting the $2\times 2$ blocks and possibly changing some by a sign. This is the symmetry group of an $n$-dimensional cube $W = S_n \rtimes \mathbb{Z}_2^n$. By our above proposition, we know that,

$$H^*(BSO(2n+1);\mathbb{Q}) = H^*(BT;\mathbb{Q})^W = \mathbb{Q}[x_1,\dots,x_n]^W,$$

for $|x_i| = 2i$. Note that $W$ acts on $\mathbb{Z}[x_1,\dots,x_n]$ by permuting the $x_i$ and changing them by a sign. In order that a polynomial in $\mathbb{Z}[x_1,\dots,x_n]$ be invariant under a change of any $x_i$ by a sign it must be a polynomial in the squares $x_1^2,\dots,x_n^2$. Hence, $H^*(BSO(2n+1);\mathbb{Q})$ is the algebra of symmetric rational polynomials in $x_1^2,\dots,x_n^2$. Let $p_i$ denote the $i$th symmetric polynomial function of the $x_i$, so that,

$$(1+x_1^2)\cdots(1+x_n^2) = 1 + p_1 + \cdots + p_n,$$

where $p_i$ is the homogeneous part of degree $4i$. Hence,

$$H^*(BSO(2n+1);\mathbb{Q}) = \mathbb{Q}[p_1,\dots,p_n] \quad \text{where} \quad |p_i| = 4i.$$

Now we deal with the even case $BSO(2n)$. The maximal torus is then,

$$T = \left\{ \begin{pmatrix} R_{\theta_1} & & \\ & \ddots & \\ & & R_{\theta_n} \end{pmatrix} : \theta_1, \dots \theta_n \in S^1 \right\}.$$

The Weyl group $W$ consists of transformations permuting $2\times 2$ blocks with sign changes as before, with the new restriction that the total number of sign changes must be even. Let $P$ be the convex hull of $2^{n-1}$ vertices of an $n$-dimensional cubes which pairwise do not share an edge. $W$ is the symmetry group of $P$, isomorphic to $S_n \rtimes \mathbb{Z}_2^{n-1}$. Again,

$$H^*(BSO(2n);\mathbb{Q}) = H^*(BT;\mathbb{Q})^W = \mathbb{Q}[x_1,\dots,x_n]^W,$$

for $|x_i| = 2i$. All the symmetric polynomials in the $x_i^2$ will be in the cohomology as before. Additionally, since any element of $W$ acts by an even number of sign changes on the $x_i$, the element $x_1\cdots x_n$ is also in the cohomology. By a dimension count, this generator along with the symmetric polynomials gives the entire cohomology. Let $p_i$

denote the $i$th elementary polynomial in $x_i$'s as in the odd case. Let $e = x_1 \cdots x_n$. We conclude,

$$H^*(BSO(2n); \mathbb{Q}) = \mathbb{Q}[p_1, \ldots, p_n, e]/(p_n - e^2) \quad \text{where} \quad |p_i| = 4i,\ |e| = 2n.$$

For even and odd $SO$ groups, the element $p_i$ is called the *ith universal Pontryagin class* and for an oriented vector bundle $E$ it corresponds to a characteristic class $p_i(E)$ called *ith Pontryagin class of $E$.* The element $e$ in the cohomology of $BSO(2n)$ is called the *universal Euler class.*

Note we have already defined the Euler class for an oriented bundle as an element of the integral cohomology class. We will later show that this Euler class is the old Euler class reduced over $\mathbb{Q}$. The fact there is no Euler class in the rational cohomology of $BSO(2n+1)$ tells us the Euler class is torsion for odd rank bundles (in fact we will show later it is always 2-torsion in odd dimensions).

Note also that we have defined Pontryagin classes as integral classes of quaternionic bundles and rational classes of oriented real bundles. To distinguish the real and quaternionic cases, we will call the Pontryagin classes for $BSp$ *quaternionic Pontryagin classes.* They are sometimes also called *symplectic classes.*

To make matters more confusing, the usual definition of Pontryagin classes is as integral classes of any bundle. More precisely, given a real vector bundle $E \to X$, we define the *ith Pontryagin class of $E$* to be $p_i(E) = (-1)^i c_{2i}(\mathbb{C}E) \in H^{4i}(X; \mathbb{Z})$, where $\mathbb{C}E$ denotes the complexification of the real bundle $E$. From a classifying space perspective, there is a map $O(n) \to U(n)$ sending an $n \times n$ orthogonal matrix $A$ to the same matrix thought of as unitary (i.e. the complexification of the linear map). This gives a map $f : BO(n) \to BU(n)$. The pullback of the universal Chern class $c_{2i}$ by $f$ multiplied by $(-1)^i$ defines an element of the $4i$th cohomology group of $BO(n)$ which will be called the $i$th universal Pontryagin class $p_i$ and corresponds to the $i$th Pontryagin class of real bundles.

We will show that under pullback by the double cover map $BSO(n) \to BO(n)$ induced by inclusion the class $p_i$ reduced over $\mathbb{Q}$ will coincide with the universal Pontryagin class for oriented bundles defined above. Our argument above show that modulo torsion, the Pontryagin classes and Euler class generate the homology of $BSO(n)$ (note this is not true over $\mathbb{Z}$, see the remark below).

We summarize our findings in the following definition.

**Definition 6.22: Characteristic Classes**

Let $E \to X$ be a real rank $n$ vector bundle. To it we can associate,

$$\textit{Stiefel–Whitney classes:} \quad w_i(E) \in H^i(X;\mathbb{Z}_2),$$
$$\textit{Pontryagin classes:} \quad p_i(E) \in H^{4i}(X;\mathbb{Z}).$$

If $E$ is oriented we can also associate the

$$\textit{Euler class:} \quad e(E) \in H^n(X;\mathbb{Z}).$$

If $E$ is a complex bundle we furthermore can associate,

$$\textit{Chern classes:} \quad c_i(E) \in H^{2i}(X;\mathbb{Z}_2).$$

Finally, if $E$ is a quaternionic bundle we associate,

$$\textit{quaternionic Pontryagin classes:} \quad p_i(E) \in H^{4i}(X;\mathbb{Z}).$$

Remark 6.23. One may reasonably complain that for real and real oriented bundles we only classified characteristic classes over $\mathbb{Z}_2$ and $\mathbb{Q}$ respectively. In fact, there is no real content in passing to working over $\mathbb{Z}$. Firstly, our description of the generators of $H^*(BSO(n);\mathbb{Q})$ continues to hold if we pass to $\mathbb{Z}[1/2]$. Also,

$$H^*(BSO(n);\mathbb{Z}_2) = \mathbb{Z}_2[w_2,\dots,w_n],$$

(we will see later that $w_1$ vanishes iff a bundle is orientable).

Let $\beta : H^i(X;\mathbb{Z}_2) \to H^{i+1}(X;\mathbb{Z})$ denote the Bockstein homomorphism associated to $\mathbb{Z} \xrightarrow{\times 2} \mathbb{Z} \to \mathbb{Z}_2$. We define the $(i+1)$*st integral Stiefel–Whitney class* of a real bundle $E \to X$ as,

$$W_{i+1}(E) = \beta(w_i(E)) \in H^{i+1}(X;\mathbb{Z}).$$

These characteristic classes correspond to universal elements $W_i \in H^i(BO;\mathbb{Z})$ obtained by applying $\beta$ to the cohomology of $BO$. More generally, given a product $w_{i_1}\cdots w_{i_k}$ living in $H^*(BO(n);\mathbb{Z}_2)$, we can apply $\beta$ to it to obtain an integral cohomology element. Under pullback by the map induced by inclusion $BSO \to BO$, the images under $\beta$ of universal elements may also be thought of as classes in the integral homology of $BSO$.

It happens that $H^*(BSO(n);\mathbb{Z})$ is generated by the universal Pontryagin classes $p_i$ for $i \le (n-1)/2$, the universal Euler class $e$, and the Bockstein homomorphism of a product of even degree universal Stiefel–Whitney classes: $\beta(w_{2i_1}\cdots w_{2i_k})$ for $0 < i_1 < \cdots < i_k \le (n-1)/2$. Similarly $H^*(BO(x);\mathbb{Z})$ is generated by the universal Pontryagin classes $p_i$ for $i \le n/2$, and the Bockstein homomorphism of a product of even degree universal Stiefel–Whitney classes possibly along with $w_1$: $\beta(w_1^{\varepsilon}w_{2i_1}\cdots w_{2i_k})$ for $\varepsilon \in \{0,1\}$,

$0 < i_1 < \cdots < i_k \leq n/2$. These algebra are not free on the given generators and a complete set of relations can be given explicitly, we will see some of the relations later.

We will not need these facts and so not prove them. The arguments are not so complicated and can be found in a short paper of Brown Jr. [**Brown**]. For our purposes, it is enough to say that if one understands the relevant elements in Definition 6.22 and one can describe the cohomology ring and Bockstein homomorphism $\beta$ of the base, then one knows everything about the characteristic classes of the vector bundle. △

To make computations with characteristic classes it is highly useful to know their algebraic and geometric properties. In fact one often introduces the classes axiomatically as uniquely satisfying some list of properties, as in [**MS**]. We will discuss these properties/axioms in due time, as well as geometric constructions of the characteristic classes as first obstructions to extending sections. Quickly, we discuss one important property shared by most of the characteristic classes.

We say a characteristic class $c$ for $G$-bundles is *stable* if $c(V \oplus T) = c(V)$ for any vector bundle $V$ and any trivial rank one bundle $T$ in the $G$-bundle sense. Note that the maps induced by inclusion $BG(n) \hookrightarrow BG(n+1)$, for $G = O, SO, U, Sp$ induce a map from rank $n$ $G$-bundles to rank $(n+1)$ $G$-bundles given by summing with a rank one trivial $G$-bundle. Thus for a characteristic class $c \in H^*(BG(n))$ to be stable is the same as for it to be a pullback of an element in $c \in H^*(BG(\infty))$.

Note that our description of the cohomology of $BO(n), BSO(n), BU(n)$ and $BSp(n)$ extends easily to the classifying spaces of the corresponding infinite dimensional groups $BO, BSO, BU, BSp$. And the pullback maps on cohomology are the given by inclusions (since we know the Schubert cell decomposition respects these inclusions). Thus we have proved the following corollary.

PROPOSITION 6.24. *The Stiefel–Whitney, Pontryagin, Chern, and quaternionic Pontryagin classes are all stable.*

In contrast, recall from Corollary 4.51 that the Euler class is unstable. This should be unsurprising: we know if $E \to X$ is rank $n$ that $e(E) \in H^n(X)$ and $e(E \oplus \underline{\mathbb{R}}) \in H^{n+1}(X)$, so these two cannot be the same whenever $e(E)$ is non-zero. This should not necessarily be seen as a deficit of the Euler class; it allows us to distinguish non-isomorphic but stably isomorphic bundles. For example $TS^2$ is stably trivial after direct summing with the trivial normal

bundle of the embedding $S^2 \subset \mathbb{R}^3$. But $TS^2$ is non-trivial as we saw before because $2 = \chi(S^2) = \langle e(TS^2), [S^2] \rangle$.

**6.3. Characteristic Classes and First Obstructions.** Let us recall from our discussion of obstruction theory that if one has a homotopically simple bundle $E \to B$ over a CW base with $k$-connected fibre $F$, then one obtains a unique cohomology class $C(E) \in H^{k+2}(B; \pi_k(F))$ representing any attempt to extend a section of the $(k+1)$-skeleton of $B$ to the $(k+2)$-skeleton. This class is natural with respect to pullbacks, hence a functorial association of a vector bundle $\xi$ on $B$ to such a fibre bundle defines a characteristic class for $\xi$.

For example, we saw before that to an oriented rank $n$ vector bundle $\xi$ on $B$, we can associate the first obstruction $C(S\xi) \in H^n(B)$ of the sphere bundle $S\xi$. We called this the Euler class $e(\xi)$ of $\xi$ and showed it can be recovered as the restriction of the Thom class $t_\xi$ to the base, or equivalently in the smooth case as Poincaré dual to the fundamental class of the zero set of a generic section of $\xi$. We will show below that this agrees with the Euler class as defined above in the cohomology of $BSO(n)$.

*Defining Stiefel–Whitney and Chern Classes as Obstructions.* Let $\xi \to B$ be a complex vector bundle of rank $n$. To it, we can associate the bundle $F_{n-m+1}(\xi)$ of $(n-m+1)$ frames of $\xi$ which is a fibre bundle with fibre $\mathbb{C}V(n, n-m+1)$. Consider the sequence of fibrations for $n > m$,

$$\begin{aligned}
\mathbb{C}V(n-1, n-m) &\hookrightarrow \mathbb{C}V(n, n-m+1) \to S^{2n-1} \\
\mathbb{C}V(n-2, n-m-1) &\hookrightarrow \mathbb{C}V(n-1, n-m) \to S^{2n-3} \\
&\vdots \\
\mathbb{C}V(m, 1) &\hookrightarrow \mathbb{C}V(m+1, 2) \qquad \to S^{2m+1}.
\end{aligned}$$

Note that $\mathbb{C}V(m, 1) = S^{2m-1}$. Recursively applying the homotopy long exact sequence to these fibrations we conclude that $\mathbb{C}V(n, n-m+1)$ is $(2m-2)$-connected and that $\pi_{2m-1}(\mathbb{C}V(n-m+1)) = \mathbb{Z}$. We thus obtain a first obstruction to the existence of a $(n-m+1)$-frame of $\xi$ over the $2m$-skeleton of $B$,

$$c_m(\xi) := C(F_{n-m+1}(\xi)) \in H^{2m}(B; \mathbb{Z}).$$

We show soon this agrees with the $m$th Chern class.

We point out two special cases. For $m = 1$, we are studying $n$-frames, i.e. bases of the fibre. So, $c_1(\xi)$ measures the obstruction to the existence of a trivialization of the bundle over the 2-skeleton. From this we deduce for example that any bundle $\xi$ on a 2 dimensional complex with $c_1(\xi) = 0$ is trivial.

For $m = n$, we are studying 1-frames, i.e. non-vanishing sections. So, $c_n(\xi)$ measures the obstruction to the existence of a non-vanishing section of $\xi$ over the $n$-skeleton. But this obstruction class is already familiar to us, which implies the following important corollary.

**Theorem 6.25**

Let $\xi \to X$ be a complex rank $n$ vector bundle, then $c_n(\xi) = e(\xi_{\mathbb{R}})$, where $\xi_{\mathbb{R}}$ denotes the underlying oriented real bundle.

Note that if we are dealing with a line bundle $\xi$, there is one Chern class $c_1(\xi)$ measuring the first obstruction to defining a section of $\xi$ over the two skeleton of the base. But note since the structure group of $\xi$ is $U(1)$, which has trivial higher homotopy groups, there are no further obstructions. Thus a trivialization over the two-skeleton extends to one on the entire base. And hence $c_1(\xi) = 0$ implies the bundle trivializes. This reproduces our earlier result that complex line bundles are classified by $c_1$.

The case for quaternionic bundles is exactly analogous. We redefine $p_m(\xi)$ for a quaternionic rank $n$ bundle $\xi \to B$ as the first obstruction to the existence of a quaternionic $(n - m + 1)$-frame of $\xi$ over the $2m$-skeleton of $B$.

Now consider $\xi \to B$ a rank $n$ real vector bundle. To it we associate the bundle $F_{n-m+1}(\xi)$ of $n - m + 1$ frame of $\xi$ which is a fibre bundle with fibre $V(n, n - m + 1)$. Aw with the complex case, we have a sequence of fibrations,

$$\begin{aligned} V(n-1, n-m) &\hookrightarrow V(n, n-m+1) \to S^{n-1} \\ V(n-2, n-m-1) &\hookrightarrow V(n-1, n-m) \to S^{n-2} \\ &\vdots \\ V(m, 1) &\hookrightarrow V(m+1, 2) \to S^m. \end{aligned}$$

Note that $V(m, 1) = S^{m-1}$. We conclude recursively from the homotopy long exact sequences of the fibrations that $V(n, n - m + 1)$ is $m - 2$ connected. Furthermore from the last exact sequence, which is a Gysin sequence,

$$0 \to \pi_m(S^m) = \mathbb{Z} \xrightarrow{\times \chi(S^m)} \mathbb{Z} = \pi_{m-1}(V(m, 1)) \to \pi_{m-1}(V(m+1, 2)) \to 0 = \pi_{m-1}(S^m).$$

So we conclude,

$$\pi_{m-1}(V(n, n-m+1)) = \begin{cases} \mathbb{Z}_2 & m < n \text{ even or } m = 1, \\ \mathbb{Z} & m > 1 \text{ odd or } m = n. \end{cases}$$

The bundle $F_{n-m+1}$ may not be homotopically simple, and the obstruction class may have twisted coefficients whenever $\pi_{m-1} = \mathbb{Z}$, but we will be okay

if we work over $\mathbb{Z}_2$. We thus obtain a first obstruction to the existence of a $(n-m+1)$ frame of $\xi$ over the $m$-skeleton of $B$,

$$w_m(\xi) := C(F_{n-m+1}(\xi)) \in H^m(B; \mathbb{Z}_2).$$

We will show this agrees with the $m$th Stiefel–Whitney class.

For $m = n$, $w_n(\xi)$ is the first obstruction to a 1-frame over the $n$-skeleton of $B$. Hence we have the following as with the complex case.

PROPOSITION 6.26. *Let $\xi \to X$ be a real rank $n$ vector bundle, then $w_n(\xi) = e(\xi) \pmod 2$. Here we define $e(\xi)$ in the unoriented case as the restriction of the mod 2 Thom class to the base $X$.*

For $m = 1$, $w_1(\xi)$ is the first obstruction to an $n$-frame, i.e. a trivialization, over the 1-skeleton of $B$.

*Proving Equivalence.* Let us prove that the obstruction theoretic definitions of our characteristic classes agree with the previous cohomological definitions.

Note that by naturality of first obstructions it suffices to show that the obstruction class and cohomology class agree for the universal bundle $EG \to BG$. Actually to make things simpler we will use a version of *the splitting principle*, which we explain in generality later, that allows us to reduce to a sum of tautological bundles on $BT$, for $T$ the maximal torus of $G$.

To see what we mean, let's begin with the Stiefel–Whitney classes. Recall we had a fibration $Bi : (\mathbb{R}P^\infty)^n \to BO(n)$ induced by the inclusion $i : \mathbb{Z}_2^n \to O(n)$. The pullback of the universal bundle $EO(n)$ by $Bi$ can be seen to be a direct sum of copies of the tautological line bundle $\xi^1$ on each copy of $\mathbb{R}P^\infty$. This is because $(Bi)^*EO(n)$ is the fibration $V(\infty, n) \to (\mathbb{R}P^\infty)^n$ sending an orthonormal $n$-frame $(v_1, \ldots, v_n)$ to the lines $\operatorname{span}(v_1), \ldots, \operatorname{span}(v_n) \in (\mathbb{R}P^\infty)^n$.

As we showed before $H^*((\mathbb{R}P^\infty)^n; \mathbb{Z}_2) = \mathbb{Z}_2[x_1, \ldots, x_n]$, $H^*(BO(n); \mathbb{Z}_2) = \mathbb{Z}_2[w_1, \ldots, w_n]$ and the pullback $(Bi)^*$ is injective on cohomology sending $w_i$ to the $i$th symmetric polynomial in the $x_i$. Because this pullback is injective if the obstruction class definition of $w_i(\xi^1 \oplus \cdots \oplus \xi^1)$ agrees with the cohomological class $\sigma_i(x_1, \ldots, x_n)$, then we will know that the obstruction theory and cohomological definitions of Stiefel–Whitney classes agree.

To understand $w_m(\xi^1 \oplus \cdots \oplus \xi^1)$, lets evaluate it on an element of homology $[\mathbb{R}P^{\ell_1} \times \cdots \times \mathbb{R}P^{\ell_n}]$ with $\sum \ell_i = m$. We claim that,

$$\left\langle w_m(\xi^1 \oplus \cdots \oplus \xi^1), [\mathbb{R}P^{\ell_1} \times \cdots \times \mathbb{R}P^{\ell_n}] \right\rangle = \begin{cases} 1 \pmod 2 & \ell_i \leq 1 \ \forall i, \\ 0 & \text{else.} \end{cases}$$

To see this note if some $\ell_i > 1$, because $\sum \ell_j = m$, there must be at least $n - m + 1$ of the $\ell_j$ which are zero. But the direct summand of the bundle $\xi^1 \oplus \cdots \oplus \xi^1$ trivializes on $[\mathbb{R}P^{\ell_1} \times \cdots \times \mathbb{R}P^{\ell_n}]$ for each factor with $\ell_j = 0$. In particular if some $\ell_i > 1$, we may find an $(n - m + 1)$ frame of the bundle on $[\mathbb{R}P^{\ell_1} \times \cdots \times \mathbb{R}P^{\ell_n}]$ and hence the obstruction $w_m(\xi^1 \oplus \cdots \oplus \xi^1)$ must vanish on this class, as claimed. Conversely, if $\ell_i \leq 1$ for all $i$, then the bundle will look like $(\xi^1)^{\oplus m} \oplus (\underline{\mathbb{R}})^{\oplus(n-m)}$ on $[(\mathbb{R}P^1)^m \times \{\mathrm{pt}\}^{n-m}]$. Note $w_m((\xi^1)^{\oplus m}) = e(\xi^{\oplus m}) = 1 \pmod 2$ since this is a direct sum of copies of Möbius strips. Thus $(\xi^1)^{\oplus m}$ has no non-vanishing section and $(\xi^1)^{\oplus m} \oplus (\underline{\mathbb{R}})^{\oplus(n-m)}$ does not have a $(n - m + 1)$ frame on $[(\mathbb{R}P^1)^m \times \{\mathrm{pt}\}^{n-m}]$ and so the obstruction class must be the non-trivial element of $\mathbb{Z}_2$ on this class.

Note that the cohomology of $(\mathbb{R}P^\infty)^n$ is given by,

$$\langle x_1^{i_1} \cdots x_n^{i_n}, [\mathbb{R}P^{\ell_1} \times \cdots \times \mathbb{R}P^{\ell_n}] \rangle = \delta_{i_1,\ell_1} \cdots \delta_{i_n,\ell_n} \pmod 2.$$

Hence we have,

$$w_m(\xi^1 \oplus \cdots \oplus \xi^1) = \sum_{i_1 < \cdots < i_m} x_{i_1} \cdots x_{i_m} = \sigma_m(x_1, \ldots, x_n).$$

Hence the obstruction theoretic Stiefel–Whitney class agrees with the cohomological definition.

Now we move onto the complex/Chern case. As with above, we have a fibration $Bi : (\mathbb{C}P^\infty)^n \to BU(n)$ induced by the inclusion $i : T^n \to U(n)$. The pullback of the universal bundle $EU(n)$ by $Bi$ is the direct sum of copies of the tautological complex line bundle $\xi^1$ on each copy of $\mathbb{C}P^\infty$. This is because $(Bi)^* EU(n)$ is the fibration $\mathbb{C}V(\infty, n) \to (\mathbb{C}P^\infty)^n$ sending an orthonormal $n$-frame $(v_1, \ldots, v_n)$ to the complex lines $\mathrm{span}(v_1), \ldots, \mathrm{span}(v_n) \in (\mathbb{C}P^\infty)^n$. Because the induced map $(Bi)^*$ is injective on cohomology, it will suffice to show $c_i(\xi^1 \oplus \cdots \oplus \xi^1)$ defined obstruction theoretically agrees with $\sigma_i(y_1, \ldots, y_n)$ where $H^*(BU(n)) = \mathbb{Z}[y_1, \ldots, y_n]$. We claim for $\sum \ell_j = m$,

$$\left\langle c_m(\xi^1 \oplus \cdots \oplus \xi^1), [\mathbb{C}P^{\ell_1} \times \cdots \times \mathbb{C}P^{\ell_n}] \right\rangle = \begin{cases} (-1)^m & \ell_i \leq 1 \ \forall i, \\ 0 & \text{else.} \end{cases}$$

To see this, if some $\ell_i > 1$, then as before at least $n - m + 1$ of the $\ell_j$ must be zero. Then the direct summand of the bundle $\xi^1 \oplus \cdots \oplus \xi^1$ trivializes on $[\mathbb{C}P^{\ell_1} \times \cdots \times \mathbb{C}P^{\ell_n}]$ for each factor with $\ell_j = 0$ and we may find an $(n-m+1)$ frame of the bundle on $[\mathbb{C}P^{\ell_1} \times \cdots \times \mathbb{C}P^{\ell_n}]$ and hence the obstruction $c_m(\xi^1 \oplus \cdots \oplus \xi^1)$ must vanish on this class.

If all $\ell_i \leq 1$, then the bundle will look like $(\xi^1)^{\oplus m} \oplus (\underline{\mathbb{C}})^{\oplus(n-m)}$ on $[(\mathbb{C}P^1)^m \times \{\mathrm{pt}\}^{n-m}]$. Note that $\xi^1 \to \mathbb{C}P^1$ is the degree $-1$ bundle so that $c_1(\xi^1) = -1$ (the fact that $c_1$ defined cohomologically and obstruction-theoretically for a

line bundle on $S^2$ can be prove easily directly). If one takes a direct sum of these bundles, the obstruction class will multiply so that

$$\begin{aligned}\left\langle c_m((\xi^1)^{\oplus m} \oplus (\underline{\mathbb{C}})^{\oplus(n-m)}), [(\mathbb{C}P^1)^m \times \{\text{pt}\}^{n-m}]\right\rangle &= \left\langle c_m((\xi^1)^{\oplus m}), [(\mathbb{C}P^1)^m]\right\rangle \\ &= \left\langle c_1(\xi^1), [\mathbb{C}P^1]\right\rangle^m \\ &= (-1)^m.\end{aligned}$$

Hence as for the real case we have,

$$c_m(\xi^1 \oplus \cdots \oplus \xi^1) = \sum_{i_1<\cdots<i_m} y_{i_1} \cdots y_{i_m} = \sigma_m(y_1, \ldots, y_n).$$

So the obstruction theory and cohomological Chern classes agree. The quaternionic case is identical and we omit it.

Lastly, we should check that the Pontryagin classes as defined cohomologically in $H^*(BSO(n); \mathbb{Q})$ agree with their definition as $p_i(\xi) = (-1)^i c_{2i}(\mathbb{C}\xi)$.

There is a pullback map $H^*(BSO(n)) \to H^*(BT^{\lfloor n/2 \rfloor}) = \mathbb{Z}[x_1, \ldots, x_{\lfloor n/2 \rfloor}]$ which is injective. Using our splitting principle it is enough to show the universal class $p_i$ pulled back to $BT^{\lfloor n/2 \rfloor}$ is $\sigma_i(x_1^2, \ldots, x_i^2)$ and we should also show the universal Euler class $e$ pulled back to $BT^{\lfloor n/2 \rfloor}$ is $x_1 \cdots x_n$.

The above pullback corresponds to a map $BT^{\lfloor n/2 \rfloor} \to BSO(n)$ induced by inclusion of the maximal torus of $SO(n)$. The pullback of the universal bundle $ESO(n)$ under this map to a bundle on $(\mathbb{C}P^\infty)^{\lfloor n/2 \rfloor}$ is a direct sum of $\lfloor n/2 \rfloor$ oriented 2-plane bundles given as $\xi^1_{\mathbb{R}}$ for $\xi^1$ the tautological complex line bundle on $\mathbb{C}P^\infty$. Note that the complexification of $\xi^1_{\mathbb{R}}$ is $\xi^1 \oplus (\xi^1)^*$, where $V^*$ denotes the complex dual bundle (in general $\mathbb{C}(V_{\mathbb{R}}) = V \oplus V^*$ for any complex bundle $V$). Hence we compute,

$$p_m((\xi^1_{\mathbb{R}})^{\oplus n}) = (-1)^m[c((\xi^1 \oplus (\xi^1)^*) \oplus \cdots \oplus (\xi^1 \oplus (\xi^1)^*))]_{2m},$$

where $[\cdot]_{2m}$ denotes the part of degree $2m$. As we will show, total Chern classes are multiplicative under direct summing,

$$= (-1)^m[c(\xi^1)c((\xi^1)^*) \cdots c(\xi^1)c((\xi^1)^*)]_{2m}.$$

We have suppressed that each $\xi^1$ comes from a different copy of $\mathbb{C}P^\infty$ in the base. By above, we know $\xi^1$ over the $i$th $\mathbb{C}P^\infty$ will have $c(\xi^1) = 1 + x_i$. Since taking the complex dual multiplies the degree by minus one, $c((\xi^1)^*) = 1 - x_i$. Hence,

$$\begin{aligned}&= (-1)^m[(1 - x_1^2) \cdots (1 - x_n^2)]_{2m} \\ &= \sigma_m(x_1^2, \ldots, x_n^2).\end{aligned}$$

The fact our result holds for this bundle on the maximal torus implies it for every other bundle. For a complex rank $n$ bundle, we know $e = c_{2n}$. We immediately see that

$$e((\xi^1)^{\oplus n}) = c_n((\xi^1)^{\oplus n}) = \sigma_n(x_1, \ldots, x_n) = x_1 \ldots x_n.$$

Again by our splitting principle, the Euler class is as claimed.

*Some Geometric Interpretations of Specific Characteristic Classes.* The presentation of characteristic classes as geometric obstructions is an important perspective on the meaning and utility of them for classification problems. We now present some specific examples of how characteristic classes can be interpreted geometrically.

**Theorem 6.27**

Given a real rank $n$ bundle $\xi$ and a complex rank $n$ bundle $\eta$, one has

$$w_1(\xi) = w_1(\det \xi) \quad \text{and} \quad c_1(\eta) = c_1(\det_{\mathbb{C}} \eta),$$

where $\det \xi = \Lambda^n \xi$ and $\det_{\mathbb{C}} \eta = \Lambda^n_{\mathbb{C}} \eta$. In particular, $w_1(\xi)$ vanishes if and only if $\xi$ is orientable (reduces to structure group $SO(n)$) and $c_1$ vanishes if and only if $\eta$ reduces to structure group $SU(n)$.

Proof. We know $w_1(\xi)$ is the first obstruction for a section of the orthogonal $n$-frame bundle of $\xi$. But note this $n$-frame bundle is isomorphic to $\det \xi$ by sending a frame $(v_1, \ldots, v_n)$ to $v_1 \wedge \cdots \wedge v_n$. So $w_1(\xi)$ is the first obstruction to finding a section of the line bundle $\det \xi$. But that is nothing but $e(\det \xi)$ (mod 2) which equals $w_1(\det \xi)$. The same analysis holds for $c_1$ noting that the unitary $n$ frame bundle of $\eta$ is isomorphic to $\det_{\mathbb{C}} \eta$.

We have $w_1(\xi)$ vanishes if and only if $w_1(\det \xi) = 0$, if and only if $\det \xi$ is trivial, if and only if a global orientation of $\xi$ exists. The same analysis holds for $c_1$.

Note one can also see the orientation condition more directly, as we proved many eons ago. An orientation is a reduction of structure group via a lift of classifying map $X \to BO(n)$ for the map $BSO(n) \to BO(n)$. But this fits into a fibration $\mathbb{Z}_2 \to BSO(n) \to BO(n)$. So by obstruction theory, since $\mathbb{Z}_2$ has no higher homotopy groups, the first obstruction to a lift (which is $w_1(\xi)$) is the only obstruction. And $w_1(\xi) = 0$ if and only if such a lift exists. The same goes for the complex case vis à vis the fibration $U(1) \to BSU(n) \to BU(n)$. □

Proposition 6.28. *If a rank $n$ real bundle $\xi$ admits $m$ linearly independent sections, then $w_n(\xi) = w_{n-1}(\xi) = \cdots = w_{n-m+1}(\xi) = 0$. If a rank $n$ complex*

*bundle $\eta$ admits $m$ linearly independent sections, then $w_n(\eta) = w_{n-1}(\eta) = \cdots = w_{n-m+1}(\eta) = 0$.*

PROOF. A collection of $m$ linearly independent sections provides an $m$-frame of $\xi$ over the total base space. And so the obstruction classes related to $1, 2, \cdots, m$-frames, i.e. $w_n, w_{n-1}, \ldots, w_{n-m+1}$, must all vanish. The same holds for the complex case. □

Recall a spin structure on an oriented bundle $E \to X$ given by a classifying map $X \to BSO(n)$ is given by a lift to $X \to BSpin(n)$. A spin$^{\mathbf{c}}$ structure is a lift to $X \to BSpin^{\mathbf{c}}(n)$.

**Theorem 6.29**

A rank $n$ oriented bundle $E \to X$ has a spin structure if and only if $w_2(E) = 0$. If $E$ is spin, the set of spin structures is an affine space modelled on $H^1(X; \mathbb{Z}_2)$.

PROOF. We have a fibration $\mathbb{Z}_2 \hookrightarrow BSpin(n) \to BSO(n)$ and we are looking for a lift $X \to BSO(n)$ to $X \to BSpin(n)$. Since the higher homotopy groups of $\mathbb{Z}_2$ are trivial, the only obstruction to a lift is the first obstruction, which is an element of $H^2(X; \mathbb{Z}_2)$. This obstruction must define a characteristic class (non-trivial since there are non-spin bundles) and so must be $w_2$. Hence a lift exists iff $w_2(E) = 0$.

The different possible lifts to a map $X \to BSpin(n)$ are parameterized by difference cochains. Again looking at the homotopy groups of $\mathbb{Z}_2$, the only non-vanishing difference cochain lives in $H^1(X; \mathbb{Z}_2)$ checking if two maps $X \to BSpin(n)$ are homotopic relative to $X \to BSO(n)$ on the 1-skeleton. □

REMARK 6.30. It turns out that a 7 dimensional manifold admits a $G_2$ structure if and only if it is spin. Hence the same obstructions apply to $G_2$ structures. △

**Theorem 6.31**

A rank $n$ oriented bundle $E \to X$ has a spin$^{\mathbf{c}}$ structure if and only if $w_2(E)$ is the reduction mod 2 of an integral class. If $E$ is spin$^{\mathbf{c}}$, the set of spin$^{\mathbf{c}}$ structures is an affine space modelled on $H^2(X; \mathbb{Z}_2)$.

PROOF. We have a fibration $U(1) \hookrightarrow BSpin^{\mathbf{c}}(n) \to BSO(n)$ and we are looking for a lift $X \to BSO(n)$ to $X \to BSpin^{\mathbf{c}}(n)$. Studying the homotopy

groups of $U(1)$, the only obstruction is an element of $H^3(X;\mathbb{Z})$. This must be a characteristic class and so by the classification which we stated but did not prove, it must be of the form $aW_3(E)$ for integer $a$. In fact $W_3(E)$ is 2-torsion since $2\beta(w_2) = \beta(2w_2) = 0$. And since the obstruction class is non-trivial (since there are non-spin$^{\mathbf{c}}$ bundles) we conclude the first and only obstruction is $W_3(E)$. Recall $W_3$ comes from the Bockstein exact sequence,

$$H^2(X;\mathbb{Z}) \xrightarrow{r} H^2(X;\mathbb{Z}_2) \xrightarrow{\beta} H^3(X;\mathbb{Z}),$$

as $W_3 = \beta(w_2)$. Since this is exact, $W_3 = 0$ if and only $w_2$ is in the image of $r$, i.e. the reduction mod 2 of an integral class.

Remark 6.32. One way to interpret this is that $w_2 \in H^2(X;\mathbb{Z}_2)$ measures the obstruction to being a spin bundle. And we can obtain spin$^{\mathbf{c}}$ bundle if and only if there is a complex line bundle $\gamma$ with first Chern class $c_1(\gamma) \in H^2(X;\mathbb{Z})$ so $r(c_1(\gamma)) = w_2$. So that adding the complex line bundle "cancels out" or mends the failure of local $Spin(n)$ charts to patch together. △

Again analyzing the homotopy of $U(1)$, the different lifts to a map $X \to BSpin^{\mathbf{c}}(n)$ are parameterized by difference cochains, and the only non-vanishing ones live in $H^2(X;\mathbb{Z})$, checking if two maps $X \to BSpin^{\mathbf{c}}(n)$ are homotopic relative to $X \to BSO(n)$ on the 2-skeleton. □

Remark 6.33. One can show for every oriented manifold of dimension less than or equal to four, the second Stiefel–Whitney class of the tangent bundle is the mod 2 reduction of an integral class and so every oriented manifold of dimension four or less is spin$^{\mathbf{c}}$. This is clear in dimensions one and two since $W_3 \in H^3(X) = 0$. If $X$ is three dimensional, then $H^3(X) = 0$ or $\mathbb{Z}$ and has no two-torsion, so again $W_3 = 0$. The case of dimension four is slightly more tricky, proofs can easily be found online. △

Note that $SO(n)$ is the identity component of $O(n)$, and $Spin(n) \to SO(n)$ is the universal covering of $SO(n)$. We can continue this to form the Whitehead tower for $O(n)$,

$$\cdots \to FiveBrane(n) \to String(n) \to Spin(n) \to SO(n) \to O(n).$$

The two connected and three connected coverings of $O(n)$ turn out to be topological groups called the string group and five-brane group respectively. Unsurprisingly, these are of interest in string theory. A lift of a bundle classified by $X \to BO(n)$ via the map $BString(n) \to BO(n)$ is called a *string structure* and a further lift via the map $BFiveBrane(n) \to BO(n)$ is called a *five-brane structure*. It turns out these higher lifts are also classified by characteristic classes. More precisely, a spin bundle $E$ has a string structure if and only if

$\frac{1}{2}p_1(E) = 0$ and a string bundle $E$ has a five-brane structure if and only if $\frac{1}{6}p_2(E) = 0$. One may climb even further up the Whitehead tower similarly if they are so inclined (I'll stick to climbing at Benchmark).

**6.4. Properties of Characteristic Classes and Computations.** We now move towards some important computations of characteristic classes. We begin with one final definition of Stiefel–Whitney classes and show that in certain cases they are in fact homotopical invariants.

*Axiomatic Treatment of Stiefel–Whitney Classes.* Let us see how computations of Stiefel–Whitney classes can be made from a few simple axioms.

**Theorem 6.34**

There exists characteristic classes $w_i(\xi) \in H^i(X; \mathbb{Z}_2)$ for $i \geq 0$ associated to real vector bundles $\xi \to X$ which are uniquely determined by the following properties:

**Dimension:** $w_0(\xi) = 1$ and $w_i(\xi) = 0$ for $i$ greater than the rank of $\xi$.

**Non-Triviality:** If $\gamma^1$ is the tautological line bundle on $\mathbb{R}P^1$ then $w_1(\gamma^1) \neq 0$.

**Whitney Product Formula:** Given bundles $\xi, \eta$, one has,

$$w_i(\xi \oplus \eta) = \sum_{j=0}^{i} w_j(\xi) \smile w_{i-j}(\eta).$$

Proof. We begin with uniqueness; suppose $w_i$ are characteristic classes satisfying the axioms above. Let $H^*(\mathbb{R}P^1; \mathbb{Z}_2) = \mathbb{Z}_2[a]/(a^2)$. We must have by the first two axioms that the total Stiefel–Whitney class of tautological bundle $\gamma^1$ is $w(\gamma^1) = 1 + a$.

Let $\xi^1$ be the tautological line bundle on $\mathbb{R}P^\infty$. Note the standard embedding $i : \mathbb{R}P^1 \to \mathbb{R}P^\infty$ gives $\gamma^1$ as the pullback of $\xi^1$. Since $H^*(\mathbb{R}P^\infty; \mathbb{Z}_2) = \mathbb{Z}_2[b]$ and $i^*(b) = a$, we must have by naturality that $w(\xi^1) = 1 + b$.

Now consider the Cartesian product $\zeta = \xi^1 \times \cdots \times \xi^1$ over $(\mathbb{R}P^\infty)^n$. We have $H^*((\mathbb{R}P^\infty)^n; \mathbb{Z}_2) = \mathbb{Z}_2[b_1, \ldots, b_n]$ so that the $i$th inclusion $j_i : \mathbb{R}P^\infty \to (\mathbb{R}P^\infty)^n$ has $j_i^*(b_k) = \delta_{i,k}\, b$. By the Whitney product formula, and naturality with respect to the inclusion of $\mathbb{R}P^\infty$ factors,

$$w(\zeta) = (1 + b_1) \cdots (1 + b_n).$$

Since the map $Bi : (\mathbb{R}P^\infty)^n \to BO(n)$ induces an injective pullback map in cohomology, we deduce that,

$$w(EO(n)) = 1 + w_1 + \cdots + w_n,$$

where $H^*(BO(n); \mathbb{Z}_2) = \mathbb{Z}_2[w_1, \ldots, w_n]$ and $(Bi)^* w_i = \sigma_i(b_1, \ldots, b_n)$. But if we know the characteristic class on the universal bundle we know it on all other bundles by naturality and the fact maps to $BO(n)$ classifies vector bundles. So $w_i$ are unique.

Now we prove existence. We have already given some equivalent definitions of the Stiefel–Whitney class which we wish to show obey these axioms. The dimension axiom is really a matter of definition (since we only defined $w_i(\xi)$ for $1 \leq i \leq \mathrm{rk}(\xi)$). The non-triviality axiom follows from the fact $\gamma^1$ is the pullback of $EO(1)$ under the inclusion $\mathbb{R}P^1 \hookrightarrow \mathbb{R}P^\infty = BO(1)$. We know what this maps does on cohomology; it will send $w_1$ to the generator of $H^1(\mathbb{R}P^1; \mathbb{Z}_2)$. The Whitney product formula can be deduced as follows. Given bundles $\xi$ and $\eta$ on $X$ classified by maps $X \to BO(n)$ and $X \to BO(m)$, the bundle $\xi \oplus \eta$ is classified by the map $X \to B(O(n) \times O(m)) \to BO(n \times m)$, where the first map combines the above two classifying maps and the second is the classified functor applied to the inclusion $O(n) \times O(m) \to O(n+m)$ of block diagonal matrices. Our usual map $(\mathbb{R}P^\infty)^{n+m} \to BO(n+m)$ must factor through $B(O(n) \times O(m))$ via the map above. It follows immediately from the combinatorics of elementary symmetric functions that the Whitney relations hold between the generators of the homology of $BO(n+m)$, $BO(n)$ and $BO(m)$. By naturality it holds when pulled back to $X$ as well. Thus all the axioms are satisfied and the Stiefel–Whitney classes as previously defined are the unique classes given by the above axioms.

Not contenting ourselves with this, we give a separate definition of the Stiefel–Whitney classes, which will also satisfy the axioms and hence agree with our previous definitions. Given a rank $n$ real vector bundle $\xi \to X$ let $t_\xi \in H^n(\xi, \xi \setminus X; \mathbb{Z}_2)$ be the unoriented Thom class. Recall cupping with the Thom class defines an isomorphism,

$$\Phi : H^i(X; \mathbb{Z}_2) \xrightarrow{\cong} H^{n+i}(\xi, \xi \setminus X; \mathbb{Z}_2).$$

**Definition 6.35**

The *ith Stiefel–Whitney class* of a rank $n$ bundle $\xi \to X$ is given by,

$$w_i(\xi) = \Phi^{-1} Sq^i(t_\xi) = \Phi^{-1} Sq^i \Phi(1).$$

The fact this is really a characteristic class, i.e. it satisfies naturality, is immediate from the fact the maps involved are natural. Let us show the axioms hold.

The dimension axiom is clear from the definition of the Steenrod squares. The Whitney product formula follows from Cartan's identity for Steenrod squares and the fact the Thom isomorphism distributes over a product of bundles. For the non-triviality axiom, note that the total space of $\gamma^1$ is a Möbius strip $M$ which is homeomorphic to $\mathbb{R}P^2 \setminus D^2$ for an embedded disk $D^2$. By excision, $H^i(M, \partial M; \mathbb{Z}_2) \cong H^i(\mathbb{R}P^2, D^2; \mathbb{Z}_2) \cong H^i(\mathbb{R}P^2; \mathbb{Z}_2)$ for $i > 0$. The Thom class $t_{\gamma^1}$ certainly cannot be zero by Thom isomorphism, so under the above identification it must correspond to the generator $a \in H^1(\mathbb{R}P^2; \mathbb{Z}_2)$. But then $Sq^1(a) = a^2$. This is non-zero in $H^2(\mathbb{R}P^2; \mathbb{Z}_2)$. And so $Sq^1(t_{\gamma^1}) \neq 0$. Since $\Phi$ is an isomorphism, $w_1(\gamma^1) = \Phi^{-1} Sq^1(t_{\gamma^1}) \neq 0$, as desired. □

*Wu Formula.* In the above proof we gave a new construction of Stiefel–Whitney classes as $w_i(\xi) = \Phi^{-1} Sq^i \Phi(1)$. We can refine this construction for the case of tangent bundles as follows.

**Definition 6.36**

Let $M$ be a smooth closed $n$-manifold. The *ith Wu class* $\nu_i \in H^i(M; \mathbb{Z}_2)$ is uniquely defined so that,

$$\nu_i \smile x = Sq^i(x) \quad \text{for all} \quad x \in H^{n-i}(M; \mathbb{Z}_2).$$

Note that $Sq^i \in \mathrm{Hom}(H^{n-i}(M; \mathbb{Z}_2), \mathbb{Z}_2) \cong H_{n-i}(M; \mathbb{Z}_2)$. So by Poincaré duality the Wu class is well defined.

With this definition, we obtain the following important result.

**Theorem 6.37: Wu Formula**

For a closed smooth $n$-manifold $M$, the Stiefel–Whitney classes of its tangent bundle are given by

$$w_i(TM) = \sum_{j=0}^{i} Sq^{i-j}(\nu_j).$$

The proof, following [MS, §11], is a little long and will require a couple lemmas. First we need to define a new operation in (co)homology.

**Definition 6.38**

Given two spaces $X, Y$, there is an operation called the *slant product* defined as the composite,

$$/ : H^*(X \times Y) \otimes H_*(Y) \xrightarrow{\text{Künneth}} H^*(X) \otimes H^*(Y) \otimes H_*(Y) \xrightarrow{\mathrm{id} \otimes \langle , \rangle} H^*(X).$$

This satisfies the identities $(a \times b)/\beta = a\langle b, \beta\rangle$ and $[(a \times 1) \smile p]/\beta = a \smile (p/\beta)$.

Let $t_\Delta$ denote the $\mathbb{Z}_2$ Thom class of the normal bundle to the diagonal embedding $\Delta : M \to M^2$. By excision and then restriction we obtain from $t_\Delta$ a class $u \in H^n(M \times M; \mathbb{Z}_2)$. It follows from local computations on a slice $\{x\} \times M$ that $u/[M] = 1$ and that $(a \times 1) \smile u = (1 \times a) \smile u$.

LEMMA 6.39. *For any homogeneous basis $b_1, \ldots, b_r$ of $H^*(M; \mathbb{Z}_2)$ there is a dual basis $c_1, \ldots, c_r$ so that $\langle b_i \smile c_j\rangle = \delta_{i,j}$. In this case, one has,*

$$u = \sum_i (-1)^{|b_i|} b_i \times c_i.$$

PROOF. By the Künneth formula we know that $u$ is a sum of products $u = \sum_i b_i \times e_i$ for $b_i, e_i \in H^*(M; \mathbb{Z}_2)$ of complimentary dimensions. We compute for arbitrary $a$,

$$\begin{aligned}[(1 \times a) \smile u]/[M] &= [(a \times 1) \smile u]/[M] \\ &= a \smile u/[M] \\ &= a.\end{aligned}$$

We can plug the expansion of $u$ in the product basis into the original expression,

$$\begin{aligned}a &= \sum_j (-1)^{|a|\cdot|b_j|}(b_j \times (a \smile e_j))/[M] \\ &= \sum_j (-1)^{|a|\cdot|b_j|} b_j \langle a \smile e_j, [M]\rangle\end{aligned}$$

Taking $b_i = a$, we conclude that $c_i = (-1)^{|b_i|} e_i$ satisfies the lemma. □

LEMMA 6.40. *For a closed smooth $n$-manifold $M$, one has,*

$$w_i(TM) = Sq^i(u)/[M].$$

PROOF. Note that the Stiefel–Whitney class $w_i(TM)$ is equivalently the $i$th Stiefel–Whitney class of the normal bundle to the embedding $\Delta : M \to M \times M$. From our formula for Stiefel–Whitney classes we have,

$$w_i(TM) \smile t_\Delta = Sq^i(t_\Delta).$$

Following through the isomorphism to $H^*(M^2, M^2 \setminus \Delta; \mathbb{Z}_2)$ and then the restriction to $H^*(M^2; \mathbb{Z}_2)$, this identity becomes,

$$(w_i(TM) \times 1) \smile u = Sq^i(u).$$

Taking the slant product with $[M]$, this becomes,

$$w_i(TM) = Sq^i(u)/[M].$$

□

PROOF. (of Theorem 6.37) Pick a homogeneous basis $b_1, \ldots, b_r$ of $H^*(M; \mathbb{Z}_2)$. By Lemma 6.39 we may find a dual basis $c_1, \ldots, c_r$. Let $\nu = 1 + \nu_1 + \cdots + \nu_n$ be the *total Wu class*. Similarly define the *total Steenrod square* $Sq = 1 + Sq^1 + \cdots + Sq^n$. We must have,

$$\nu = \sum_i b_i \langle \nu \smile c_i, [M] \rangle.$$

By definition of the Wu classes,

$$= \sum_i b_i \langle Sq(c_i), [M] \rangle.$$

Taking the total steenrod square of both sides,

$$\begin{aligned} Sq(\nu) &= \sum_i Sq(b_i) \langle Sq(c_i), [M] \rangle \\ &= \sum_i [Sq(b_i) \times Sq(c_i)]/[M]. \end{aligned}$$

By Lemma 6.39,

$$= Sq(u)/[M].$$

By Lemma 6.40,

$$= w(TM).$$

□

COROLLARY 6.41. *If $f : M \to N$ is a continuous map of closed $n$ manifolds which is an isomorphism on $\mathbb{Z}_2$-homology, then $TM$ and $TN$ have the same Stiefel–Whitney classes.*

PROOF. The Wu classes and Steenrods squares are defined intrinsically from the structure of the $\mathbb{Z}_2$ cohomology. The Wu formula shows that this suffices to recover the Stiefel–Whitney classes. □

This is a surprising result, because defining the tangent bundle explicitly requires knowing the smooth structure. Nevertheless, we learn for example that if a topological manifold $M$ admits many distinct or "exotic" smooth structures, all its smooth structures must have the same Stiefel–Whitney classes on $TM$. The Wu formula also has some nice geometric consequences, like the following.

PROPOSITION 6.42 (Stiefel). *Every closed orientable three-manifold is parallelizable*

PROOF. Let $M$ be a closed oriented three-manifold. Because our manifold is orientable, a 2-frame of the tangent bundle can always be extended to a 3-frame, i.e. a trivialization. Thus it suffices to construct a section of the 2-frame bundle $F_2$ on $M$; this is a fibre bundle with fibre $V(3,2) \cong \mathbb{R}P^3$. Note that $\pi_1(\mathbb{R}P^3) = \mathbb{Z}_2$ and $\pi_2(\mathbb{R}P^3) = \pi_2(S^3) = 0$.

The first obstruction to finding a section of $F_2$ on the 2-skeleton is by definition the second Stiefel–Whitney class $w_2(TM) \in H^2(M;\mathbb{Z}_2)$. If $w_2(TM) = 0$, the next obstruction to extending a section to the 3-skeleton lives in the group $H^3(M;\pi_2(V(3,2))) = 0$. Since $M$ is three dimensional, this is the last obstruction. Thus it suffices to show the first obstruction $w_2(TM) = 0$.

By the Wu formula, $w_1(TM) = \nu_1$ and since $M$ is orientable $\nu_1 = 0$. Recall $\nu_2$ is defined so that $\nu_2 \smile x = Sq^2(x)$ for any $x \in H^1(M;\mathbb{Z}_2)$. But $Sq^2 : H^1(M;\mathbb{Z}_2) \to H^3(M;\mathbb{Z}_2)$ is axiomatically the zero operation. So, $\nu_2 = 0$. Then,

$$w_2(TM) = Sq^1(\nu_1) + Sq^0(\nu_2) = \nu_1^2 + \nu_2 = 0.$$

□

There is another result also called the Wu formula to compute the Steenrod squares of Stiefel–Whitney classes.

**Theorem 6.43: Wu Formula for Stiefel–Whitney Classes of Steenrod Squares**

For any bundle, the Steenrod squares of the Stiefel–Whitney classes obey the formula,

$$Sq^i(w_j) = \sum_{k=0}^{i} \binom{j+k-i-1}{k} w_{i-k} w_{j+k}.$$

PROOF. Let $x$ be the generator of $H^1(\mathbb{R}P^\infty;\mathbb{Z}_2)$. We must have $Sq(x) = x + x^2$. From the Cartan formula we obtain, $Sq(x^j) = (x + x^2)^j$ and so $Sq^i(x^j) = \binom{j}{i} x^{j+1}$.

By linearity and the Cartan formula, we can extend to determine the action of $Sq$ on the cohomology of $(\mathbb{R}P^\infty)^n$. This determines the action of $Sq$ on the cohomology of $BO(n)$ using our usual splitting principle argument and we should obtain the formula above. And then by naturality, this determines the action of $Sq$ on the Stiefel–Whitney classes of any bundle. □

*Splitting Principle and Other Properties.* We have informally invoked the splitting principle several times. Let us be a bit more formal about what is going on and use it to describe how characteristic classes behave under direct sums and tensor products. We will then describe a few more important properties.

Let us deal with all cases simultaneously; so let $G$ denote $O(n), U(n)$ or $Sp(n)$ and $T$ the maximal torus of $G$. For the $O(n)$ case, $\mathbb{Z}_2$ coefficients in cohomology are to be understood. As we have stated many times before $Bi : BT \to BG$ is a fibration with injective pullback in cohomology. Suppose $E \to X$ is a bundle induced by a map $g : X \to BG$, we can pull back the above fibration along $g$.

$$\begin{array}{ccccc} f^*E & \longrightarrow & Y & \longrightarrow & BT \\ \downarrow & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle Bi} \\ E & \longrightarrow & X & \xrightarrow{g} & BG \end{array}$$

We thus obtain a space $Y$ and a $G/T$ fibre bundle $f : Y \to X$ so the above commutes. Moreover, we can pullback $E \to X$ to obtain a bundle $f^*E$. Note because the classifying map $f \to BG$ factors through $BT$, the bundle $f^*E$ must be a direct sum of $n$ line bundles.

Recall that we had an isomorphism,

$$H^*(BT) \cong H^*(BG) \otimes H^*(G/T),$$

since the Serre spectral sequence of the fibration collapsed immediately.

Using the morphism of spectral sequences induced by $g$, the spectral sequence associated to $f$ must also collapse:

$$H^*(Y) \cong H^*(X) \otimes H^*(G/T),$$

and moreover the pullback $f^* : H^*(X) \to H^*(Y)$ is the edge morphism $x \mapsto x \otimes 1$. In particular, the pullback is injective. We have thus proved the following.

**Theorem 6.44: Generalized Splitting Principle**

Given a $G$-bundle $E \to X$, there is a bundle $f : Y \to X$ with fibre $G/T$ so that $f^*E$ has structure group $T$ and $f^*$ is injective on cohomology.

One of the upshots of this, is that if we can prove some algebraic fact about the characteristic classes of the pullback bundle with reduced structure group $T$, it must hold for the original bundle in the cohomology of $X$. In particular, we deduce that *to show an algebraic relation in terms of the Stiefel–Whitney, Chern, or quaternionic Pontryagin classes holds for a general bundle it will suffice to do so for a direct sum of line bundles.* Some more general examples of the splitting principle can be found here in a note of May.

Let us now use this to study the behaviour of characteristic classes under direct sums.

**Theorem 6.45: Whitney Sum Formula**

Given two bundles $\xi, \eta$, the total characteristic classes of their direct sum obey a product formula. So if they are real bundles,

$$w(\xi \oplus \eta) = w(\xi) \smile w(\eta).$$

If they are oriented,

$$e(\xi \oplus \eta) = e(\xi) \smile e(\eta).$$

If they are complex,

$$c(\xi \oplus \eta) = c(\xi) \smile c(\eta).$$

The same holds for Pontryagin classes modulo 2,

$$2p(\xi \oplus \eta) = 2p(\xi) \smile p(\eta)$$

Proof. We have already proved this for Stiefel–Whitney classes, essentially using the splitting principle and the fact $(BT)^n = B(T^n)$ with the expect isomorphism in cohomology. Exactly the same idea holds for Chern classes. It also holds for Pontryagin classes provided we work in a ring with $1/2$.

Here's an alternate proof for Pontryagin classes. Note first for a real bundle that the odd Chern classes of $\mathbb{C}\xi$ are 2-torsion; we will prove this below in Corollary 6.51. We have,

$$\begin{aligned} 2p_k(\xi \oplus \eta) &= (-1)^k 2 \cdot c_{2k}(\mathbb{C}(\xi \oplus \eta)) \\ &= (-1)^k \sum_{i \leq 2k} 2 \cdot c_i(\mathbb{C}\xi) \smile c_{2k-i}(\mathbb{C}\eta). \end{aligned}$$

Ignoring the odd Chern classes, which disappear under 2-torsion,

$$\begin{aligned}
&= (-1)^k \sum_{i \leq k} 2 \cdot c_{2i}(\mathbb{C}\xi) \smile c_{2k-2i}(\mathbb{C}\eta) \\
&= (-1)^k \sum_{i \leq k} 2(-1)^i p_i(\xi) \smile (-1)^{k-i} p_{k-i}(\eta) \\
&= 2 \sum_{i \leq k} p_i(\xi) \smile p_{k-i}(\eta)
\end{aligned}$$

For the Euler case, note that one can study the Thom class of the product bundle $\xi \times \eta \to B^2$ and it will be the cross product of the Thom classes of the two individual bundles by a local calculation. Restricting to the base and then pulling back by the diagonal embedding $\Delta : B \to B^2$ recovers on one hand the Euler class of $\xi \oplus \eta$ and on the other hand the cup product of $e(\xi)$ and $e(\eta)$. □

Note the following important corollary of this. Elements in the ring of $\mathbb{Z}$ or $\mathbb{Z}_2$ power series are invertible if and only if their degree zero term is one. This is a simple exercise that can be proved inductively, or by construction a germ of the reciprocal of the power series at zero using Taylor series. Since the total Stiefel–Whitney and Chern classes have this form, $w(\xi)$ **and** $c(\xi)$ **are invertible in the cohomology ring**. This means that if we have a bundle $\eta$ with simple characteristic classes and so that the sum $\xi \oplus \eta$ has simple characteristic classes then we can invert the characteristic classes of $\eta$ and use the Whitney sum formula to find the characteristic classes of $\xi$.

We will apply this later to the case where $\xi = TM$ is the tangent bundle of a manifold $M$ and $\eta = \nu_M$ is the normal bundle to an embedding into Euclidean space $\mathbb{R}^n$. Almost by definition, $TM \oplus \nu_M = T\mathbb{R}^n|_M$, the latter of which is a trivial bundle and hence has total Stiefel Whitney class $w = 1$. We thus obtain the following useful result.

Proposition 6.46 (Whitney Duality Formula). *If $M$ is a smooth manifold and $f : M \to \mathbb{R}^n$ is a smooth embedding with normal bundle $\nu_M$ then,*

$$w(TM) = w(\nu_M)^{-1}.$$

Note that although the same Whiteny sum formula holds for the Euler class, the Euler class will not be a unit in the cohomology ring and may in fact be a zero divisor. So we cannot usually solve for the Euler class of a summand just using the sum formula.

As with the direct sum, we can perform similar splitting principle analysis for the tensor product of two bundles, but the result will not be as clean. We state it here just for reference.

**Theorem 6.47: Characteristic Classes of Tensor Products**

For $\xi, \eta$ real bundles of ranks $n$ and $m$,

$$w(\xi \otimes \eta) = p_{n,m}(w_1(\xi), \dots, w_n(\xi), w_1(\eta), \dots, w_m(\eta)).$$

And if they are complex bundles,

$$c(\xi \otimes \eta) = p_{n,m}(c_1(\xi), \dots, c_n(\xi), c_1(\eta), \dots, c_m(\eta)).$$

Here $p_{n,m}$ is the unique polynomial so that if $\sigma_i$ and $\sigma_i'$ are the elementary symmetric polynomials in the variables $t_1, \dots, t_n$ and $s_1, \dots, s_m$ respectively then,

$$p_{n,m}(\sigma_1, \dots, \sigma_n, \sigma_1', \dots \sigma_m') = \prod_{i=1}^{n} \prod_{j=1}^{m} (1 + t_i + s_j).$$

In particular, if the bundles are real,

$$w_1(\xi \otimes \eta) = m w_1(\xi) + n w_1(\eta),$$

and if they are complex,

$$c_1(\xi \otimes \eta) = m c_1(\xi) + n c_1(\eta).$$

There is another equivalent way to find the classes of a tensor product. We will just look at the complex case, which is more commonly studied.

**Definition 6.48: Chern Character**

Suppose $\xi = L_1 \oplus \cdots \oplus L_n$ is a direct sum of complex line bundles. We define its *Chern character* to be the power series in the cohomology ring,

$$\mathrm{ch}(\xi) := e^{c_1(L_1)} + \cdots + e^{c_1(L_n)}.$$

By the splitting principle, this definition extends to an arbitrary complex vector bundle. Explicitly the first few terms of the Chern character are given by,

$$\mathrm{ch}(\xi) = \mathrm{rk} + c_1 + \frac{1}{2}(c_1^2 - c_2) + \frac{1}{6}(c_1^3 - 3c_1c_2 + 3c_3) + \cdots.$$

It satisfies the formulae,

$$\mathrm{ch}(\xi \oplus \eta) = \mathrm{ch}(\xi) + \mathrm{ch}(\eta),$$
$$\mathrm{ch}(\xi \otimes \eta) = \mathrm{ch}(\xi)\mathrm{ch}(\eta).$$

Using this definition, we see that,

$$\begin{aligned}\mathrm{ch}(\xi\otimes\eta) &= (\mathrm{rk}(\xi)+c_1(\xi)+\frac{1}{2}(c_1^2(\xi)-c_2(\xi))+\cdots)\\ &\qquad\cdot(\mathrm{rk}(\eta)+c_1(\eta)+\frac{1}{2}(c_1^2(\eta)-c_2(\eta))+\cdots)\\ &= \mathrm{rk}(\xi)\mathrm{rk}(\eta)+\mathrm{rk}(\eta)c_1(\xi)+\mathrm{rk}(\xi)c_1(\eta)+c_1(\xi)c_1(\eta)\\ &\qquad+\frac{\mathrm{rk}(\xi)}{2}(c_1^2(\eta)-c_2(\eta))+\frac{\mathrm{rk}(\eta)}{2}(c_1^2(\xi)-c_2(\xi))+\cdots.\end{aligned}$$

From this we recover our formula for the first Chern class above. We also obtain the identity,

$$c_1(\xi\otimes\eta)^2-c_2(\xi\otimes\eta)=2c_1(\xi)c_1(\eta)+\mathrm{rk}(\xi)\left[c_1^2(\eta)-c_2(\eta)\right]+\mathrm{rk}(\eta)\left[c_1^2(\xi)-c_2(\xi)\right].$$

Since we know $c_1(\xi\otimes\eta)$, this tells us $c_2$. In the abstract this computation is a little crazy, but for concrete $\xi$ and $\eta$ it may not be so bad.

The Chern character is important for the following $K$-theory result. Note that the properties of the Chern character and the stability of characteristic classes means it defines a ring homomorphism,

$$\mathrm{ch}_{\mathbb{Q}} : K^0(X)\otimes\mathbb{Q}\to H^{\mathrm{even}}(X;\mathbb{Q}).$$

Using stability under suspension of homology and $K$-theory, this extends in a natural way to a $\mathbb{Z}_2$-graded ring homomorphism,

$$\mathrm{ch}_{\mathbb{Q}} : K^*(X)\otimes\mathbb{Q}\to H^*(X;\mathbb{Q}).$$

**Theorem 6.49**

For $X$ a finite CW complex, the Chern character induces a rational isomorphism,

$$\mathrm{ch}_{\mathbb{Q}} : K^*(X)\otimes\mathbb{Q}\cong H^*(X;\mathbb{Q}).$$

Proof. Consider the Atiyah–Hirzebruch spectral sequence for the $K$-theory of $X$ tensored with $\mathbb{Q}$. As we have discussed before, there are no non-trivial stable cohomology operations on rational homology, and hence the spectral sequence must immediately degenerate. Since $K^*(\mathrm{pt})\otimes\mathbb{Q}=\mathbb{Q}$ in degree zero and zero in degree, we obtain the desired isomorphism. The fact this isomorphism is realized by $\mathrm{ch}_{\mathbb{Q}}$ can be argued using the naturality of the Atiyah–Hirzerbruch spectral sequence and studying the $K$-theory of $BU$. □

We now discuss elementary transformations of bundles.

**Theorem 6.50**

If $\xi$ is a real bundle with dual $\xi^*$, then $\xi$ and $\xi^*$ have the same characteristic classes.

If $\xi$ is an oriented bundle, and $-\xi$ is the same bundle with opposite orientation, then $e(-\xi) = -e(\xi)$.

If $\xi$ is a complex bundle with dual $\xi^*$ and the bundle $\overline{\xi}$ is the same underlying bundle as $\xi$ equipped with the opposite complex structure, then,

$$c_i(\overline{\xi}) = c_i(\xi^*) = (-1)^i c_i(\xi).$$

PROOF. Picking a metric on a real bundle $\xi$ gives a canonical identification of $\xi$ and $\xi^*$. So these bundles are isomorphic and hence have the same characteristic classes.

Now suppose $\xi$ is oriented. Note that if we reverse the orientation of $\xi$, this by definition will reverse the sign of the fundamental class of the fibres in the Thom space of $\xi$. Hence the Thom class, which is defined to restrict to the fundamental class of each fibre, will satisfy $t_{-\xi} = -t_\xi$. Since the Euler class is the restriction of the Thom class to the base we conclude $e(-\xi) = -e(\xi)$. As an alternate proof, note that the Euler classes of $\xi$ and $-\xi$ are both defined as the primary obstruction in $H^n(X; \pi_{n-1}(S^{n-1}))$ to the existence of a section of the underlying unoriented bundle. Changing the orientation simply changes the fibrewise identification of $\pi_{n-1}(S^{n-1})$ with $\mathbb{Z}$ by a sign, and so will change the Euler class by a sign.

Lastly, let $\xi$ be complex. Note that a choice of Hermitian metric gives a canonical isomorphism of $\xi^*$ and $\overline{\xi}$. Hence it will suffice to relate $c(\overline{\xi})$ and $c(\xi)$. Note that if $\xi$ is a complex line bundle, then $c_1(\xi) = e(\xi)$. If we use the conjugate bundle $\overline{\xi}$, it will be the same underlying real bundle but with opposite orientation (since an old $\mathbb{R}$-basis $v, iv$ for a fibre of $\xi$ becomes $v, -iv$ under conjugation). By the Euler computation above we know,

$$c_1(\overline{\xi}) = e(\overline{\xi}) = -e(\xi) = -c_1(\xi).$$

Now suppose $\xi = L_1 \oplus \cdots \oplus L_n$ is a direct sum of complex line bundles. We have,

$$\begin{aligned} c(\overline{\xi}) &= \prod_{i=1}^{n}(1 + c_1(\overline{L_i})) \\ &= \prod_{i=1}^{n}(1 - c_1(L_i)) \\ &= 1 - c_1(\xi) + c_2(\xi) - \cdots \pm c_n(\xi). \end{aligned}$$

We deduce the result from the splitting principle. □

COROLLARY 6.51. *If $\xi$ is an orientable odd dimensional bundle, then $e(\xi)$ is 2-torsion.*

*If $\eta$ is a complex bundle which is self-dual (for example if $\eta$ is the complexification of a real bundle) than its odd Chern classes are 2-torsion.*

PROOF. If $\xi$ is any bundle, there is an isomorphism given by multiplication by minus one on the fibres: $(p, v) \mapsto (p, -v)$. If $\xi$ has odd rank and is orientable, this isomorphism will reverse the orientation. Thus, $\xi = -\xi$ and so $e(\xi) = e(-\xi) = -e(\xi)$. So $2e(\xi) = 0$. The same argument works for odd Chern classes in the self-dual case. □

*Computations for Smooth Manifolds.* We will not do very many computations, but the canonical example is to find the Stiefel–Whitney classes of the tangent bundle of $\mathbb{R}P^n$ and the Chern class of the tangent bundle of $\mathbb{C}P^n$. We should note the convention that we refer to the characteristic classes of a manifold to mean the characteristic classes of its tangent bundle. In this case we write $w(M)$ in place of $w(TM)$.

Let us write the homology of $\mathbb{R}P^n$ as $\mathbb{Z}[a]/(a^{n+1})$ and the homology of $\mathbb{C}P^n$ as $\mathbb{Z}[x]/(x^{n+1})$ where $|a| = 1$ and $x$ is Poincaré dual to $[\mathbb{C}P^{n-1}]$.

**Theorem 6.52**

The total Stiefel–Whitney class of $\mathbb{R}P^n$ is,

$$w(\mathbb{R}P^n) = (1 + a)^{n+1}.$$

The total Chern classes of $\mathbb{C}P^n$ is,

$$c(\mathbb{C}P^n) = (1 + x)^{n+1}$$

PROOF. We claim the tangent bundle $T = T\mathbb{R}P^n$ is isomorphic to the bundle $\mathrm{Hom}(L, \underline{\mathbb{R}}^{n+1}/L)$, where $L$ is the tautological line bundle on $\mathbb{R}P^n$ and $\underline{\mathbb{R}}^{n+1}$ is a trivial bundle.

By lifting to $S^n$, we see that over a point $\ell \in \mathbb{R}P^n$ intersecting $x, -x$ on the unit circle, the fibre of $T$ consists of pairs $\{(x, v), (-x, -v)\}$ for a choice of $v \in \mathbb{R}^{n+1}$ so that $x \cdot v = 0$, i.e. $v \in \ell^\perp$. On the other hand, over $\ell$, the fibre of $\mathrm{Hom}(L, \underline{\mathbb{R}}^{n+1}/L)$ is given by the space of linear maps from the one-dimensional vector space $\ell$ to the vector space $\mathbb{R}^{n+1}/\ell$. Considering the standard metric on $\mathbb{R}^{n+1}$ (and hence on $\underline{\mathbb{R}}^{n+1}$), we may identify $\mathbb{R}^{n+1}/\ell$ with $\ell^\perp \subset \mathbb{R}^{n+1}$. Thus our fibre over $\ell$ is $\mathrm{Hom}(\ell, \ell^\perp)$. Since $\ell$ is one-dimensional, the elements $\varphi \in \mathrm{Hom}(\ell, \ell^\perp)$ are in one to one correspondence with pairs $\{(x, v), (-x, -v)\}$ above via $\varphi \mapsto \{(x, \varphi(x)), (-x, -\varphi(x))\}$. This fives the desired isomorphism.

We now use this to find the Stiefel-Whitney classes. Note that we have $\mathrm{Hom}(L, L) = \underline{\mathbb{R}}$ since the identity map provides a non-vanishing section of the line bundle $\mathrm{Hom}(L, L)$. We then compute,

$$T \oplus \underline{\mathbb{R}} = \mathrm{Hom}(L, \underline{\mathbb{R}}^{n+1}/L) \oplus \mathrm{Hom}(L, L).$$

We have an equality,

$$\begin{aligned} &= \mathrm{Hom}(L, (\underline{\mathbb{R}}^{n+1}/L) \oplus L) \\ &= \mathrm{Hom}(L, \underline{\mathbb{R}}^{n+1}) \\ &= \bigoplus_{i=1}^{n+1} \mathrm{Hom}(L, \underline{\mathbb{R}}) \end{aligned}$$

Since $L$ is isomorphic to its dual,

$$= \bigoplus_{i=1}^{n+1} L.$$

We claim the line bundle $L$ has total Stiefel-Whitney class $w(L) = 1+a$. To see this, note that the embedding $i : \mathbb{R}P^1 \to \mathbb{R}P^n$ gives the Möbius bundle $M$ over the circle as the pullback $i^*L$. By naturality, we have $w_1(M) = i^* w_1(L)$. By the non-triviality axiom of the Stiefel-Whitney classes we know that $w_1(M) = 1 + i^*a$ and so $w(L) = 1 + a$.

Since the Stiefel-Whitney classes are stable,

$$\begin{aligned} w(\mathbb{R}P^n) &= w(T \oplus \underline{\mathbb{R}}) \\ &= w\left(\bigoplus_{i=1}^{n+1} L\right). \end{aligned}$$

By the properties of Stiefel-Whitney classes under direct sum,

$$\begin{aligned} &= w(L)^{n+1} \\ &= (1+a)^{n+1}. \end{aligned}$$

The proof in the Chern case is essentially identical. We obtain as above, as identification

$$T\mathbb{C}P^n \oplus \underline{\mathbb{C}} = \bigoplus_{i=1}^{n+1} \mathrm{Hom}(\gamma, \underline{\mathbb{C}}) = \bigoplus_{i=1}^{n+1} \overline{\gamma},$$

where $\gamma$ is the tautological line bundle on $\mathbb{C}P^n$. Let $x = -c_1(\gamma) \in H^2(\mathbb{C}P^n)$. We have,

$$c(\mathbb{C}P^n) = c(\overline{\gamma})^{n+1} = (1+x)^{n+1}.$$

To determine $x$, we note,

$$e(\mathbb{C}P^n) = c_n(\mathbb{C}P^n) = (n+1)x^n.$$

We have,

$$\chi(\mathbb{C}P^n) = \langle e(\mathbb{C}P^n), [\mathbb{C}P^n] \rangle = (n+1)\langle x^n, [\mathbb{C}P^n] \rangle.$$

But the Euler characteristic of $\mathbb{C}P^n$ is $n+1$, so $\langle x^n, \mathbb{C}P^n \rangle = 1$, meaning that $x$ is the compatibly oriented generator of $H^1(\mathbb{C}P^n)$, i.e. the Poincaré dual of $[\mathbb{C}P^{n-1}]$. □

COROLLARY 6.53. *$\mathbb{R}P^{2^r}$ cannot be immersed in $\mathbb{R}^N$ for $N < 2^{r+1} - 1$.*

PROOF. Note if $n = 2^r$, then $w(\mathbb{R}P^n) = 1 + a + a^n$. If $f : \mathbb{R}P^n \to \mathbb{R}^N$ is an immersion with normal bundle $\nu$, the Whitney duality formula implies

$$w(\nu) = w(\mathbb{R}P^n)^{-1} = 1 + a + a^2 + \cdots + a^{n-1}.$$

But note the rank of $\nu$ is $N - n$, and since $w_{n-1}(\nu) \neq 0$ we must have $N - n \geq n - 1$, or equivalently $N \geq 2^{r+1} - 1$. □

REMARK 6.54. Note the Whitney immersion theorem says any compact $n > 1$ manifold can be immersed in $\mathbb{R}^{2n-1}$. This result is proved optimal in certain dimensions by the spaces $\mathbb{R}P^{2^r}$. △

Here is one more theoretical result.

**Theorem 6.55**

Compact non-orientable surfaces cannot be smoothly embedded in $\mathbb{R}^3$.

PROOF. Let $\Sigma$ be a compact non-orientable surface and let $f : \Sigma \to \mathbb{R}^3$ be a smooth embedding with normal bundle $\nu$. We have, by Whitney duality,

$$w(\Sigma) = w(\nu)^{-1} = 1 + w_1(\nu).$$

Since $\Sigma$ is non-orientable, $w_1(\Sigma) \neq 0$ and so $w_1(\nu) \neq 0$ as well. This implies $\nu$ is a trivial line bundle. As we discussed previously, the Euler class of a normal bundle to an embedding measures the self-intersection of a surface. Hence, for $w_1(\nu) \neq 0$ we must have that the (un-oriented) self intersection number of $\Sigma$ inside of $\mathbb{R}^3$ is non-zero. But this is clearly not true. If we perturb the image of $\Sigma$ inside of $\mathbb{R}^3$ to $\Sigma'$, then we may "drag" $\Sigma'$ away from $\Sigma$ and off to infinity. This will provide a null cobordism of $\Sigma \cap \Sigma'$. Hence $w_1(\nu) = 0$. □

*Characteristic Numbers of Smooth Manifolds.* From the characteristic classes of a smooth manifold, we can extract certain numerical invariants. These allow for direct comparisons of characteristic classes.

**Definition 6.56: Characteristic Numbers**

Let $M$ be a closed $n$-manifold with $\mathbb{Z}_2$ fundamental class $[M]$. Let $r_1, r_2, \ldots, r_n$ be a sequence of non-negative integers so that $r_1 + 2r_2 + \cdots + nr_n = n$. Then we define a *Stiefel–Whitney number* of $M$,

$$w_{r_1,\ldots,r_n}(M) = \langle w_1(M)^{r_1} \cdots w_n(M)^{r_n}, [M] \rangle \in \mathbb{Z}_2.$$

If $M$ is a closed oriented $4n$-manifold with oriented fundamental class $[M]$, and $r_1, r_2, \ldots, r_n$ a sequence of non-negative integers so that $r_1 + 2r_2 + \cdots + nr_n = n$ then we define a *Pontryagin number* of $M$,

$$p_{r_1,\ldots,r_n}(M) = \langle p_1(M)^{r_1} \cdots p_n(M)^{r_n}, [M] \rangle \in \mathbb{Z}.$$

If $M$ is a closed almost complex $2n$-manifold with fundamental class $[M]$, and $r_1, r_2, \ldots, r_n$ a sequence of non-negative integers so that $r_1 + 2r_2 + \cdots + nr_n = n$ then we define a *Chern number* of $M$,

$$c_{r_1,\ldots,r_n}(M) = \langle c_1(M)^{r_1} \cdots c_n(M)^{r_n}, [M] \rangle \in \mathbb{Z}.$$

EXAMPLE 6.57. Consider an odd projective space $\mathbb{R}P^{2n-1}$. Then $w(\mathbb{R}P^{2n-1}) = (1+a)^{2n} = (1+a^2)^n$. In particular, $w_i(\mathbb{R}P^{2n-1}) = 0$ for $n$ odd. But every Stiefel–Whitney number $w_{r_1,\ldots,r_{2n-1}}(\mathbb{R}P^{2n-1})$ must have some $r_i \neq 0$ for $i$ odd. Hence all the Stiefel–Whitney numbers of $\mathbb{R}P^{2n-1}$ vanish. ▲

One major use of characteristic numbers is as bordism invariants. Consider the following proposition.

PROPOSITION 6.58. *If a compact $n$-manifold $M$ is null-bordant, then its Stiefel–Whitney numbers all vanish.*

PROOF. Suppose $M = \partial W$. By the tubular neighbourhood theorem, $TW|_M$ decomposes as a direct sum of $TM$ and a trivial line bundle. Thus under the inclusion $i : M \hookrightarrow W$, $i^* w(W) = w(M)$.

Note that the $\mathbb{Z}_2$ fundamental class $[M]$ is the boundary (in the LES of a pair) of the relative fundamental class $[W, M]$. We have,

$$w_{r_1,\dots,r_j}(M) = \langle i^* w_{r_1,\dots,r_j}(W), \partial[M, W]\rangle = \langle w_{r_1,\dots,r_j}(W), i_*\partial[M, W]\rangle = 0,$$

since $i_*\partial = 0$ is the composition of consecutive maps in the LES of a pair. □

This result extends to other characteristic numbers. And remarkably, the converse is also true. The proof of this is closely linked to the computation of the cobordism rings which we discussed earlier.

**Theorem 6.59: Milnor, Novikov, Thom, Wall**

Two compact $n$-manifolds are cobordant if and only if their Stiefel–Whitney numbers agree.

Two compact oriented $n$-manifolds are oriented cobordant if and only if their Stiefel–Whitney numbers and Pontryagin numbers agree.

Two compact stably almost complex $n$-manifolds are complex cobordant if and only if their Chern numbers agree.

**6.5. Thom Spectra and the Pontryagin–Thom Construction.** Earlier we described a generalized homology theory based on singular manifolds up to cobordism called the bordism groups. By Brown representability, we know the bordism groups must be representable, i.e. the groups are given by homotopy classes of maps into some sequence of spaces. A very nice geometric model for these spaces and a proof they reproduce the bordism groups is due to Pontryagin and Thom and goes by the name of the *Pontryagin–Thom construction*.

**Definition 6.60: Thom Spectrum**

Let $\nu_n \to BO(n)$ be the tautological bundle, i.e. the associated bundle of the universal bundle $EO(n) \to BO(n)$. Define $MO(n)$ to be the Thom space $T\nu_n$. Note that the map induced by inclusion $BO(n) \to BO(n+1)$ induces a map $\nu_n \oplus \underline{\mathbb{R}} \to \nu_{n+1}$ and thus a map on the corresponding Thom spaces. The Thom space of $\nu_n \oplus \underline{\mathbb{R}}$ is the same thing as the suspension of $T\nu_n$. Hence we have a map,

$$\Sigma MO(n) \to MO(n+1).$$

The sequence of spaces $\{MO(n)\}_{n\in\mathbb{N}}$ along with these suspension maps is called the *unoriented Thom spectrum*.

If $\nu_n^+$ is the tautological bundle on $BSO(n)$ then $MSO(n) = T\nu_n^+$. If $\nu_n^{\mathbb{C}}$ is the tautological bundle on $BU(n)$ then $MU(n) = T\nu_n^{\mathbb{C}}$. We get corresponding suspension maps,

$$\Sigma MSO(n) \to MSO(n+1)$$
$$\Sigma^2 MU(n) \to MU(n+1).$$

These are called the *oriented Thom spectrum* and *complex Thom spectrum* respectively. For any sufficiently nice topological group $G$ ($Sp(n), SU(n), Spin(n)$ etc.) we have a corresponding $G$-structure Thom spectrum $MG$ defined analogously.

There is a multiplicative structure on the Thom spectra. Consider the map $BO(n)\times BO(m) \to BO(n+m)$ obtained by applying the classifying functor to the inclusion of block diagonal orthogonal matrices. This classifies a map of vector bundles $\nu_n \oplus \nu_m \to \nu_{n+m}$. Taking the Thom space of both sides produces a map,

$$MO(n) \wedge MO(m) \to MO(n+m).$$

Analogous constructions will work for any Thom spectrum $MG(n)$.

Now suppose we have some singular compact $q$-manifold $f : M^q \to X$. By Whitney embedding, we can find an embedding $i : M^q \hookrightarrow \mathbb{R}^{n+q}$. By compactifying the codomain, we can turn this into an embedding into $S^{n+q}$. Let $\nu_i$ be the normal bundle of the embedding, which we identify with a tubular neighbourhood of $i(M^q) \subset S^{n+q}$. We now have the "Pontryagin–Thom collapse map." This is defined as,

$$PT : S^{n+q} \to S^{n+q}/(S^{n+q} \setminus \nu_i) \to T\nu_i,$$

where the first map is the quotient identifying the complement of the tubular neighbourhood to a point and the second is an identification of this compactified tubular neighbourhood with the Thom space of the bundle. We may extend the map $f : M^q \to X$ to the normal bundle constantly on the fibres. We thus obtain a smash product,

$$f \wedge PT : S^{n+q} \to X_+ \wedge T\nu_i.$$

The normal bundle $\nu_i$ is classified by some map $M \to BO(n)$. Looking at the map on total spaces and taking Thom spaces yields a map of $T\nu_i \to MO(n)$. Composing with the map above defines a map,

$$PT_f : S^{n+q} \to X_+ \wedge MO(n).$$

There was a choice made in constructing this map beyond just what $f$ was; namely, the embedding we used and its dimension. Recall the fact that stably, all embeddings are isotopic. To see this, given embedddings $i_1, i_2 : M \hookrightarrow \mathbb{R}^{n_1,n_2}$, we can stabilize to embeddings $\widetilde{i_1}, \widetilde{i_2} : M \hookrightarrow \mathbb{R}^{n_1+n_2}$ and then there is an isotopy,

$$\Upsilon(x,t) = \widetilde{i}_1(x)t + (1-t)\widetilde{i}_2(x).$$

If we stabilize the embedding $i : M^q \hookrightarrow \mathbb{R}^{n+q}$ we chose earlier, the map $PT_f$ is suspended to the map,

$$\Sigma PT_f : S^{n+q+1} \to X_+ \wedge MO(n+1).$$

If we have two embeddings, stabilized to be isotopic, the isotopy will define a homotopy of the two versions of the map $PT_f$. Thus taking the colimit under stabilization and considering the map only up to homotopy, our map $PT_f$ will be a canonical element independent of embedding,

$$[\Sigma^\infty PT_f] \in \lim_{n\to\infty} [S^{n+q}, X_+ \wedge MO(n)].$$

Here the colimit is defined over the maps,

$$[S^{n+q}, X_+ \wedge MO(n)] \xrightarrow{\Sigma} [S^{n+q+1}, X_+ \wedge \Sigma MO(n)] \to [S^{n+q+1}, X_+ \wedge MO(n+1)],$$

where the second map comes from the Thom spectrum map.

Note that if the map $f : M \to X$ is cobordant to another map $g : N \to X$, then this cobordism will define a homotopy between the classes $[\Sigma^\infty PT_f]$ and $[\Sigma^\infty PT_g]$. Hence this Pontryagin–Thom map descends to a map,

$$PT : \Omega_q^O(X) \to \lim_{n\to\infty} [S^{n+q}, X_+ \wedge MO(n)].$$

Actually this is a group homomorphism. Here the group structure on the right comes from the co-$H$-space structure on $S^{n+q}$ which is compatible with the diagram in the colimit. To see this, if we take the cobordism class of two singular manifolds $f, g$, the corresponding collapse map $PT_{f,g}$ up to homotopy is given by gluing the domains of the collapse maps $PT_f$ and $PT_g$ under the co-$H$-space map $S^{n+q} \to S^{n+q} \vee S^{n+q}$. It is thus immediate that the $PT$ map respects group structures.

We now construct a candidate inverse to this map. Suppose we have some map $f : S^{n+q} \to X_+ \wedge MO(n)$. Compose with projection to obtain a map $S^{n+q} \to MO(n)$. Up to homotopy this map can be made to land in $\gamma_n^r \subset MO(n)$, where $\gamma_n^r$ is the tautological $n$-plane bundle on $G(r,n)$ for sufficiently large $r$. By the Thom transversality theorem, we may homotope $f$ to obtain a smooth map $S^{n+q} \to \gamma_n^r$ which is transverse to the zero section. Hence the preimage of the zero locus will be a smooth compact $q$-manifold $M^q$ naturally embedded

in $\mathbb{R}^{n+q} \subset S^{n+q}$. Restricting $f$ and projecting to $X$ then defines a singular $q$-manifold $PT_f^{-1} : M^q \to X$. Note that if we take a homotopy of our map $f$ to a map $g$, it will induce a cobordism between $PT_f^{-1}$ and $PT_g^{-1}$ by looking (after a homotopy of homotopies) at the zero locus of the induced map $S^{n+q} \times I \to \gamma_n^R$. And we may freely stabilize the map $f$ to a map $S^{n+q+1} \to X_+ \wedge MO(n+1)$ and we will obtain the same $q$-manifold $M^q$ (up to cobordism) but now embedded in $\mathbb{R}^{n+q+1}$. We conclude there is a reverse map,

$$PT^{-1} : \lim_{n\to\infty} [S^{n+q}, X_+ \wedge MO(n)] \to \Omega_q^O(X).$$

If we have two maps $f : S^{n+q} \to X_+ \wedge MO(n), g : S^{n+r} \to X_+ \wedge MO(n)$, then the combined map $f \vee g$ can be seen to have (up to bordism) $PT_{f\vee g}^{-1} = PT_f^{-1} \sqcup PT_g^{-1}$. Hence this inverse is also a group homomorphism.

Let us show these are indeed inverses. If $f : M^q \to X$ is a singular manifold, then $PT_f$ is a map $S^{n+q} \to X_+ \wedge MO(n)$. If we take the preimage of the zero section in $MO(n)$ under this map, we reobtain $M^q \to X$. So, $PT^{-1} \circ PT = \mathrm{id}$. To conclude these are inverse isomorphisms, it will be enough to show $PT^{-1}$ is injective. Suppose we have $f : S^{n+q} \to X_+ \wedge MO(n)$ so that $M^q = PT_f^{-1}$ is null-cobordant. After increasing the dimension $n$ and stabilizing our map, we may assume $M^q \subset S^{n+q}$ bounds a manifold $W^{q+1}$ which is embedded in $D^{n+q+1}$. As before $f$ is homotopic after projection to a map $g : S^{n+q} \to T\gamma_n^r$ Let $U = g^{-1}(\gamma_n^r)$, which can be taken after homotopy to be a tubular neighbourhood of $M^q$ in $S^{n+q}$ so that $g : U \to \gamma_n^r$ is a map of total spaces of vector bundles. This can be extended by a relative tubular neighbourhood theorem to a define a map from a tubular neighbourhood $V$ of $W$ to $\gamma_n^r$ which is also a map of total space of vector bundles. We can then extend this to a map $\widetilde{g} : D^{n+1} \to T\gamma_n^r$ by setting it equal to $\infty$ outside $W$ and then to a map $\widetilde{G} : D^{n+1} \to X_+ \wedge MO(n)$. Then $\widetilde{G}$ by its construction provides a null homotopy of $f$.

We could have done much the same construction for more complicated $G$-structures. In the forward direction the stable $G$-structure on the normal bundle of the singular manifold provides a map to $MG(n)$ instead of $MO(n)$. And conversely, taking the zero locus of a map into $S^{n+q} \to X_+ MG(n)$ would endow the resulting manifold $M \subset S^{n+q}$ with a $G$-structure on its normal bundle coming from an identification with a tubular neighbourhood of the zero section of $EG(n)$. We have thus proved the following fundamental result.

**Theorem 6.61: Pontryagin–Thom Construction**

There is an isomorphism,

$$\Omega_q^O(X) = \lim_{n\to\infty}[S^{n+q}, X_+ \wedge MO(n)].$$

Similarly, there are isomorphisms,

$$\Omega_q^{SO}(X) = \lim_{n\to\infty}[S^{n+q}, X_+ \wedge MSO(n)],$$
$$\Omega_q^U(X) = \lim_{n\to\infty}[S^{2n+q}, X_+ \wedge MU(n)].$$

Analogous results hold for any nice topological group $G$ (e.g. $G = Sp, Spin, SU, etc.$).

In the case where $X$ is a point, we obtain the $n$-dimensional bordism ring. And in fact we can recover the ring structure from this construction. Using the multiplication maps we defined for $MO$, there is a natural map,

$$[S^{n+q}, MO(n)] \times [S^{m+r}, MO(m)] \xrightarrow{\wedge} [S^{n+q} \wedge S^{m+r}, MO(n) \wedge MO(m)]$$
$$\to [S^{n+m+q+r}, MO(n+m)].$$

We claim this coincides with the multiplication in $\Omega_*^O$ given by Cartesian product. Indeed, if we have two manifolds $N, M$, embedded into $S^{n+q}$ and $S^{m+r}$ respectively, then the product $N \times M$ can be embedded into $S^{n+q+m+r}$. The normal bundle $\nu_{N\times M}$ of this embedding is the direct sum of the two normal bundles $\nu_N, \nu_M$. And so the Pontryagin–Thom collapse map sends $S^{n+q+m+r}$ to $T\nu_{N\times M}$ and then to $MO(n+m)$ precisely via the Thom multiplication map $MO(n) \times MO(m) \to MO(n+m)$ induced by the classifying space map corresponding to direct summing vector bundles. Thus we see we have a ring isomorphism. Of course the same holds for the homology of a point in any other Thom spectrum $MG$.

Dual to our generalized homology theory is a generalized cohomology theory, the cobordism groups. Using the Pontryagin–Thom construction, these can be defined by,

$$\Omega_O^n(X) = \lim_{n\to\infty}[\Sigma^n X, MO(n+q)],$$
$$\Omega_{SO}^n(X) = \lim_{n\to\infty}[\Sigma^n X, MSO(n+q)],$$
$$\Omega_U^n(X) = \lim_{n\to\infty}[\Sigma^{2n} X, MU(n+q)],$$

and analogously for other $G$-structures. Note that there is always a graded ring map on the cobordism groups using our smash map on the Thom spectra.

In the case where $X$ is a closed $k$-manifold, there is a generalized Poincaré duality for cobordism giving,

$$\Omega_n^O(X) = \Omega_O^{k-n}(X),$$

and analogously for oriented and complex cobordism if $X$ is oriented or has a stable almost complex structure. The multiplication on cohomology then becomes a map in homology analogous to the intersection product,

$$\Omega_n^O(X) \times \Omega_m^O(X) \to \Omega_{m+n-k}^O(X).$$

This map agrees with the multiplication we described earlier coming from looking generically at the preimage of the diagonal under the product of two singular manifolds in $X$.

Thom spectra are actually quite general phenomena and we describe now how a Thom-like construction recovers a familiar generalized homology theory.

**Definition 6.62**

A *normal framing* of an embedding $i : M^n \hookrightarrow \mathbb{R}^{n+m}$ of an $n$-manifold is an isomorphism of the normal bundle $\nu_i$ with a trivial rank $m$ bundle.

A *stable normal framing* of a $n$-manifold $M$ is an isomorphism class of normal framings up to homotopy of framings and stabilization. I.e. given two embeddings $i : M \to \mathbb{R}^{n+a}$ and $j : M \to \mathbb{R}^{n+b}$ with normal framings, we identify their framings if $\nu_i \oplus \underline{\mathbb{R}}^k \cong \underline{\mathbb{R}}^{a+k}$ and $\nu_j \oplus \underline{\mathbb{R}}^\ell \cong \underline{\mathbb{R}}^{b+\ell}$ are homotopic for some $k$ and $\ell$. A manifold is called *stably framed* if it is equipped with a choice of stable normal framing.

We may define the *framed bordism groups* of a space $X$,

$\Omega_n^{\text{fr}}(X) = \{\text{stably framed singular } n\text{-manifold}\}/\{\text{stably framed bordism}\}.$

That is, we consider singular $n$-manifolds $M \to X$ with $M$ stably framed and we identify two such maps if there is a stably framed singular $(n+1)$-manifold with boundary $W \to X$ which restricts to the two singular manifolds with their given framings on the boundary.

Note that not all manifolds possess a normal framing. By the Whitney duality formula one needs all the characteristic classes of the manifold to vanish, in particular $M$ should be orientable and spin.

Now consider the analogous Pontryagin–Thom construction for the framed bordism group. If we begin with a stably framed singular $q$-manifold $f : M^q \to X$ and a choice of embedding $i : M^q \to \mathbb{R}^{n+q}$ with a normal framing

$\nu_i \cong \underline{\mathbb{R}^n}$ , we obtain a Pontryagin–Thom collapse map,

$$PT_f : S^{n+q} \to X_+ \wedge T\nu_i$$

Because we have a preferred trivialization of $\nu_i$, we have an isomorphism $\nu_i \to M \times \mathbb{R}^n$. Taking the Thom space, this becomes an isomorphism $T\nu_i \to S^n \wedge M_+$. We may compose the above map with this one and project out the factor of $M$ to get the following,

$$PT_f : S^{n+q} \to X_+ \wedge S^n = \Sigma^n X.$$

In the same manner as we showed before, up to bordisms of $f$ and a different choice of normal framing in the same stable normal framing class, the homotopy class of the stabilized version of $PT_f$ will not change. I.e. we obtain a well defined element of,

$$\lim_{n\to\infty} [S^{n+q}, \Sigma^n X] = \lim_{n\to\infty} \pi_{n+q}(\Sigma^n X) = \pi_q^{\mathrm{st}}(X).$$

Further, we can show this map is a homomorphism and we may obtain by the same procedure an inverse map. Hence one has the following result.

**Theorem 6.63**

There is an isomorphism,

$$\Omega_n^{\mathrm{fr}}(X) = \pi_n^{\mathrm{st}}(X).$$

Stably framed bordism defines the same generalized homology theory as stable homotopy. In particular, we have,

$$\Omega_n^{\mathrm{fr}}(*) = \pi_n^{\mathrm{st}},$$

so the stable homotopy groups of spheres agree with the stably framed cobordism groups.

In general, understanding bordism groups is difficult so this is not an effective method for computing homotopy groups. Nevertheless, it does suggest deeper connections between cobordism theory and stable homotopy theory that allow for theoretical studies.

Let us conclude by showing how this theorem lets us compute stable homotopy groups of spheres in low dimensions. A normal framing of a connected 0-manifold, i.e. a point, is just an element of $GL_n(\mathbb{R})$. Up to homotopy, there are two such stable normal framings given by the determinant of the framing of the point. A line segment is a cobordism between a pair of points with opposite framings. Hence $\Omega_0^{\mathrm{fr}}(\mathrm{pt}) = \pi_0^{\mathrm{st}} = \mathbb{Z}$ counting the number of positively framed minus negatively framed points.

There are four stable normal framings of a circle. This can be seen from the fact $\pi_0 \mathrm{Map}(S^1, GL(\mathbb{R}^\infty)) = \mathbb{Z}_2 \times \pi_1(GL(\mathbb{R}^\infty)) = \mathbb{Z}_2 \times \mathbb{Z}_2$.[3] There are two positively oriented framings, one coming from the standard embedding $i : S^1 \hookrightarrow \mathbb{R}^2$ and the other coming from the left invariant framing of $S^1$ as a Lie group (note this second is a tangent framing, which will be equivalent to a normal framing). The second framing can be thought of as associated to the embedding $j : S^1 \hookrightarrow \mathbb{R}^3$ as a figure-eight. There are two negatively-oriented framings given by multiplying these two framings by a sign.

The stable framing from $i$ is clearly null-bordant. And the stable framing from $j$ is 2-torsion in the bordism group using a cylinder bordism. Also multiple copies of the circle are always bordant to a single copy using a pair of pants with many leg holes. We conclude that the framed bordism group consists of two elements: the identity element associated to $i$ and the non-trivial element associated to the Lie framing or $j$. Hence, $\Omega_1^{\mathrm{fr}}(\mathrm{pt}) = \pi_1^{\mathrm{st}} = \mathbb{Z}_2$.

Proving that $\Omega_2^{\mathrm{fr}}(\mathrm{pt}) = \pi_2^{\mathrm{st}} = \mathbb{Z}_2$ is a little more delicate. In fact, when Pontryagin first published this construction, he erroneously computed that $\Omega_2^{\mathrm{fr}}(\mathrm{pt})$ was trivial. A generator of the second framed cobordism group is the torus $T^2$ with its Lie group framing. Since this is the product in the framed cobordism ring of two copies of $S^1$ with its Lie framing, we know it must be two-torsion. The fact that it is a non-vanishing framed cobordism class, and that there are no other generators of the group requires more work, which we do not explain here. The standard proof comes via defining the topological Arf invariant. It turns out that the Lie group framing of $S^1$ has multiplicative order four in the framed cobordism ring (it's order can be no bigger since $\pi_4^{\mathrm{st}} = 0$).

The generator of $\Omega_3^{\mathrm{fr}}(\mathrm{pt}) = \pi_3^{\mathrm{st}} = \mathbb{Z}_{24}$ also comes from a Lie group framing, namely that of $S^3 = Sp(1)$. There is a delightful argument that this generator is 24-torsion attributed to Atiyah. Recall the *K3 surfaces* are an important class of compact Kähler surfaces, uniquely specified by the fact they are simply connected with trivial canonical bundle. All such surfaces are homeomorphic; a basic example is the Fermat quartic given as the vanishing locus of $x^4 + y^4 + z^4 + w^4$ in $\mathbb{C}P^3$. It is a famous computation to work out the Hodge diamond of a K3, and one finds its cohomology is concentrated in even degrees with Betti numbers $b_0 = b_4 = 1$ and $b_2 = 22$. Hence, $\chi(K3) = 24$.

Consider a vector field $\xi$ on a smooth $K3$ surface with only non-degenerate and positively oriented zeroes (there will be 24 such zero by the Euler characteristic computation). Let $M$ be the 4-manifold obtained by removing small balls

[3] We are using that stable normal framings on $X$ are a torsor for homotopy classes of isomorphisms of a trivial bundle on $X$ up to stabilization. This is given by $\pi_0 \mathrm{Map}(X, GL(\mathbb{R}^\infty))$.

around each zero of $\xi$ in the $K3$ surface; the boundary of $M$ consists of 24 disjoint 3-spheres. K3 surfaces are in fact hyper-Kähler and thus there are three complex multiplication operators $I, J, K$ on the tangent bundle of our surface. Hence, $\{\xi, I\xi, J\xi, K\xi\}$ provides a framing of the tangent bundle of $M$. We can arrange it so that $\xi$ will be normal to each boundary 3-sphere of $M$ and so then each boundary component of $M$ is given a stabilization of the $Sp(1)$ Lie framing $\{I\xi, J\xi, K\xi\}$. Hence $M$ provides a null-cobordism of 24 copies of the Lie framing on $S^3$, and we see this generator must be 24-torsion.

So far, all the generators of stable homotopy groups have been from Lie groups framings. The next two stable homotopy groups are trivial and one also has that $\Omega_6^{\mathrm{fr}} = \pi_6^{\mathrm{st}} = \mathbb{Z}_2$ is generated by the Lie framing on $Sp(1) \times Sp(1)$. However in dimension 7, $\Omega_7^{\mathrm{fr}} = \pi_7^{\mathrm{st}} = \mathbb{Z}_{240}$ is generated by the Hopf map $S^{15} \to S^8$ which does not correspond to a Lie framing. One can see this as arising from the failure of their to be a Lie group structure on $S^7$ as an octonionic analogue to the unit quaternion group $Sp(1) = S^3$.

Many other generators in low dimensions arise from Lie framings, for example $SU(3), U(3), G_2$ and $U(1) \times G_2$ give rise to 2-torsion in groups 8,9,14, and 15 respectively, and $Sp(2)$ and $Sp(1) \times Sp(2)$ give rise to 3-torsion in groups 10 and 13. However, there are many other components of these stable homotopy groups which arise from more sophisticated forms of cobordism, versions of the $J$-homomorphism, and more mysterious non-geometric sources.

CHAPTER 3

# Partial Differential Equations

## 1. Functional Analysis

**1.1. Important Facts.** For reference, we review a few important fundamental results without proofs that I often forget.

**Definition 1.1: Topologies**

- Let $E$ be a Banach space. The usual topology on $E$ is called the *strong topology*. One has $x_n \to x$ strongly iff $\|x_n - x\| \to 0$.
- The *weak topology* on $E$ is the coarsest topology so that the linear functionals are continuous. One has $x_n \rightharpoonup x$ weakly iff $\varphi(x_n - x) \to 0$ for all $\varphi \in E^*$.
- The *weak* topology* on $E^*$ is the coarsest topology so that the linear functionals coming from $E \subset E^{**}$ are continuous. One has $\varphi_n \overset{*}{\rightharpoonup} \varphi$ in the weak* topology iff $\varphi_n(x) \to \varphi(x)$ for all $x \in E$.
- A Banach space $A$ is *separable* if it has a countable dense subset. It is *reflexive* if $X^{**} \cong X$ (via the tautological inclusion).
- A Banach space $E$ is *uniformly convex* if, $\forall \varepsilon \geq 0$ $\exists ¿ 0$ so that for $x, y$ in the unit ball of $E$, $\|x - y\| > \varepsilon$ implies $\|x + y\| \leq 2 - \delta$.

**Fundamental Theorems**

**Analytic Hanh–Banach:** Let $f : E \to \mathbb{R}$ be a sublinear functional on a topological vector space, i.e. $f(\lambda x) = \lambda f(x)$ and $f(x+y) \leq f(x)+f(y)$. Suppose $g : G \subset \mathbb{R}$ is a linear functional defined on a subspace $G \subset E$ and $g(x) \leq f(x)$. Then $g$ may be extended to a linear functional $\widetilde{g} : E \to \mathbb{R}$ so that $\widetilde{g}(x) \leq f(x)$.

**Geometric Hanh–Banach:** Given disjoint non-empty convex subsets $A, B \subset E$ at least one of which is open, $A$ and $B$ are separated by a closed affine hyperplane.

**Krein–Milman:** If $K \subset E$ is a closed convex subset of a normed vector space, $K$ is the closed convex hull of its *extremal points* (points that can't be written as a weighted sum of other points of $K$).

---

For the remaining theorems assume the spaces are Banach.

**Uniform Boundedness Principle:** Given a family $T_i$ of bounded linear operators, if for each $x$ of unit norm $\{\|T_i x\|\}$ is bounded, then it is bounded independent of $x$. Thus the $T_i$ have uniformly bounded operator norm.

**Banach–Steinhaus:** If a sequence of bounded linear operators converges pointwise on a dense set and is pointwise bounded, then the pointwise limit is a globally defined linear operator.

**Open Mapping Theorem:** Let $L$ be a surjective bounded linear operator. The image under $L$ of the unit ball contains an open ball around zero. Hence $L$ is open, and if $L$ is bijective its inverse is bounded.

**Closed Graph Theorem:** A linear map $L : E \to F$ is continuous iff its graph is closed in $E \times F$.

**Convexity:** A convex set in $E$ is weakly open if and only if it is strongly open.

**Banach–Alaoglu:** The closed unit ball in $E^*$ is always weak*-compact. Hence, weak-compact if $E$ is reflexive.

**Kakutani:** If the closed unit ball of $E$ is weak-compact, then $E$ is reflexive.

**Mazur:** If $E$ is reflexive, any convex closed subset is weakly closed.

**Milman–Pettis:** Every uniformly convex space $E$ is reflexive.

## Important $L^p$ Inequalities

- **Young's Inequality:** For $a, b > 0$ and $p, q \in (1, \infty)$ *Hölder conjugate*, i.e. $\frac{1}{p} + \frac{1}{q} = 1$,
$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q.$$
- **Minkowski's Inequality:** The $L^p$ norms satisfy the triangle inequality.

- **Hölder's inequality:** If $f \in L^p$ and $g \in L^q$ with $p, q \in [1, \infty]$ Hölder conjugate,
$$\|fg\|_{L^1} \leq \|f\|_{L^p} \cdot \|g\|_{L^q}.$$
- **Young's Convolution Inequality**: If $f \in L^1(\mathbb{R}^N)$ and $g \in L^p(\mathbb{R}^N)$, then
$$\|f * g\|_{L^p} \leq \|f\|_{L^1} \cdot \|g\|_{L^p}.$$
More generally if $f \in L^p$ and $g \in L^q$ with $1/p + 1/q = 1/r \geq 1$, then
$$\|f * g\|_{L^r} \leq \|f\|_{L^r} \cdot \|g\|_{L^p}.$$
- **Finite Spaces:** On a finite measure space, if $p \leq q \in [1, \infty]$, then $L^q \subset L^p$.
- **Interpolation Inequality:** Let $1 \leq s \leq r \leq t \leq \infty$, and $\frac{1}{r} = \frac{\theta}{s} + \frac{1-\theta}{t}$. If $u \in L^s \cap L^t$, then $s \in L^r$ and
$$\|u\|_{L^r} \leq \|u\|_{L^s}^{\theta} \|u\|_{L^t}^{1-\theta}.$$

For $\Omega \subset \mathbb{R}^N$, a separable measurable subset, the spaces $L^p(\Omega)$ have the following properties [**Br**, §4.3].

| | Reflexive | Separable | Dual Space |
|---|---|---|---|
| $L^p$, $p \in (1, \infty)$ | Yes | Yes | $L^q$ where $1 = 1/p + 1/q$ |
| $L^1$ | No | Yes | $L^\infty$ |
| $L^\infty$ | No | No | Strictly bigger than $L^1$ |

Separability follows easily by taking $\mathbb{Q}$-linear combinations of characteristic functions on a countable basis of $\Omega$'s topology. Reflexivity follows from Milman–Pettis. The dual pairing is given by a version of Riesz representation (for $L^2$ it is verbatim Riesz representation) and takes the form,
$$\langle f, g \rangle = \int_\Omega f \cdot g \, \mathrm{d}\mu \quad \text{where} \quad f \in L^p, \ g \in L^q.$$
Finiteness of this expression is guaranteed by Hölder. For finite $p$, every functional on $L^p$ is of this form; but that is not true for $p = \infty$.

Similarly for the sequence spaces [**Br**, §11.3].

| | Reflexive | Separable | Dual Space |
|---|---|---|---|
| $\ell^p$, $p \in (1, \infty)$ | Yes | Yes | $\ell^q$ where $1 = 1/p + 1/q$ |
| $\ell^1$ | No | Yes | $\ell^\infty$ |
| $c_0$ | No | Yes | $\ell^1$ |
| $c$ | No | Yes | $\ell^1 \times \mathbb{R}$ |
| $\ell^\infty$ | No | No | Strictly bigger than $\ell^1$ |

Here $c_0 \subset c$ are bounded sequences that converge to zero or something finite respectively.

Recall the Arzelà–Ascoli theorem characterizes precompact subsets of the space of continuous functions on a compact metric space as those which are bounded and uniformly equicontinuous. One can generalize to spaces of $L^p$ functions using an $L^p$ version of equicontinuity.

**Theorem 1.2: Fréchet–Kolmogorov** [Br, Thm. 4.26, Cor. 4.27]

Let $\mathcal{F} \subset L^p(\mathbb{R}^N)$ be bounded and $p \in [1, \infty)$. If

(equicontinuity) $\lim_{|h|\to 0} \|f(\cdot + h) - f(\cdot)\|_{L^p} = 0$ uniformly in $f \in \mathcal{F}$,

then $\mathcal{F}|_\Omega$ is precompact in $L^p(\Omega)$ for any finite measure $\Omega \subset \mathbb{R}^N$. Moreover if

(equitightness) $\lim_{r\to\infty} \int_{|x|>r} |f|^p \, d\mu = 0$ uniformly in $f \in \mathcal{F}$,

then $\mathcal{F}$ is precompact in $L^p(\mathbb{R}^N)$.

**1.2. Riesz Representation and Lax–Milgram.** We can identify Hilbert spaces with their dual:

**Theorem 1.3: Riesz Representation** [Br, Thm. 5.5]

Let $H$ be a Hilbert space and $\varphi \in H^*$. Then there is a unique $u \in H$ so that $\varphi(v) = \langle u, v\rangle$ for all $v \in H$; also, $|v| = \|\varphi\|_{H^*}$.

Proof. Let $M \subset H$ be the closed (proper if $\varphi \neq 0$) subspace of vectors annihilated by $\varphi$. Let $a$ be a unit norm vector in the orthogonal complement of $M$. Then $u = \varphi(a)a$ works since,

$$v - \frac{\varphi(v)}{\varphi(a)^2} u \in M \quad \text{hence} \quad \langle u, v\rangle = \varphi(v).$$

Uniqueness is clear. □

The following result will be important for solving linear elliptic PDEs. One can view it as saying that the conclusion of the above theorem didn't require using the inner product; any sufficiently nice bilinear form will do, in the following sense.

**Definition 1.4**

Let $A$ be a bilinear form on a Hilbert space $H$. $A$ is *continuous* if,

$$|A(u,v)| \leq C|u| \cdot |v|,$$

and *coercive* if,

$$|A(u,u)| \geq C'|u|^2.$$

**Theorem 1.5: Lax–Milgram** [Br, Cor. 5.8]

If $A$ is continuous and coercive, then for any $\varphi \in H^*$ there is a unique $u \in H$ so that,

$$\varphi(v) = A(u,v) \quad \forall v \in H.$$

Moreover if $A$ is symmetric, then $u$ is characterized as the unique minimum of the functional $\frac{1}{2}A(v,v) - \varphi(v)$.

Proof. Define $F : H \to H$ so that $A(u,v) = \langle F(u), v\rangle$. Such an $F$ exists by Riesz representation and is clearly linear and continuous (by continuity of $A$). Lax–Milgram is then the statement that $F$ is a bijection.

$F$ must be injective since coercivity implies its kernel is trivial. The image of $F$ is closed since $C'|v| \leq |F(v)|$ by coercivity and Cauchy-Schwarz (consider convergent sequence $F(v_i)$ and then apply coercivity to show $v_i \to v$ and hence $F(v_i) \to F(v)$ by continuity). The image of $F$ is also dense since $\langle F(u), v\rangle = 0$ for all $u$ implies $v = 0$ (so $\overline{\mathrm{im}(F)} = (\mathrm{im}(F)^\perp)^\perp = H$). Hence $F$ is surjective.

The characterisation in the symmetric case is just elementary calculus of variations. □

**1.3. Compact Operators.** We consider some special operators on Hilbert spaces that are almost finite.

**Definition 1.6**

A bounded linear operator $T : G \to H$ is *compact* if the image of the unit ball is precompact in $H$ (in the strong topology). $T$ is *finite rank* if its range is finite dimensional. We will write $\mathcal{L}(G,H)$, $\mathcal{K}(G,H)$, and $\mathcal{R}(G,H)$ for the spaces of bounded, compact, and finite rank operators respectively.

Note $\mathcal{R}(G,H) \subset \mathcal{K}(G,H)$.

PROPOSITION 1.7. *The compact operators are closed in the strong topology. Hence, convergent sequences of finite rank operators converge to compact operators. If $G, H$ are Hilbert spaces, then $\mathcal{K}(G,H) = \overline{\mathcal{R}(G,H)}$.*

LEMMA 1.8 (Riesz). *The unit ball of a normed vector space $E$ is compact iff $E$ is finite.*

PROOF. One direction is clear. If $E$ is infinite, inductively construct a sequence $\{u_n\}$ of unit norm vectors as follows. Write $E_n = \bigoplus_{i<n} \operatorname{span}(u_i)$. Choose $u_n$ so that $d(u_n, E_n) \geq 1/2$. This can be done by picking any $v$ outside of $E_n$ (alright since $E$ infinite) and letting $u = (v-x)/\|v-x\|$ where $x \in E_n$ has $\|v-x\| \leq 2d(v, E_n)$. Then for any $y \in E_n$,

$$\|u - y\| = \left\| \frac{v-x}{\|v-x\|} - y \right\| \geq \frac{d(v, E_n)}{\|v-x\|} \geq 1/2.$$

Now clearly $(u_n)$ is a unit norm sequence with no Cauchy subsequence. □

**Theorem 1.9: Fredholm Alternative** [Br, Thm. 6.6]

Let $T \in \mathcal{K}(E)$. Then,
(a) $\ker(I-T)$ is finite.
(b) $\operatorname{im}(I-T) = \ker(I-T^*)^\perp$ is closed.
(c) $\ker(I-T) = 0$ iff $\operatorname{im}(I-T) = E$.
(d) $\dim \ker(I-T) = \dim \ker(I-T^*)$.

REMARK 1.10. What one should really remember is the following:

(1) The Fredholm alternative is commonly invoked in solving inhomogeneous linear problems as follows. Let $L = I - T$ as above. Then $Lu = f$ has a solution iff $f \in \ker(L^*)^\perp$.

(2) In the language we will shortly develop, the remaining substance of the above theorem is that compact perturbations of the identity are Fredholm of index zero.

△

PROOF. We write $L = I - T$. We just prove the first couple of these claims.

(a) The $F \subset E$ be the kernel of $L$. If $B_E$ is the unit ball of $E$, $B_E \cap F = T(B_E \cap F) \subset T(B_E)$. So the unit ball of $F$ is compact. By Lemma 1.8, $F$ is finite dimensional.

(b) Suppose $x_n = L(u_n)$ converges to $x$; we want $x \in \mathrm{im}(L)$. Let $d_n = d(u_n, \ker(L))$. this kernel is finite hence closed and so $d_n = \|u_n - v_n\|$ for some $v_n \in \ker(L)$. Note $x_n = L(u_n - v_n)$.

If $\|u_n - v_n\|$ was unbounded, without loss of generality we could work with a subsequence diverging to $\infty$. Then $w_n = (u_n - v_n)/\|u_n - v_n\|$ satisfies $L(w_n) = x_n/\|u_n - v_n\|$ which must go to zero. Since $T$ is compact, we may pass again to a subsequence so that the $w_n$ and hence the $T(w_n)$ converge to some $w \in ker(L)$. So on one hand $d(w_n, \ker(L)) \to d(w, \ker(L)) = 0$ while on the other,

$$d(w_n, \ker(L)) = \frac{d(u_n \ker(L))}{\|u_n - v_n\|} = \frac{d_n}{\|u_n - v_n\|} = 1.$$

We obtain a contradiction; hence $\|u_n - v_n\|$ is bounded and by passing to a subsequence we may assume $T(u_n - v_n)$ converges to $y$. Then,

$$u_n - v_n = (L + T)(u_n - v_n) \to x + y.$$

Note $L(u_n - v_n)$ must converge to $L(x + y)$ but above we said it converged to $x$. Hence $x = L(x + y) \in \mathrm{im}(L)$.

The equality is trivial noting $\mathrm{im}(L)^\perp = \ker(L^*)$ for any (unbounded) operator $L$.

□

**Definition 1.11**

The (real) *spectrum* $\sigma(T)$ of $T \in \mathcal{L}(E)$ is the set of $\lambda \in \mathbb{R}$ so that $T - \lambda I$ fails to be bijective. If $T$ is a $\mathbb{C}$-linear operator, the complex spectrum is defined analogously.

PROPOSITION 1.12. *The spectrum of a bounded operator $T$ on $E$ is compact and bounded by $\|T\|$. The complex spectrum is non-empty.*

PROOF. If $|\lambda| > \|T\|$, then the contraction mapping principle implies there is a unique solution to $u = \lambda^{-1}(Tu - f)$ for any $f \in E$.

If $Tu - \lambda_0 u = f$ has a unique solution, then applying the contraction mapping principle to,

$$u = (T - \lambda_0 I)^{-1}[f + (\lambda - \lambda_0)u],$$

we see $Tu - \lambda u = f$ has a unique solution for $\lambda$ close enough to $\lambda_0$ (independent of $f$). Hence the complement of the spectrum is open and thus $\sigma(T)$ is compact.

Over $\mathbb{C}$, $(T-\lambda I)^{-1}$ is a bounded holomorphic function of $\lambda$ on $\mathbb{C}\setminus\sigma(T)$. By Liouville's theorem, $\sigma(T)$ must be non-empty. $\square$

PROPOSITION 1.13. *Let* $\dim E=\infty$. *If* $T\in\mathcal{K}(E)$, *then* $0\in\sigma(T)$ *and* $0$ *is the only possible accumulation point of* $\sigma(T)$. *Hence* $\sigma(T)$ *is either finite or a sequence converging to* $0$.

PROOF. If $T$ were surjective, the image of the unit ball would contain the closed ball of radius $\|T\|$. Hence this ball would be compact, which contradicts Lemma 1.8 given that $E$ is infinite.

Note that any non-zero $\lambda\in\sigma(T)$ iff $\ker(T-\lambda I)\neq 0$, since if $T-\lambda I$ is injective (c) of the Fredholm alternative implies it's also surjective. Fix $\lambda_0>0$ and let $F\subset E$ be the subspace spanned by all vectors in the kernel of $\ker(T-\lambda I)$ for $|\lambda|>\lambda_0$. The image of the unit ball of $F$ under $T$ contains the closed ball in $F$ of radius $\lambda_0/2$. Hence the unit ball of $F$ is compact and by Lemma 1.8, $F$ is finite. In particular there are only finitely many $|\lambda|>\lambda_0$ in the spectrum of $T$ and no $|\lambda|>\lambda_0$ can be an accumulation point of $T$. $\square$

**Theorem 1.14: Spectral Theorem for Compact Self-Adjoint Operators**

Let $H$ be a separable Hilbert space and $T\in\mathcal{L}(H)$ compact self-adjoint. Then $H$ has an orthonormal (Hilbert) basis of eigenvectors of $T$.

PROOF. If $H$ is finite, we are done by the usual spectral theorem. Otherwise, Proposition 1.13 implies the eigenvectors of $T$ forms a possibly finite sequence $(\lambda_n)$.

Note the eigenspaces of $\lambda_n\neq 0$ are finite dimensional by the above proof. Further, the eigenspaces are mutually orthogonal by the standard argument. Since $H$ is separable, $\ker(T)$ has an orthonormal basis. We can also pick an orthonormal basis for the finite dimensional eigenspaces of every $\lambda_n\neq 0$. Thus we obtain an orthonormal basis for the subspace $G$ defined as the direct sum of the eigenspaces.

It now suffices to prove the closure of $G$ is $H$. We have $T(G)\subset G$ and so $T(G^\perp)\subset G^\perp$ by self-adjointness. Note $T$ restricted to $G^\perp$ is compact self-adjoint with no eigenvalues (since all the eigenvalues of $T$ lie in $G$ and $G\cap G^\perp=\{0\}$).

We claim if $S$ is a compact self-adjoint operator, $M=\sup_{\|u\|=1}\langle S(u),u\rangle$ is an eigenvalue of $S$. By Cauchy-Schwarz (applied to symmetric bilinear form

$\langle Mu - Su, v\rangle$),

$$|\langle Mu - Su, v\rangle| \leq \sqrt{\langle Mu - Su, u\rangle}\sqrt{\langle Mv - Sv, v\rangle}.$$
$$\|Mu - Su\|^2 \leq \sqrt{\langle Mu - Su, u\rangle}\sqrt{\langle M^2u - S^2u, Mu - Su\rangle}.$$
$$\leq \|MI - S\|_{op} \cdot \langle Mu - Su, u\rangle.$$

We can chose a sequence $(u_n)$ of unit vectors so that $\langle Su_n, u_n\rangle \to M$ and hence $\|Mu_n - Su_n\| \to 0$. But then $M \in \sigma(S)$, since otherwise $\|u_n\| = (MI - S)^{-1}(Mu_n - Su_n) \to 0$ which yields a contradiction. Since $M \neq 0$, the proof of Proposition 1.13 yields $M$ is an eigenvalue.

Applying this result to $S = T|_{G^\perp}$, we obtain a contradiction unless $G^\perp$ is zero-dimensional. Hence $G^\perp = 0$ and $\overline{G} = H$. □

**1.4. Fredholm Operators.** We briefly discuss these important class of operators connected to the study of elliptic PDEs.

**Definition 1.15**

An operator $T \in \mathcal{L}(E, F)$ is *Fredholm* if $\ker(T)$ is finite dimensional and $\mathrm{im}(T)$ is finite co-dimensional. It is often assumed $\mathrm{im}(T)$ is closed, but this is automatic. We write $\mathcal{F}(E, F)$ for the space of Fredholm operators. In this case, the *index* of $T$ is given as,

$$\mathrm{ind}(T) = \dim \ker(T) - \dim \mathrm{coker}(T).$$

We summarize some important facts about Fredholm operators.

**Fredholm Operators**

- Fredholm operators are precisely those which are invertible modulo finite rank operators.
- $\mathcal{F}(E, F)$ is an open subset of $\mathcal{L}(E, F)$ and the index is locally constant.
- Compact perturbations of a Fredholm operator $T$ are Fredholm with the same index as $T$.
- The composition of Fredholm operators is Fredholm and the index is additive.

## 2. Sobolev Spaces

To apply functional analysis to the study of PDEs, we should study differential operators acting on suitable Banach spaces. Sobolev spaces provide these as spaces consisting of $L^p$ functions that can be differentiated in some sense.

**2.1. Basic Properties.** Let $\Omega \subset \mathbb{R}^N$ be open and consider $u \in L^p(\Omega)$. We say $u$ has a *weak derivative* $\nabla u = (g_1, \dots, g_N)$ if,

$$\int_\Omega u \frac{\partial \varphi}{\partial x_i} = -\int_\Omega g_i \varphi \quad \text{for all} \quad \varphi \in C_c^\infty(\Omega).$$

That is, $u$ and the $g_i$ satisfy integration by parts for any test function $\varphi$.

**Definition 2.1**

The *Sobolev space* $W^{1,p}(\Omega)$ is the space of functions $u \in L^p(\Omega)$ whose weak derivative $\nabla u = (\frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_N})$ exists with all $\frac{\partial u}{\partial x_i} \in L^p(\Omega)$. We equip these spaces with the *Sobolev norm*,

$$\|u\|_{W^{1,p}} = \|u\|_p + \sum_{i=1}^N \left\| \frac{\partial u}{\partial x_i} \right\|_p.$$

We write $H^1$ for $W^{1,2}$ and equip it with the inner product,

$$\langle u, v \rangle_{H^1} = \langle u, v \rangle_{L^1} + \sum_{i=1}^N \left\langle \frac{\partial u}{\partial x_i}, \frac{\partial v}{\partial x_i} \right\rangle_{L^2} = \int_\Omega uv + \sum_{i=1}^N \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_i}.$$

PROPOSITION 2.2. *$W^{1,p}(\Omega)$ is a Banach space for $p \in [1, \infty]$. It is reflexive if $p \neq 1, \infty$ and separable for $p \neq \infty$. $H^1(\Omega)$ is a separable Hilbert space.*

PROOF. This follows relatively straightforwardly from the same results for $L^p(\Omega)$. □

**Sobolev Facts**

- If $u \in C^1(\Omega)$ and $u$ and its derivatives are in $L^p(\Omega)$, then $u \in W^{1,p}(\Omega)$ and the weak derivatives are the actual derivatives.
- If $(u_n)$ is a sequence in $W^{1,p}$ so that $u_n \to u$ in $L^p$ and $\nabla u_n$ converges in $(L^p)^N$, then $u \in W^{1,p}$ is the limit of $u_n$. If $p > 1$, we just need $\nabla u_n$ bounded.

Here's a very general density result.

PROPOSITION 2.3 (Meyers–Serrin). *Let $u \in W^{1,p}(\Omega)$ for $p$ finite. Then there is a sequence $u_n \in C_c^\infty(\mathbb{R}^N)$ for which $u_n$ converges to $u$ in $L^p(\Omega)$ and $\nabla u_n$ converges to $\nabla u$ in $L^p(\omega)^N$ for any open $\omega$ which is precompact and has $\overline{\omega} \subset \Omega$. (We say $\omega \Subset \Omega$ is* strongly included*).*

PROOF. A proof is in [Br, pp. 266–267]. We'll skip it in favour of proving a result later. □

**Theorem 2.4**

We generalize results of calculus to $W^{1,p}(\Omega)$ for any $p \in [1, \infty]$.

- If $u, v \in L^p(\Omega) \cap L^\infty(\Omega)$, $uv \in W^{1,p}(\Omega)$ and the usual product rule holds for weak derivatives.
- If $u \in W^{1,p}(\Omega)$ and $G \in C^1(\mathbb{R})$ with $G(0) = 0$ and bounded derivative, then $G \circ u \in W^{1,p}(\Omega)$ and the usual chain rule holds for weak derivatives.
- If $H : \Omega' \to \Omega$ is a $C^1$ bijection with $C^1$ inverse so that the Jacobians of $H, H^{-1}$ are in $L^\infty$, then for $u \in W^{1,p}(\Omega)$, $u \circ H \in W^{1,p}(\Omega')$ and the usual change of variables formula holds.

PROOF. In each case, we approximate our $W^{1,p}$ functions with functions in $C_c^\infty$ using Proposition 2.3. We then use the corresponding results for smooth functions and apply dominated convergence to conclude. □

**Definition 2.5**

We can inductively define higher order *Sobolev space* $W^{m,p}(\Omega)$ for $p \in [1, \infty]$ and $m \in \mathbb{N}$ as the space of $W^{m-1,p}(\Omega)$ functions with weak derivatives in $W^{m-1,p}(\Omega)$. In this case we can define weak derivatives of up to $m$th order. The Sobolev norm,

$$\|u\|_{W^{m,p}} = \sum_{0 \le |\alpha| \le m} \|D^\alpha u\|_{L^p},$$

here written in multi-index notation, makes $W^{m,p}(\Omega)$ a Banach space and for $p = 2$ the corresponding inner product makes $H^m(\Omega) = W^{m,2}(\Omega)$ a Hilbert space.

**2.2. Extension Operators.** We now study ways to extend functions in $W^{1,p}(\Omega)$ to $\mathbb{R}^N$, where its properties are more tractable. We being with a method of reflecting functions, for which we need some notation.

We will decompose $x = (x', x_N) \in \mathbb{R}^N$ with $x' \in \mathbb{R}^{N-1}$ the projection of $x$. We denote $\mathbb{R}^N_+$ the right half-space with $x_N > 0$ and $Q$ the polydisc of those $x$ with $|x'|, |x_N| < 1$. We also take $Q_+ = Q \cap \mathbb{R}^N_+$ the right half-polydisc and $Q_0 = Q \cap \{x_N = 0\}$.

LEMMA 2.6. *Let $u \in W^{1,p}(Q_+)$ for $p \in [1,\infty]$. Let $u^*$ be its even extension by reflection across $\{x_N = 0\}$. Then $u^* \in W^{1,p}(Q)$ satisfies,*

$$\|u^*\|_{L^p(Q)} \le 2\|u\|_{L^p(Q_+)} \quad \textit{and} \quad \|u^*\|_{W^{1,p}(Q)} \le 2\|u\|_{W^{1,p}(Q_+)}.$$

PROOF. Indeed we show,

$$\frac{\partial u^*}{\partial x_i} = \left(\frac{\partial u}{\partial x_i}\right)^* i = 1,\dots,N-1 \quad \text{and} \quad \frac{\partial u^*}{\partial x_N} = \left(\frac{\partial u}{\partial x_N}\right)^\circ,$$

where $f^\circ$ denotes an odd rather than even extension over the reflection.

Let $\eta \in C^\infty(\mathbb{R})$ be zero for $t < 1/2$ and one for $t > 1$. Define $\eta_k(t) = \eta(kt)$ for $k \in \mathbb{N}$. For a test function $\varphi \in C_c^1(Q)$,

$$\int_Q u^* \frac{\partial \varphi}{\partial x_i} = \int_{Q_+} u \frac{\partial \psi}{\partial x_i} \qquad \text{where} \quad \psi(x', x_N) = \varphi(x', x_N) + \varphi(x', -x_N).$$

We can make $\psi$ a test function by multiplying by $\eta_k(x_N)$,

$$-\int_{Q_+} \frac{\partial u}{\partial x_i}\eta_k\psi = \int_{Q_+} u\frac{\partial}{\partial x_i}(\eta_k\psi) = \int_{Q_+} u\eta_k \frac{\partial \psi}{\partial x_i}.$$

Taking $k \to \infty$ and applying dominated convergence,

$$\int_Q u^* \frac{\partial \varphi}{\partial x_i} = \int_{Q_+} u \frac{\partial \psi}{\partial x_i} \qquad = -\int_{Q_+} \frac{\partial u}{\partial x_i}\psi = -\int_Q \left(\frac{\partial u}{\partial x_i}\right)^* \varphi.$$

Let $\chi(x', x_N) = \varphi(x', x_N) - \varphi(x', -x_N)$. Note $|\chi(x', x_N)| \le C|x_N|$ on $Q$ since $\chi$ vanishes on $Q_0$. We have,

$$\left|\int_{Q_+} uk\eta'(kx_N)\chi\right| \le C' \int_{0<x_N<1/k} k|u|x_N \,\mathrm{d}x \le C' \int_{0<x_N<1/k} |u| \,\mathrm{d}x \to 0 \quad \text{as} \quad k \to \infty.$$

Also,

$$-\int_{Q_+} \frac{\partial u}{\partial x_N}\eta_k\chi = \int_{Q_+} u\frac{\partial}{\partial x_N}(\eta_k\chi) = \int_{Q_+} u\eta_k\frac{\partial \chi}{\partial x_N} + uk\eta'(kx_N)\chi.$$

Hence, taking $k \to \infty$ and using dominated convergence plus the above computation,

$$\int_Q u^* \frac{\partial \varphi}{\partial x_N} = \int_{Q_+} u\frac{\partial \chi}{\partial x_N} = -\int_{Q_+} \frac{\partial u}{\partial x_N}\chi = \int_Q \left(\frac{\partial u}{\partial x_N}\right)^\circ \varphi.$$

□

**Theorem 2.7: Sobolev Extension Theorem** [Br, Thm. 9.7]

Let $\Omega$ be a $C^1$ open domain so that either $\Omega = \mathbb{R}^N_+$ or $\partial\Omega$ is bounded. For $p \in [1, \infty]$, there is a linear *extension operator*,

$$P : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^N),$$

which satisfies for $u \in W^{1,p}(\Omega)$ and $C$ depending only on $\Omega$,

(i) $Pu|_\Omega = u$,
(ii) $\|Pu\|_{L^p(\mathbb{R}^N)} \leq C\|u\|_{L^p(\Omega)}$,
(iii) and $\|Pu\|_{W^{1,p}(\mathbb{R}^N)} \leq C\|u\|_{W^{1,p}(\Omega)}$.

PROOF. The previous lemma implies the result for $\mathbb{R}^N_+$, so we take $\Omega$ to be bounded. Since $\Omega$ is $C^1$ and bounded, its boundary is covered by finitely many open sets $U_1, \ldots, U_k$ with $C^1$ manifold with boundary charts $H_i : Q \to U_i$. Introduce $U_0 = \mathbb{R}^N \setminus \partial\Omega$ Let $\psi_i$ be a partition of unity subordinate to $\{U_i\}_{i=0}^k$ with all $\psi_i$ except $\psi_0$ of compact support (it will have compact support on $\Omega$).

Write $u_i = \psi_i u$. Let $\overline{u}_0$ be the extension of $u_0$ by zero. Note $\nabla\psi_0 = -\sum_{i=1}^k \nabla\psi_i$ has compact support. Hence we obtain,

$$\overline{u}_0 \in W^{1,p}(\mathbb{R}^N) \quad \text{and} \quad \frac{\partial}{\partial x_i}\overline{u}_0 = \psi_0 \overline{\frac{\partial u}{\partial x_i}} + \frac{\partial \psi_0}{\partial x_i}\overline{u}.$$

Now define $w_i$ on $U_i$ by pulling back $u$ from $U_i \cap \Omega$ to $Q_+$, extending evenly by reflection to $Q$, and then pushing forward to $U_i$. Clearly $w_i|_{U_i\cap\Omega} = u_i$ and the previous lemma plus change of variables theorem for weak derivatives gives $w_i \in W^{1,p}(U_i)$ with,

$$\|w_i\|_{W^{1,p}(U_i)} \leq C\|u\|_{W^{1,p}(U_i\cap\Omega)},$$

with $C$ depending only on the choice of manifold atlas and partition of unity.

Now let $\widehat{u}_i \in W^{1,p}(\mathbb{R}^N)$ be the extension by zero of $\psi_i w_i$. Clearly $\widehat{u}_i = u_i$ on $\Omega$ and

$$\|\widehat{u}_i\|_{W^{1,p}(\mathbb{R}^N)} \leq C\|u\|_{W^{1,p}(U_i\cap\Omega)}.$$

Then, $Pu = \overline{u}_0 + \sum_{i=1}^k \widehat{u}_i$ has all the properties of the extension operator. □

REMARK 2.8. This whole construction was quite simple. We cannot simply extend $u \in W^{1,p}(\Omega)$ to $\mathbb{R}^N$ by zero, since weak differentiability may fail at the boundary. Instead we cover $\partial\Omega$ with manifold with boundary charts, reflect $u$ across the chart boundaries and use a partition of unity to kill the reflected $u$ away from the charts. This gives a a more gentle extension of $u$ which is weak differentiable. △

This theorem gives us another density result.

**Theorem 2.9: Smooth Approximation** [Br, Cor. 9.8]

Let $\Omega$ be a $C^1$ open domain and $u \in W^{1,p}(\Omega)$ for $p \in [1, \infty)$. Then there is a sequence $(u_n)$ in $C_c^\infty(\mathbb{R}^N)$ so that $u_n|_\Omega \to u$ in $W^{1,p}(\Omega)$. Hence $C_c^\infty(\mathbb{R}^N)|_\Omega$ is a dense subspace of $W^{1,p}(\Omega)$.

PROOF. First take $\partial\Omega$ bounded. By the extension theorem, $u$ extends to an element $Pu \in W^{1,p}(\mathbb{R}^N)$. Let $f_n$ be a sequence of smooth mollifiers converging to the delta function and $g_n$ a sequence of smooth cutoff functions converging to a function which is constantly one.

Convolving $Pu$ with $f_n$ and multiplying by $g_n$ gives a sequence of smooth compactly supported functions converging to $u$ in $W^{1,p}$.

If $\partial\Omega$ is unbounded, we can multiply $u$ by the cutoffs $g_n$ to get a sequence $u_n$ of functions in $W^{1,p}$ of domains with bounded boundary. These $u_n$ approximate $u$ in $W^{1,p}$ and hence so do their smooth approximations. □

**2.3. Traces.** We introduced Sobolev spaces as a method for finding weak solutions of PDEs. We may want these weak solutions to also satisfy boundary conditions in some sense. The boundary of a domain is measure-zero so asking for it to have given values doesn't make sense in $L^p$ spaces. Instead, we now define the trace operator.

**Theorem 2.10: Trace Theorem** [Evans, Thm. 5.5.1]

Let $\Omega$ be a bounded $C^1$ open domain and $p \in [1, \infty)$. Then there exists,

$$T : W^{1,p}(\Omega) \to L^p(\partial\Omega)$$

a bounded linear operator called the *trace operator* so that,

$$Tu = u|_{\partial\Omega} \quad \text{if} \quad u \in W^{1,p}(\Omega) \cap C(\overline{\Omega}) \quad \text{and} \quad \|Tu\|_{L^p(\partial\Omega)} \leq C\|u\|_{W^{1,p}(\Omega)},$$

for $C$ depending only on $p$ and $\Omega$.

PROOF. Recalling the notation above, we begin with $u \in C^1(\overline{Q_+})$. Consider a smooth bump function $f$ on $Q$ which is non-negative and is constantly one on the polydisc of radius $1/2$, $Q_{1/2}$. Let $\Gamma$ be the intersection of $\partial Q_+$ with

$Q_{1/2}$. One has,

$$\begin{aligned}\int_\Gamma |u|^p \,\mathrm{d}x' &\leq \int_{Q_0} f|u|^p \,\mathrm{d}x' \\ &= -\int_{Q_+} \frac{\partial}{\partial x_n}(f|u|^p)\,\mathrm{d}x \\ &= -\int_{Q_+} |u|^p \frac{\partial f}{\partial x_n} + p|u|^{p-1}\mathrm{sgn}(u)\frac{\partial u}{\partial x_n} f \,\mathrm{d}x\,.\end{aligned}$$

By Young's inequality ($a^{p-1}b \leq C(a^p + b^p)$),

$$\leq C\int_{Q_+} |u|^p + |Du|^p \,\mathrm{d}x\,.$$

More generally now take $u \in C^1(\overline{\Omega})$. If $U$ is a manifold with boundary chart for $\Omega$ and $\Gamma = U \cap \partial\Omega$, then the change of variables theorem gives,

$$\int_\Gamma |u|^p \leq C\int_U |u|^p + |Du|^p \,\mathrm{d}x\,.$$

Covering $\partial\Omega$ with finitely many manifold with boundary charts $\{U_i\}_{i=1}^n$ with boundaries $\Gamma_i$, one has for each $i$,

$$\|u\|_{L^p(\Gamma_i)} \leq C\|u\|_{W^{1,p}(\Omega)}.$$

Hence setting $Tu = u|_{\partial\Omega}$,

$$\|Tu\|_{L^p(\partial\Omega)} \leq C\|u\|_{W^{1,p}(\Omega)}.$$

This $C$ depended only on $p$ and our covering of $\Omega$, not $u$.

Lastly, take a general $u \in W^{1,p}(\Omega)$. By our smooth approximation theorem, we can find $u_m \in C^\infty(\overline{\Omega})$ converging to $u$ in $W^{1,p}$. We then have from above,

$$\|Tu_m - Tu_\ell\|_{L^p(\partial\Omega)} \leq C\|u_m - u_\ell\|_{W^{1,p}(\Omega)}. \tag{2.1}$$

Clearly $Tu_m$ is then a Cauchy sequence in $L^p(\partial\Omega)$ and so converges to some limit in $L^p(\partial\Omega)$ which we call $Tu$. This limit is independent of the smooth approximation by (2.1). This extends $T$ to a linear bounded operator on $W^{1,p}(\Omega)$

If $u$ was continuous on $\overline{\Omega}$, the smooth approximation can be taken to converge uniformly to $u$ hence $Tu = u|_{\partial\Omega}$. The bound $\|Tu\|_{L^p} \leq C\|u\|_{W^{1,p}}$ follows from density of smooth functions. □

**Definition 2.11**

For $p \in [1, \infty)$ we define $W_0^{1,p}(\Omega)$ to be the closure of $C_c^1(\Omega)$ (or $C_c^\infty(\Omega)$) in $W^{1,p}(\Omega)$.

**Theorem 2.12: Trace Zero Functions in $W^{1,p}$ [Evans, Thm. 5.5.2]**

Let $\Omega$ be bounded and $C^1$. The kernel of the trace operator on $W^{1,p}(\Omega)$ is $W_0^{1,p}(\Omega)$.

REMARK 2.13. This proof is a little technical and probably not worth remembering. △

PROOF. If $u \in W_0^{1,p}(\Omega)$, then there is a sequence $u_m$ in $C_c^\infty(\Omega)$ converging to $u$ in $W^{1,p}(\Omega)$. We know $Tu_m = 0$ for each $m$ by the trace theorem. Since $T$ is continuous, $Tu = 0$.

For the converse, assume $Tu = 0$. Using the standard partitions of unit argument, we may reduce to the case where our domain is $\mathbb{R}^N_+$ with $u$ of compact support.

We can find a $C^1(\overline{\mathbb{R}^N_+})$ sequence $u_m$ approximating $u$ in $W^{1,p}(\mathbb{R}^N_+)$ which must have $Tu_m \to 0$ in $L^p(\mathbb{R}^{N-1})$. For $(x', x_N) \in \overline{\mathbb{R}^N_+}$ we have,

$$|u_m(x', x_N)| \leq |u_m(x', 0)| + \int_0^{x_n} \left| \frac{\partial}{\partial x_n} u_m(x', t) \right| \mathrm{d}t \,.$$

Hence, taking $L^p$ norms and applying Hölder to the $\mathrm{d}t$ integral,

$$\int_{\mathbb{R}^{N-1}} |u_m(x', x_N)|^p \,\mathrm{d}x' \leq C \Bigg( \int_{\mathbb{R}^{N-1}} |u_m(x', 0)|^p \,\mathrm{d}x' + x_N^{p-1} \int_0^{x_N} \int_{\mathbb{R}^{N-1}} |Du_m(x', t)|^p \,\mathrm{d}x' \,\mathrm{d}t \Bigg).$$

Taking $m \to \infty$,

$$\int_{\mathbb{R}^{N-1}} |u(x', x_N)|^p \,\mathrm{d}x' \leq C x_N^{p-1} \int_0^{x_N} \int_{\mathbb{R}^{N-1}} |Du(x', t)|^p \,\mathrm{d}x' \,\mathrm{d}t \,. \tag{2.2}$$

Now we take a smooth bump function $0 \leq \rho \leq 1$ on $\mathbb{R}$ which is one on $[0, 1]$ and zero outside $[0, 2]$. We define $\rho_m(x) = \rho(m x_N)$ and $w_m = u(x)(1 - \rho_m)$. We compute derivatives,

$$\frac{\partial}{\partial x_N} w_m = \frac{\partial u}{\partial x_N} - m u \rho' \quad \text{and} \quad D_{x'} w_m = D_{x'} u (1 - \rho_m).$$

We then have,

$$\int_{\mathbb{R}^N_+} |Dw_m - Du|^p \, \mathrm{d}x \leq C \int_{\mathbb{R}^N_+} |\rho_m|^p |Du|^p \, \mathrm{d}x + Cm^p \int_0^{2/m} \int_{\mathbb{R}^{N-1}} |u|^p \, \mathrm{d}x' \, \mathrm{d}t \, .$$

The first term on the right here goes to zero as $m \to \infty$ since the support of $\rho_m$ lies in $[0, 2/m]$. The second term has by (2.2),

$$\begin{aligned} Cm^p \int_0^{2/m} \int_{\mathbb{R}^{N-1}} |u(x', t)|^p \, \mathrm{d}x' \, \mathrm{d}t &\leq Cm^p \left( \int_0^{2/m} t^{p-1} \, \mathrm{d}t \right) \left( \int_0^{2/m} \int_{\mathbb{R}^{N-1}} |Du|^p \, \mathrm{d}x' \, \mathrm{d}x_N \right) \\ &\leq C \int_0^{2/m} \int_{\mathbb{R}^{N-1}} |Du|^p \, \mathrm{d}x' \, \mathrm{d}x_N \, . \end{aligned}$$

This also goes to zero as $m \to \infty$. Hence $w_m \to u$ in $W^{1,p}(\mathbb{R}^N_+)$. By construction $w_m$ is zero on $[0, 1/m]$. We may thus mollify the $w_m$ into $u_m \in C_c^\infty(\mathbb{R}^N_+)$ converging to $u$ in $W^{1,p}(\mathbb{R}^N_+)$ which gives $u \in W_0^{1,p}(\mathbb{R}^N_+)$. □

We end this section with a few more properties of traces that we provide for reference without proof; they can be found in [**Br**, §9.8.7].

**Trace Operator Properties**

(i) If $u \in W^{1,p}(\Omega)$, then $Tu \in W^{1-(1/p),p}(\partial\Omega)$ and surjects onto this space (we won't define this fractional Sobolev space).

(ii) The trace satisfies Green's formula for $u, v \in H^1(\Omega)$:

$$\int_\Omega \frac{\partial u}{\partial x_i} v = -\int_\Omega u \frac{\partial v}{\partial x_i} + \int_{\partial\Omega} Tu \cdot Tv \, (\vec{n} \cdot \vec{e_i}) \, \mathrm{d}\sigma.$$

(iii) For $u \in W^{2,p}(\Omega)$ it make sense to talk about,

$$\frac{\partial u}{\partial n} = T(Du) \cdot \vec{n} \in L^p(\partial\Omega),$$

as $T(Du) \in L^p(\partial\Omega)^N$. Green's other formula then holds for $u, v \in H^1(\Omega)$:

$$-\int_\Omega (\Delta u) v = \int_\Omega \nabla u \cdot \nabla v - \int_{\partial\Omega} \frac{\partial u}{\partial n} v \, \mathrm{d}\sigma.$$

**2.4. Sobolev Embeddings.** We now come to perhaps the most important tool in the theory of Sobolev spaces. There is a rich variety of embeddings (sometimes compact) of one Sobolev space into another, or into a Hölder space. These embeddings will allow us to bootstrap the regularity of weak solutions of PDEs and infer the existence of smooth/strong solutions. The key tool in proving these embeddings is a series of Sobolev inequalities that tell us one

norm on a space of functions can be controlled and bounded by another. The strength of the embedding we get will depend on the power $p$, the number of derivatives $k$, and the dimension $n$. There will be critical regimes governed by the sign of the quantity $k - \frac{n}{p}$.

We quickly mention the central result in the case $n = 1$ where things are particularly simple.

**Theorem 2.14: Sobolev Embedding in One Dimension [Evans, Ex. 5.4]**

If $I \subset \mathbb{R}$ is an open interval, then $u \in W^{1,p}(I)$ if and only if $u$ has a representative which is absolutely continuous with derivative in $L^p$.

REMARK 2.15. Recall by the Lebesgue FTC, a function $f$ is absolutely continuous if and only if it has a derivative $f'$ defined almost everywhere satisfying the FTC. So our model for $W^{1,p}(I)$ should be functions that are continuous and piecewise (in a loose sense) differentiable with derivative in $L^p$. △

In dimensions greater than one, $W^{1,p}$ functions need not be continuous. Consider the function $u(x) = |x|^{-\alpha}$ defined on the unit ball in $\mathbb{R}^n$. One computes this has weak derivatives,

$$\partial_j u = -\alpha x_j |x|^{-\alpha-2}.$$

More generally for a multi-index $\nu$ we inductively find,

$$|\partial^\nu u(x)| \leq c_\nu |x|^{-\alpha-|\nu|}.$$

We know that $|x|^{-\beta}$ is integrable on the unit ball if and only if $\beta < n$. Thus we see that all derivatives up to order $k$ of $u$ have finite $L^p(B^n)$ norm when,

$$\alpha p + kp < n.$$

Thus if $kp < n$, then $|x|^{-\alpha}$ will be in $W^{k,p}(B^n)$ for $\alpha < n/p - k$ but will never be continuous at zero. If $kp > n$, then in fact all $W^{k,p}$ functions will be continuous.

**Definition 2.16**

The *Sobolev conjugate* of $p \in [1, n)$ is given by

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n} \quad \text{or} \quad p^* = \frac{np}{n-p}.$$

**Important Intuition**

Suppose we have for all $u \in C_c^\infty(\mathbb{R}^n)$ that $\|u\|_q \leq C\|\nabla u\|_p$. By stretching to obtain $u_\lambda(x) = u(\lambda x)$, we must have that,

$$\|u\|_q \leq C\lambda^{1+\frac{N}{q}-\frac{N}{p}}\|\nabla u\|_p.$$

Since this holds for all $\lambda > 0$, we need $q = p^*$ to preserve the original inequality. This means the next theorem could only be true with the Sobolev conjugate as we've defined it.

**Theorem 2.17: Gagliardo–Nirenberg –Sobolev Inequality** [Br, Thm. 9.9]

If $p \in [1, n)$, then there is a constant $C(p, n)$ so that for all $u \in W^{1,p}(\mathbb{R}^n)$,

$$\|u\|_{L^{p^*}} \leq C\|Du\|_{L^p} \quad \text{and hence} \quad W^{1,p}(\mathbb{R}^n) \subset L^{p^*}(\mathbb{R}^n).$$

We first need a lemma.

LEMMA 2.18. *Let $n \geq 2$ and $f_1, \ldots, f_n \in L^{n-1}(\mathbb{R}^{n-1})$. If for $x \in \mathbb{R}^n$, we define $\widetilde{x}_i \in \mathbb{R}^{n-1}$ by omitting the $i$th entry of $x$ and if we define*

$$f(x) = f_1(\widetilde{x}_1)\cdots f_n(\widetilde{x}_n),$$

*then $f \in L^1(\mathbb{R}^n)$ with $\|f\|_{L^1(\mathbb{R}^n)} \leq \prod_{i=1}^n \|f_i\|_{L^{n-1}(\mathbb{R}^{n-1})}$.*

PROOF. Induct on $n$, applying the Hölder inequality and Fubini's theorem. See [Br, Lemma 9.4]. □

PROOF. (of Gagliardo–Nirenberg–Sobolev) Let $p = 1$ and $u \in C_c^1(\mathbb{R}^n)$. We have for each $i$,

$$|u(x_1, \ldots, x_n)| \leq \int_{\mathbb{R}} \left|\frac{\partial u}{\partial x_i}\right| \mathrm{d}t\,.$$

Call the right side $f_i(\widetilde{x}_i)$. We clearly have,

$$|u(x)|^n \leq \prod_{i=1}^n f_i(\widetilde{x}_i).$$

Our lemma then implies,

$$\int_{\mathbb{R}^n} |u(x)|^{n/(n-1)} \leq \prod_{i=1}^n \|f_i\|_{L^{n-1}(\mathbb{R}^n)}^{1/(n-1)} = \prod_{i=1}^n \left\|\frac{\partial u}{\partial x_i}\right\|_{L^1(\mathbb{R}^n)}^{1/(n-1)}.$$

And hence (using AM-GM),

$$\|u\|_{L^{n/(n-1)}(\mathbb{R}^n)} \leq \prod_{i=1}^{n} \left\| \frac{\partial u}{\partial x_i} \right\|_{L^1(\mathbb{R}^n)}^{1/n} \leq \|Du\|_{L^1(\mathbb{R}^n)}. \tag{2.3}$$

For the case $1 < p < n$, we can apply (2.3) to $|u|^{m-1}u$ for $m \geq 1$. We have,

$$\begin{aligned} \|u\|_{L^{mn/(n-1)}}^m &\leq m \prod_{i=1}^{n} \left\| |u|^{m-1} \frac{\partial u}{\partial x_i} \right\|_{L^1}^{1/n} \\ &\leq m \|u\|_{L^{p'(m-1)}}^{m-1} \prod_{i=1}^{n} \left\| \frac{\partial u}{\partial x_i} \right\|_{L^p}^{1/n}, \end{aligned} \tag{2.4}$$

where $p'$ is Hölder conjugate to $p$. Setting $m$ so that $p^* = mn/(n-1) = p'(m-1)$,

$$\|u\|_{L^{p^*}(\mathbb{R}^n)} \leq m \prod_{i=1}^{n} \left\| \frac{\partial u}{\partial x_i} \right\|_{L^p(\mathbb{R}^n)}^{1/n} \leq C \|Du\|_{L^p(\mathbb{R}^n)}.$$

To go from $u \in C_c^1(\mathbb{R}^n)$ to $u \in W^{1,p}(\mathbb{R}^n)$, we can use density to find a sequence in $C_c^1(\mathbb{R}^n)$ converging to $u$ in $W^{1,p}(\mathbb{R}^n)$. We may take this convergence to hold pointwise a.e.. Fatou's lemma implies that $u \in L^{p^*}$ and the above inequality extends to any $u$. □

Remark 2.19. Using the $L^p$ interpolation theorem and Young's inequality we easily deduce the following.

Corollary 2.20. *If $p \in [1, n)$, there is a continuous injection $W^{1,p}(\mathbb{R}^n) \subset L^q(\mathbb{R}^n)$ for any $q \in [p, p^*]$.*

With a little more work, we have almost the same result for the limiting case $p = n$.

Corollary 2.21. *There is a continuous injection $W^{1,n}(\mathbb{R}^n) \subset L^q(\mathbb{R}^n)$ for any $q \in [n, \infty)$.*

Note this injection does not extend to the Sobolev conjugate $n^* = \infty$, as in the theorem. For example, $\log\log(1 + 1/|x|)$ is in $W^{1,n}(B^n)$, and with the extension operator can be extended to lie in $W^{1,n}(\mathbb{R}^n)$. On the other hand it does not belong to $L^\infty$. △

Using the extension operators we also obtain the following.

**Theorem 2.22:** [Evans, Thm. 5.6.2]

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with $C^1$ boundary. If $p \in [1, n)$ and $u \in W^{1,p}(\Omega)$, then $u \in L^{p^*}(\Omega)$ and there is a constant $C(p, n, \Omega)$ so that

$$\|u\|_{L^{p^*}(\Omega)} \leq C\|u\|_{W^{1,p}(\Omega)}.$$

REMARK 2.23. We can relax $C^1$ boundary to Lipschitz, but for very bad boundary (like very sharp cusps) this result will fail. See [Ad, Thm. 4.48] for some general counterexamples. Clearly we do get for general $\Omega$ an embedding $W_0^{1,p}(\Omega) \hookrightarrow L^{p^*}(\Omega)$.

Now we move to the other case $n < p \leq \infty$. We first recall the Hölder spaces.

**Definition 2.24: Hölder Spaces**

A function $u : \Omega \to \mathbb{R}$ is *$\gamma$-Hölder continuous* for $0 < \gamma \leq 1$ if its *$\gamma$-Hölder seminorm*,

$$[u]_{C^{0,\gamma}(\overline{\Omega})} := \sup_{x \neq y \in \Omega} \left\{ \frac{|u(x) - u(y)|}{|x - y|^\gamma} \right\} \quad \text{is finite.}$$

The *Hölder space* $C^{k,\gamma}(\overline{\Omega})$ consists of all functions $u \in C^k(\overline{\Omega})$ whose *$\gamma$-Hölder norm*,

$$\|u\|_{C^{k,\gamma}(\overline{\Omega})} := \sum_{|\alpha| \leq k} \|D^\alpha u\|_{C(\overline{\Omega})} + \sum_{|\alpha| = k} [D^\alpha u]_{C^{0,\gamma}(\overline{\Omega})} \quad \text{is finite.}$$

I.e. this is the space of functions which are $k$-times continuously differentiable with bounded derivatives and $\gamma$-Hölder continuous $k$th derivatives.

THEOREM 2.25. *The Hölder spaces $C^{k,\gamma}(\overline{\Omega})$ are Banach spaces.*

**Theorem 2.26: Morrey's Inequality** [Evans, Thm. 5.6.4]

For $p \in (n, \infty)$, there is a constant $C(n, p)$ so that for all $u \in W^{1,p}(\mathbb{R}^n)$, $u$ a continuous representative satisfying,

$$\|u\|_{C^{0,\gamma}(\mathbb{R}^n)} \leq C\|u\|_{W^{1,p}(\mathbb{R}^n)} \quad \text{where} \quad \gamma = 1 - \frac{n}{p}.$$

Hence there is a continuous injection $W^{1,p}(\mathbb{R}^n) \subset C^{0,\gamma}(\mathbb{R}^n)$.

PROOF. Take $u \in C^1(\mathbb{R}^n)$. We begin by proving, there is a constant $C(n)$ so that,

$$\fint_{B(x,r)} |u(y) - u(x)| \, \mathrm{d}y \leq C \int_{B(x,r)} \frac{|Du(y)|}{|y-x|^{n-1}} \, \mathrm{d}y \,, \tag{2.5}$$

for any ball $B(x,r) \subset \mathbb{R}^n$ (recall $\fint_U$ denotes $\frac{1}{\mu(U)} \int_U$). Fix $w \in S^{n-1} \subset \mathbb{R}^n$,

$$|u(x+sw) - u(x)| = \left| \int_0^s Du(x+tw) \cdot w \, \mathrm{d}t \right| \leq \int_0^s |Du(x+tw)| \, \mathrm{d}t \,.$$

Thus,

$$\begin{aligned}
\int_{S^{n-1}} |u(x+sw) - u(x)| \, \mathrm{d}S(w) &\leq \int_0^s \int_{S^{n-1}} |Du(x+tw)| \, \mathrm{d}S(w) \,. \\
&= \int_0^s \int_{\partial B(x,t)} \frac{|Du(y)|}{t^{n-1}} \, \mathrm{d}S(y) \, \mathrm{d}t \\
&= \int_{B(x,s)} \frac{|Du(y)|}{|x-y|^{n-1}} \, \mathrm{d}y \\
&\leq \int_{B(x,r)} \frac{|Du(y)|}{|x-y|^{n-1}} \, \mathrm{d}y
\end{aligned}$$

We also compute,

$$\int_{S^{n-1}} |u(x+sw) - u(x)| \, \mathrm{d}S(w) = \frac{1}{s^{n-1}} \int_{\partial B(x,s)} |u(z) - u(z)| \, \mathrm{d}S(z) \,.$$

Combining with above and integrating $s$ from 0 to $r$ gives,

$$\int_{B(x,r)} |u(y) - u(x)| \, \mathrm{d}y \leq \frac{r^n}{n} \int_{B(x,r)} \frac{|Du(y)|}{|x-y|^{n-1}} \, \mathrm{d}y \,,$$

which proves (2.5).

To bound the $C^0$ norm, we estimate for fixed $x \in \mathbb{R}^n$,

$$|u(x)| \leq \fint_{B(x,1)} |u(x) - u(y)| \, \mathrm{d}y + \fint_{B(x,1)} |u(y)| \, \mathrm{d}y \,.$$

By (2.5) and Hölder,

$$\begin{aligned}
&\leq C \int_{B(x,1)} \frac{|Du(y)|}{|x-y|^{n-1}} \, \mathrm{d}y + C \|u\|_{L^p(B(x,1))} \\
&\leq C \|Du\|_{L^p(\mathbb{R}^n)} \cdot \left\| \frac{1}{|x-y|^{(n-1)}} \right\|_{L^{p/(p-1)}(B(x,1))} + C \|u\|_{L^p(\mathbb{R}^n)}.
\end{aligned}$$

The $L^{p/(p-1)}(B(x,1))$ norm of $|x-y|^{1-n}$ is finite,

$$\leq C \|u\|_{W^{1,p}(\mathbb{R}^n)}.$$

Thus we infer,

$$\|u\|_{C^0(\mathbb{R}^n)} \leq C\|u\|_{W^{1,p}(\mathbb{R}^n)}.$$

To bound the $C^{0,\gamma}$-seminorm, we consider any pair $x, y \in \mathbb{R}^n$. Let $r = |x-y|$ and let $W$ denote the lune $B(x,r) \cap B(y,r)$. One has,

$$|u(x) - u(y)| \leq \fint_W |u(x) - u(z)|\,\mathrm{d}z + \fint_W |u(y) - u(z)|\,\mathrm{d}z\,.$$

Focusing on the first term and applying (2.5),

$$\begin{aligned}
\fint_W |u(x) - u(z)|\,\mathrm{d}z &\leq C \fint_{B(x,r)} |u(x) - u(z)|\,\mathrm{d}z \\
&\leq C\|Du\|_{L^p(B(x,r))} \cdot \left\| \frac{1}{|x-z|^{(n-1)}} \right\|_{L^{p/(p-1)}(B(x,r))} \\
&\leq C\left( r^{n-(n-1)\frac{p}{p-1}} \right)^{\frac{p-1}{p}} \|Du\|_{L^p(\mathbb{R}^n)} \\
&= Cr^{1-\frac{n}{p}} \|Du\|_{L^p(\mathbb{R}^n)}.
\end{aligned}$$

We may do the same for the second term and combine to obtain,

$$|u(x) - u(y)| \leq C|x-y|^{1-\frac{n}{p}} \|Du\|_{L^p(\mathbb{R}^n)}.$$

By definition, we obtain,

$$[u]_{C^{0,1-n/p}(\mathbb{R}^n)} \leq C\|Du\|_{L^p(\mathbb{R}^n)}.$$

This concludes the proof in the case $u \in C^1(\mathbb{R}^n)$, to extend to any $u \in W^{1,p}(\mathbb{R}^n)$, we use the density of $C_c^\infty(\mathbb{R}^n)$. If $u_n \to u$ is such a sequence, then $u_n \to u^*$ in $C^{0,1-n/p}(\mathbb{R}^n)$ for some continuous $u^*$. Our inequality implies that $u = u^*$ in $W^{1,p}(\mathbb{R}^n)$, i.e. they agree almost everywhere. We can also easily extend to $p = \infty$. □

Remark 2.27. (i) This result extends to bounded $C^1$ domains $\Omega$ using extension operators and the same smooth approximation result.

(ii) In light of this result, for $p > n$ we always choose a continuous representative for $u \in W^{1,p}(\Omega)$. △

**Theorem 2.28: General Sobolev Inequalities [Evans, Thm. 5.6.6]**

Let $\Omega$ be $\mathbb{R}^n$ or a bounded domain with $C^1$ boundary inside $\mathbb{R}^n$. Suppose that $u \in W^{k,p}(\Omega)$.

(i) If $k < \frac{n}{p}$, then $u \in L^q(\Omega)$, where $1/q = 1/p - k/n$ and,

$$\|u\|_{L^q(\Omega)} \leq C\|u\|_{W^{k,p}(\Omega)},$$

for $C = C(k, p, n, \Omega)$ a constant.

(ii) If $k = \frac{n}{p}$ then $u \in L^q(\Omega)$ for all $q \in [p, \infty)$.

(iii) If $k > \frac{n}{p}$, then $u \in C^{k-\lfloor \frac{n}{p} \rfloor - 1, \gamma}(\overline{\Omega})$, where

$$\gamma = \begin{cases} \lfloor \frac{n}{p} \rfloor + 1 - \frac{n}{p} & \frac{n}{p} \notin \mathbb{Z} \\ \text{any } \lambda \in (0,1) & \frac{n}{p} \in \mathbb{Z}, \end{cases}$$

and there is a constant $C = C(k, p, n, \gamma, \Omega)$ so that,

$$\|u\|_{C^{k-\lfloor \frac{n}{p} \rfloor - 1, \gamma}(\overline{\Omega})} \leq C \|u\|_{W^{k,p}(\Omega)}.$$

REMARK 2.29. I'm going to assume (until told otherwise) that I shouldn't worry so much about the details of this statement, particularly case (iii). Broadly, we know if $k > n/p$ we get $\gamma$-Hölder continuity of some number of derivatives, governed by this fundamental quantity $k - n/p$. The $\gamma$ probably usually doesn't matter. △

PROOF. We just deal with $\mathbb{R}^n$ and the rest follow from the same theorems in the bounded domain case (which in turn followed from embeddings).

(i) We know $D^\alpha u \in L^p(\mathbb{R}^n)$ for all $|\alpha| < k$. Hence by Gagliardo–Nirenberg–Sobolev,

$$\|D^\beta u\|_{L^{p^*}(\mathbb{R}^n)} \leq C \|u\|_{W^{k,p}(\mathbb{R}^n)} \quad \text{if} \quad |\beta| \leq k - 1.$$

Thus $u \in W^{k-1,p^*}(\mathbb{R}^n)$. Repeating again we find $u \in W^{k-2,p^{**}}(\mathbb{R}^n)$. We note $\frac{1}{p^{**}} = \frac{1}{p^*} - \frac{1}{n} = \frac{1}{p} - \frac{2}{n}$. Continuing recursively, we find, that $u \in W^{0,q}(\mathbb{R}^n) = L^q(\mathbb{R}^n)$ for $\frac{1}{q} = \frac{1}{p} - \frac{k}{n}$. We get our inequality by combining the estimates from each stage.

(ii) We combine the above with the proof/conclusion of Corollary 2.21.

(iii) Suppose $n/p \notin \mathbb{Z}$. By the above argument we have $u \in W^{k-\ell,r}(\Omega)$ where $\frac{1}{r} = \frac{1}{p} - \frac{\ell}{n}$ as long as $\ell p < n$. choose $\ell = \lfloor n/p \rfloor$ so that $\ell < n/p < \ell + 1$. We thus have,

$$r = \frac{pn}{n - p\ell} > n.$$

By Morrey's inequality, $D^\alpha u \in C^{0,1-\frac{n}{r}}(\overline{\Omega})$ for all $|\alpha| \leq k - \ell - 1$. We have,

$$1 - \frac{n}{r} = 1 - \frac{n}{p} + \ell = \left\lfloor \frac{n}{p} \right\rfloor + 1 - \frac{n}{p}.$$

Hence, $u \in C^{k-\lfloor \frac{n}{p} \rfloor - 1, \lfloor \frac{n}{p} \rfloor + 1 - \frac{n}{p}}(\overline{\Omega})$, as claimed.

Now suppose $n/p \in \mathbb{Z}$. Taking $\ell = \frac{n}{p} - 1$, we obtain as before that $u \in W^{k-\ell,r}(\Omega)$ for $r = pn/(n-p\ell) = n$. So by Gagliardo–Nirenberg–Sobolev, $D^\alpha u \in L^q(\Omega)$ as long as $n \leq q < \infty$ and $|\alpha| \leq k - \ell - 1$. By Morrey, $D^\alpha u \in C^{0,1-\frac{n}{q}}(\overline{\Omega})$ for $n < q < \infty$ and $|\alpha| \leq k - \ell - 2$. So, $u \in C^{k-\frac{n}{p}-1,\gamma}(\overline{\Omega})$ for $\gamma \in (0,1)$. □

By a simple extension of this result, if $k > \ell$, $p < n$, and one has,

$$k - \frac{n}{p}\ell - \frac{n}{q},$$

then one obtains an embedding, $W^{k,p}(\Omega) \subset W^{\ell,q}(\mathbb{R}^n)$.

In figure 1 a useful schematic picture of the Sobolev embeddings is shown. We graph the plane $(x,y) = (1/p, k)$. A Sobolev space $W^{k,p}(\mathbb{R}^n)$ embeds into all spaces downwards along line of slope $n$ out of it. If this line crosses the $k$ axis at $y = r + \alpha$ (so that $r + \alpha = k - \frac{n}{p}$) for $\alpha \in (0,1)$ then we obtain an embedding into $C^{r,\alpha}(\mathbb{R}^n)$. The unfilled dots indicate optimal embeddings that *do not hold*, i.e. for $k = n/p$ we get an embedding into $L^q$ for any $q < \infty$ but not $q = \infty$ and for $k - \frac{n}{p} = m \in \mathbb{N}$ we get an embedding into $C^{m-1,\gamma}$ for all $\gamma < 1$ but not $C^m$. If we work on a $C^1$ bounded domain $\Omega$ instead of $\mathbb{R}^n$, we have an embedding into not just the points further down the line but all the points to the right of the line as well; we will shortly show the embedding into any points right of the line is compact.

**2.5. Compact Embeddings.** We now upgrade some of our Sobolev embeddings to compact embeddings. The big upshot of compactness comes straight from the definition; if $W^{1,p}(\Omega)$ compactly embeds in a Banach space $E$, then any bounded sequence in $W^{1,p}$ has a convergent subsequence in $E$.

Note that we need slighty more stringent conditions, in terms of the domain and the exponents, to get compactness than in our previous embeddings.

**Theorem 2.30: Rellich-Kondrachov Compactness** [Br, Thm. 9.16]

Suppose $\Omega$ is a bounded domain in $\mathbb{R}^n$ with $C^1$ boundary. We have compact embeddings,

$$\begin{cases} W^{1,p}(\Omega) \subset L^q(\Omega) & \forall q \in [1, p^*) & \text{if} \quad p < n, \\ W^{1,p}(\Omega) \subset L^q(\Omega) & \forall q \in [p, \infty) & \text{if} \quad p = n, \\ W^{1,p}(\Omega) \subset C(\overline{\Omega}) & & \text{if} \quad p > n. \end{cases}$$

In particular, there is always a compact injection $W^{1,p}(\Omega) \subset L^p(\Omega)$.

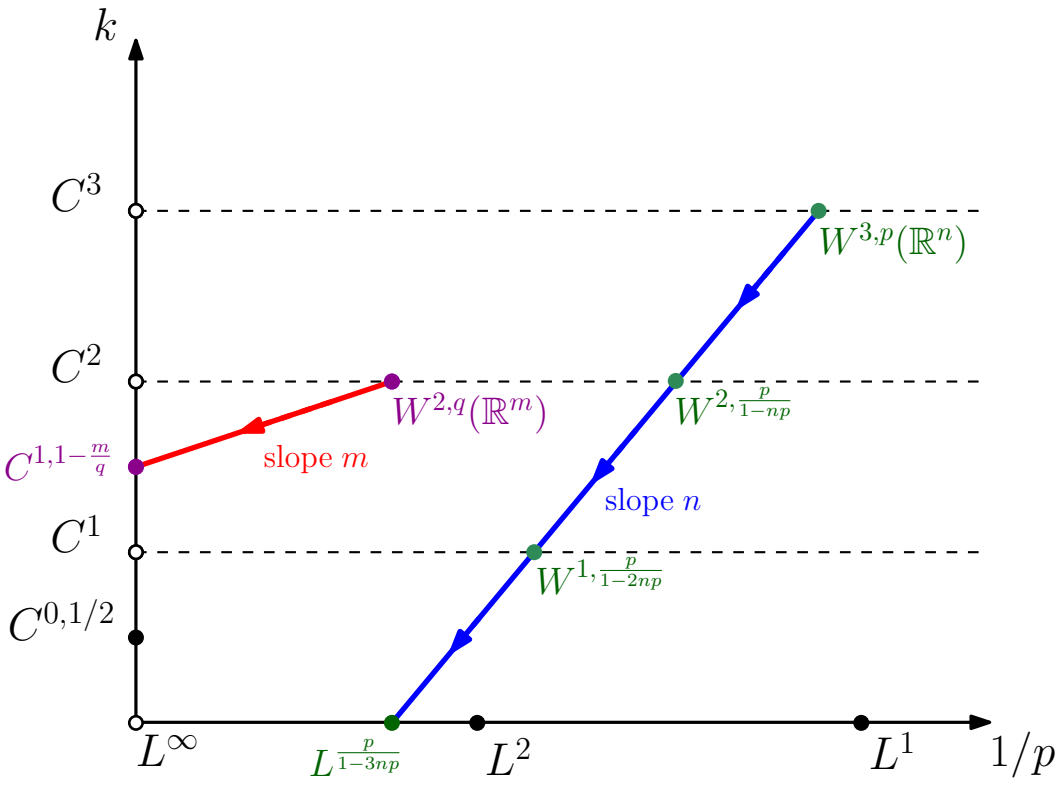



FIGURE 1. A useful diagram for encapsulating and remembering the Sobolev embeddings.

PROOF. Note that a family of functions with bounded norm in a Hölder space on a bounded domain is necessarily uniformly bounded and equicontinuous. Thus by Arzelà–Ascoli, such a family is precompact in $C(\overline{\Omega})$. So the third inclusion in our theorem follows directly from Morrey's inequality.

The second inclusion follows from the same argument as the first, so we just want to prove $W^{1,p}(\Omega) \subset L^q(\Omega)$ is compact for $q \in [1, p^*)$ and $p < n$. To do this, we invoke our $L^p$ version of Arzelà–Ascoli, the Fréchet–Kolmogorov theorem. If $\mathcal{F}$ denotes the image of the unit ball under the inclusion into $L^q(\Omega)$, clearly $\mathcal{F}$ is bounded and so we just need to show $L^q$-equicontinuity, i.e.,

$$\lim_{|h|\to 0} \|f(\cdot + h) - f(\cdot)\|_{L^q(\Omega)} = 0,$$

uniformly for $f \in \mathcal{F}$.

Let $u \in C_c^\infty(\mathbb{R}^n)$. We have, for $h \in \mathbb{R}^n$,

$$\begin{aligned}\int_{\mathbb{R}^n} |u(x+h) - u(x)|^p \,\mathrm{d}x &= \int_{\mathbb{R}^n} \left| \int_0^1 h \cdot \nabla u(x+th) \,\mathrm{d}t \right|^q \mathrm{d}x \\ &\leq |h|^p \int_0^1 \int_{\mathbb{R}^n} |\nabla u(x+th)|^p \,\mathrm{d}x \,\mathrm{d}t\,.\end{aligned}$$

Translating,

$$= |h|^p \int_0^1 \int_{\mathbb{R}^n} |\nabla u(x)|^p \,\mathrm{d}x \,\mathrm{d}t\,.$$

Hence we have $\|u(\cdot + h) - u(\cdot)\|_{L^p(\mathbb{R}^n)} \leq |h| \|\nabla u\|_{L^p(\mathbb{R})}$. This extends from $C_c^\infty(\mathbb{R}^n)$ to $W^{1,p}(\mathbb{R}^n)$ by density.

Now write $\frac{1}{q} = \frac{\alpha}{p} + \frac{1-\alpha}{p^*}$ for some $\alpha \in (0, 1]$. By the $L^p$ interpolation inequality,

$$\|f(\cdot + h) - f(\cdot)\|_{L^q(\mathbb{R}^n)} \leq \|f(\cdot + h) - f(\cdot)\|_{L^p(\mathbb{R}^n)}^\alpha \|f(\cdot + h) - f(\cdot)\|_{L^{p^*}(\mathbb{R}^n)}^{1-\alpha}.$$

By the inequality we derived above,

$$\begin{aligned}&\leq |h|^\alpha \|\nabla f\|_{L^p(\mathbb{R}^n)}^\alpha \|f(\cdot + h) - f(\cdot)\|_{L^{p^*}(\mathbb{R}^n)}^{1-\alpha} \\ &\leq |h|^\alpha \|\nabla f\|_{L^p(\mathbb{R}^n)}^\alpha \big(2\|f\|_{L^{p^*}(\mathbb{R}^n)}\big)^{1-\alpha}.\end{aligned}$$

We know $\|\nabla f\|_{L^p} \leq \|f\|_{W^{1,p}} \leq 1$ and $\|f\|_{L^{p^*}}$ is bounded by Gagliardo-Nirenberg-Sobolev, hence,

$$\leq C|h|^\alpha.$$

This constant $C$ is independent of $f$ and so we obtain $L^q$-equicontinutity. □

REMARK 2.31. Here are some ways we might hope to extend this result that fail.

(i) For $\Omega$ unbounded, the embedding $W^{1,p}(\Omega) \subset L^p(\Omega)$ need not be compact.

(ii) Even if $\Omega$ is smooth and bounded, the embedding $W^{1,p}(\Omega) \subset L^{p^*}(\Omega)$ will not be compact.

For the first statement, let $f$ be some smooth function of compact support in $W^{1,p}(\mathbb{R}^n)$. Suppose $\{x_n\}$ is a sequence in $\mathbb{R}^n$ going off to infinity. Then let $f_n(x) = f(x - x_n)$. One clearly has that $f_n$ is bounded in $W^{1,p}(\mathbb{R}^n)$. Suppose after passing to a subsequence, the $f_n$ converged in $L^p(\mathbb{R}^n)$ to some function $g$. By a standard result, after passing to a further subsequence the $f_n$ must converge pointwise a.e.. to $g$. But pointwise the $f_n$ go to zero. Hence $g = 0$, which is a contradiction since $\|f_n - g\|_{L^p} = \|f_n\|_{L^p} \nrightarrow 0$.

For the second statement, consider for $p < n$ the functions $f_k : B^n \to \mathbb{R}$,

$$f_k(x) = \begin{cases} k^{n/p^*}(1 - k|x|) & |x| < 1/k \\ 0 & |x| \geq 1/k. \end{cases}$$

We have $\partial_i f_k(x) = -k^{1+n/p^*} \frac{x_i}{|x|}$ within the ball of radius $1/k$. Then,

$$\begin{aligned} \|f_k\|_{L^p} &= k^{n/p^*} \left( \int_{B^n(1/k)} (1 - k|x|)^p \, \mathrm{d}x \right)^{1/p} \\ &= k^{n/p^*} \left( \mathrm{vol}(S^{n-1}) \int_0^{1/k} r^{n-1}(1 - kr)^p \, \mathrm{d}r \right)^{1/p} \\ &\leq C k^{n/p^*} \left(k^{-n}\right)^{1/p} \\ &= C k^{n/p^* - n/p} \\ &= \frac{C}{k}. \end{aligned}$$

Further,

$$\begin{aligned} \|\partial_i f_k\|_{L^p} &\leq k^{1+n/p^*} (\mathrm{vol}(B^n(1/k)))^{1/p} \\ &= C k^{1+n/p^*} k^{-n/p} \\ &= C. \end{aligned}$$

Thus the $W^{1,p}$ norms of the sequence $f_k$ are uniformly bounded. On the other hand,

$$\begin{aligned} \|f_k\|_{L^{p^*}} &= k^{n/p^*} \left( \int_{B^n(1/k)} (1 - k|x|)^p \, \mathrm{d}x \right)^{1/p^*} \\ &= k^{n/p^*} \left( \mathrm{vol}(S^{n-1}) \int_0^{1/k} r^{n-1}(1 - kr)^p \, \mathrm{d}r \right)^{1/p^*} \\ &\geq C k^{n/p^*} \left(k^{-n}\right)^{1/p^*} \\ &= C k^{n/p^* - n/p^*} \\ &= C > 0. \end{aligned}$$

Note since $f_k \to 0$ pointwise, if a subsequence of the $f_k$ converged in $L^{p^*}$, they would have to converge to zero and hence their $L^{p^*}$ norms would have to converge to zero. Thus the sequence $f_k$ cannot possibly converge (even on a subsequence) to a finite limit in $L^{p^*}$. If we instead consider the embedding into $L^q$ for $q$ slightly smaller than $p^*$, then we find $\|f_k\|_{L^q} \sim C k^{-\varepsilon}$ and hence we do get compactness when $q < p^*$, as the Rellich–Kondrachov theorem just showed. △

We conclude with one final result of crucial importance in elliptic PDE theory.

**Theorem 2.32: Poincaré's Inequality** [Evans, Thm. 5.6.3]

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain and $u \in W_0^{1,p}(\Omega)$ for $p \in [1, n)$. Then there is a constant $C = C(p, n, \Omega)$ so that,

$$\|u\|_{L^p(\Omega)} \leq C\|\nabla u\|_{L^p(\Omega)}.$$

PROOF. We can approximate $u$ in $W^{1,p}(\Omega)$ by a sequence $u_m \in C_c^\infty(\Omega)$. We extend the $u_m$ by zero to lie in $C_c^\infty(\mathbb{R}^n)$. From Gagliardo–Nirenberg–Sobolev,

$$\|u_m\|_{L^{p^*}(\Omega)} = \|u_m\|_{L^{p^*}(\mathbb{R}^n)} \leq C\|\nabla u_m\|_{L^p(\mathbb{R}^n)} = C\|\nabla u_m\|_{L^p(\Omega)}.$$

By boundedness, we may replace $p^*$ with any $q \in [1, p^*]$. The inequality follows for $u$ by approximation. □

REMARK 2.33. Note this is stronger than the inequality we obtained for $u \in W^{1,p}(\Omega)$ under the same conditions since it is bounded by $\|\nabla u\|_{L^p(\Omega)}$ rather than $\|u\|_{W^{1,p}(\Omega)}$. The key is that $W_0^{1,p}(\Omega)$ can be approximated by $C_c^\infty(\Omega)$ rather than just $C_c^\infty(\mathbb{R}^n)$. △

**2.6. The Dual Space of $H_0^1$.** We write $W^{-1,p'}(\Omega)$ for the dual of $W_0^1(\Omega)$ with $1 \leq p < \infty$ and $H^{-1}(\Omega)$ for the dual of $H_0^1(\Omega)$. Since we are identifying $L^2(\Omega)$ with its dual, one has continuous dense embeddings

$$H_0^1(\Omega) \subset L^2(\Omega) \subset H^{-1}(\Omega).$$

We equip $H^{-1}(\Omega)$ with the standard dual space operator norm $\|\cdot\|_{H^{-1}(\Omega)}$.

**Theorem 2.34: Characterization of $H_0^1$** [Evans, Thm. 5.9.1]

Let $f \in H^{-1}(\Omega)$. Then there are functions $f_0, \ldots, f_n \in L^2(\Omega)$ so that,

$$f(v) = \int_\Omega f_0 v + \sum_{i=1}^n \int_\Omega f_i \frac{\partial v}{\partial x_i} \quad \text{for all} \quad v \in H_0^1(\Omega). \tag{2.6}$$

Further,

$$\|f\|_{H^{-1}(\Omega)}^2 = \inf\left\{ \sum_{i=0}^n \|f_i\|_{L^2(\Omega)}^2 \,\middle|\, f_i \text{ satisfy (2.6) for } f \right\}.$$

PROOF. By Riesz representation, there is $u \in H_0^1(\Omega)$ so that $\langle u, v\rangle_{H_0^1(\Omega)} = f(v)$ for all $v \in H_0^1(\Omega)$. Setting $f_0 = u$ and $f_i = \frac{\partial u}{\partial x_i}$ for $1 \leq i \leq n$ gives (2.6) by definition of the norm on $H_0^1(\Omega)$.

To obtain the formula for $\|f\|_{H^{-1}(\Omega)}$, note if $u$ is again given by $\langle u, v\rangle_{H_0^1(\Omega)} = f(v)$ we have,

$$\|f\|^2_{H^{-1}(\Omega)} = \|u\|^2_{H^1(\Omega)} = \int_\Omega |u|^2 + |Du|^2 \, \mathrm{d}x = \sum_{i=0}^n \int_\Omega |f_i|^2 \, \mathrm{d}x \, .$$

Now suppose we pick any other set of function $g_i$ to satisfy (2.6) for $f$. Then, for $v \in H_0^1(\Omega)$,

$$\langle u, v\rangle_{H^1(\Omega)} = \int_\Omega g_0 v + \sum_{i=1}^n \int_\Omega g_i \frac{\partial v}{\partial x_i}.$$

Plugging in $u = v$,

$$\begin{aligned}\int_\Omega |u|^2 + |Du|^2 \, \mathrm{d}x &= \int_\Omega g_0 f_0 + \sum_{i=1}^n \int_\Omega g_i f_i \\ &\leq \sqrt{\sum_{i=0}^n \int_\Omega |g_i|^2 \, \mathrm{d}x} \cdot \sqrt{\sum_{i=0}^n \int_\Omega |f_i|^2 \, \mathrm{d}x}.\end{aligned}$$

And so by above,

$$\int_\Omega |u|^2 + |Du|^2 \, \mathrm{d}x \leq \sum_{i=0}^n \int_\Omega |g_i|^2 \, \mathrm{d}x \, .$$

□

REMARK 2.35. (i) Brezis [**Br**, pp. 291] claims instead,

$$\|f\|_{H^{-1}(\Omega)} = \max_{0 \leq i \leq n} \|f_i\|_{H^1(\Omega)}.$$

It's not clear at all to me why these should be the same.

(ii) A similar theorem is true for the spaces $W^{-1,p'}(\Omega)$. See [**Br**, Prop. 9.20]. △

REMARK 2.36. One should note that everything we have done throughout this chapter can be adapted to the setting of closed Riemannian manifolds and compact Riemannian manifolds with $C^1$ boundary. Simply cover the manifold with charts, use partitions of unity, and apply the change of variables theorem for weak derivatives. △

## 3. Second Order Elliptic PDEs

Now we are prepared to actually study some PDEs. We will see that general elliptic PDEs satisfying many of the celebrated properties of harmonic functions including existence, regularity, and maximum principles.

**3.1. Basic Definitions.** Consider a second order linear partial differential operator $L$ on $\Omega \subset \mathbb{R}^n$ open and bounded and a function $f : \Omega \to \mathbb{R}$. We will be interested in the inhomogeneous *Dirichlet problem* for $L$, i.e. the existence of $u : \overline{\Omega} \to \mathbb{R}$ satisfying,

$$\begin{cases} Lu = f & \text{on} \quad \Omega \\ u = 0 & \text{on} \quad \partial\Omega. \end{cases} \tag{3.1}$$

We may write $L$ in terms of coefficient functions $a^{ij}, b^i, c$ and write it in either its *divergence form*,

$$Lu = -\sum_{i,j} \frac{\partial}{\partial x_j}\left(a^{ij}(x)\frac{\partial u}{\partial x_i}\right) + \sum_i b^i(x)\frac{\partial u}{\partial x_i} + c(x)u,$$

or its *non-divergence form*,

$$Lu = -\sum_{i,j} a^{ij}\frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_i b^i(x)\frac{\partial u}{\partial x_i} + c(x)u.$$

As long as the $a^{ij}$ are $C^1$, these are clearly equivalent; they will be useful for studying different aspects of elliptic equations.

We make the assumption that the $a_{ij}$ are symmetric, which we can clearly do freely.

**Definition 3.1**

A differential operator $L$ is *uniformly elliptic* if there is $\theta > 0$ so that,

$$\sum_{i,j} a^{ij}(x)\xi_i\xi_j \geq \theta|\xi|^2,$$

for all $\xi \in \mathbb{R}^n$ and almost every $x \in \Omega$. Thus the symmetric matrix $(a^{ij}(x))$ is positive definite with minimal eigenvalue at least $\theta$.

Remark 3.2. The function, $\sigma_L : \Omega \times \mathbb{R}^n \to \mathbb{R}$ given by,

$$\sigma_L(x, \xi) = \sum_{i,j} a^{ij}(x)\xi_i\xi_j,$$

is called the *principal symbol* of $L$. We can relax the above condition to ask only that $\sigma_L > 0$ for $\xi \neq 0$, which is called *weak ellipticity*.

More generally we could work on a manifold $M$ so that our differential operator is $L : \Gamma(E) \to \Gamma(F)$ for vector bundles $E, F$ with the same local description as above for $a^{ij}, b^i, c$ now $\mathrm{Hom}(E, F)$-valued. The symbol becomes a linear map,

$$\sigma_L : S^k(T^*M) \otimes E \to F.$$

$L$ is elliptic if $\sigma_L(\xi_x, \ldots, \xi_x) \neq 0$ for any non-zero $\xi_x \in T^*M$. △

EXAMPLE 3.3. If we take $a^{ij} = \delta^{ij}$ and $b^i = c = 0$, we get the Laplacian $L = -\Delta$. The symbol of $L$ is $|\xi|^2$ and hence this is uniformly elliptic. ▲

REMARK 3.4. One physical interpretation of solutions $u$ of second order elliptic PDEs $Lu = 0$ is as the equilibrium states of densities of some substance, like a fluid, in $\Omega$.

The second order part $\mathbf{A}(x) = (a^{ij})$ represents the diffusion of $u$. The vector-valued function $\vec{\mathbf{F}} = -\mathbf{A}Du$ represents the diffusive flux density and so the elliptic condition $\vec{\mathbf{F}} \cdot Du \leq 0$ says that flow moves from regions of high to low concentration.

The first order term $\mathbf{b}(x) = (b^i)$ represents transport of the substance and the zeroth order term $c(x)$ represents the local creation and annihilation of the substance. △

Consider our elliptic PDE (3.1) in divergence form with $a^{ij}, b^i, c \in L^\infty(\Omega)$ and $f \in L^2(\Omega)$. In order to generalize from actual solutions to a notion of "weak solutions" lying in some Sobolev space, we could integrate our PDE against a test function and ask for the candidate weak solution to satisfy the corresponding integral equation. By approximation, instead of requiring our test functions be smooth we can take anything within $H^1_0(\Omega)$, where we require trace zero for integration by parts and to incorporate the boundary condition.

This motivates the construction of a bilinear form $B : H^1_0(\Omega) \times H^1_0(\Omega) \to \mathbb{R}$ associated to $L$ and defined by,

$$\boxed{B[u, v] = \int_\Omega \left( \sum_{i,j} a^{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} + \sum_i b^i \frac{\partial u}{\partial x_i} v + cuv \right) \mathrm{d}x \,.}$$

**Definition 3.5: Variational Formulation of (3.1)**

We say $u \in H^1_0(\Omega)$ is a *weak solution* of the Dirichlet problem (3.1) if,

$$B[u, v] = \langle f, v \rangle_{L^2(\Omega)} \quad \text{for all} \quad v \in H^1_0(\Omega).$$

Recalling the dual of $H_0^1(\Omega)$, we can slightly generalize to the PDE,

$$\begin{cases} Lu = f_0 - \sum_i \partial_{x_i} f_i & \text{on} \quad \Omega \\ u = 0 & \text{on} \quad \partial\Omega, \end{cases} \tag{3.2}$$

where $f_i \in L^2(\Omega)$ so that $f_0 - \sum_i \partial_{x_i} f_i$ represents some $\varphi \in H^{-1}(\Omega)$. We then say $u \in H_0^1(\Omega)$ is a weak solution of (3.2) if,

$$B[u,v] = \varphi(v) = \int_\Omega \left( f_0 v + \sum_{i=1}^n f_i \frac{\partial v}{\partial x_i} \right) \mathrm{d}x \quad \text{for all} \quad v \in H_0^1(\Omega).$$

Remark 3.6. We could also consider a more general Dirichlet problem. If $\partial\Omega$ is $C^1$, we could ask $u$ to weakly solve,

$$\begin{cases} Lu = f_0 - \sum_i \partial_{x_i} f_i & \text{on} \quad \Omega \\ u = g & \text{on} \quad \partial\Omega, \end{cases}$$

for $g \in L^2(\partial\Omega)$, where we ask $u = g$ on $\partial\Omega$ in the sense of traces. For a solution to exist, $g$ needs to be in the image of the trace operator, so say $g = Tw$ for $w \in H^1(\Omega)$. But then $\widetilde{u} = u - w$ is in $H_0^1(\Omega)$ and will solve 3.1 with source term $\widetilde{f} = f - Lw$ belonging to $H^{-1}(\Omega)$. Hence it suffices to deal with our case where $Tu = 0$. Of course, we may also be interested in the Neumann problem or other more complicated boundary conditions. △

**3.2. Existence and Uniqueness.** Since we have formulated weak solutions in terms of certain bilinear forms, we would like to apply Lax–Milgram to $B$.

Proposition 3.7 (Energy Estimates). *$B[u,v]$ is continuous and nearly coercive. That is, there are constants $\alpha, \beta > 0$, $\gamma \geq 0$ so that,*

$$|B[u,v]| \leq \alpha \|u\|_{H_0^1(\Omega)} \|v\|_{H_0^1(\Omega)}$$
$$\beta \|u\|^2_{H_0^1(\Omega)} \leq B[u,u] + \gamma \|u\|^2_{L^2(\Omega)},$$

*for all $u, v \in H_0^1(\Omega)$.*

Proof. We clearly have,

$$\begin{aligned} |B[u,v]| &\leq \sum_{i,j} \|a^{ij}\|_{L^\infty} \int_\Omega |Du| \cdot |Dv| \, \mathrm{d}x + \sum_i \|b^i\|_{L^\infty} \int_\Omega |Du| \cdot |v| \, \mathrm{d}x \\ &\qquad + \|c\|_{L^\infty} \int_\Omega |u| \cdot |v| \, \mathrm{d}x \\ &\leq \alpha \|u\|_{H_0^1(\Omega)} \|v\|_{H_0^1(\Omega)}. \end{aligned}$$

So $B$ is continuous. From ellipticity,

$$\begin{aligned}\theta \int_\Omega |Du|^2 \,\mathrm{d}x &\le \int_\Omega \sum_{i,j} a^{ij} \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_j} u \,\mathrm{d}x \\ &= B[u,u] - \int_\Omega \left( \sum_i b^i u \frac{\partial u}{\partial x_i} + cu^2 \right) \mathrm{d}x \\ &\le B[u,u] + \sum_i \|b^i\|_{L^\infty} \int_\Omega |Du| \cdot |u| \,\mathrm{d}x + \|c\|_{L^\infty} \int_\Omega u^2 \,\mathrm{d}x \,.\end{aligned}$$

Recall the inequality $ab \le \varepsilon a^2 + b^2/4\varepsilon$ can be derived from AM-GM. Taking $0 < 2\varepsilon < \theta / \sum_i \|b^i\|_{L^\infty}$,

$$\frac{\theta}{2} \int_\Omega |Du|^2 \,\mathrm{d}x \le B[u,u] + \|c\|_{L^\infty} \int_\Omega u^2 \,\mathrm{d}x \,.$$

From Poincarè's inequality we obtain,

$$\frac{\theta}{2(C+1)} \|u\|^2_{H^1_0(\Omega)} \le \frac{\theta}{2} \|Du\|^2_{L^2(\Omega)} \le B[u,u] + \gamma \|u\|^2_{L^2(\Omega)}.$$

□

**Theorem 3.8: First Existence for Weak Solutions** [**Evans**, Thm. 6.2.3]

There is $\gamma \ge 0$ so that for $\mu \ge \gamma$ and any $f \in L^2(\Omega)$, there is a unique weak solution $u \in H^1_0(\Omega)$ to,

$$\begin{cases} Lu + \mu u = f & \text{on } \Omega \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{3.3}$$

Proof. Let $B_\mu[u,v] = B[u,v] + \mu \langle u, v \rangle_{L^2(\Omega)}$. If $\mu \ge \gamma$ from Proposition 3.7 then $B_\mu$ is coercive. By Lax–Milgram, for any $f \in L^2(\Omega)$, we may find $u \in H^1_0(\Omega)$ so that $B_\mu[u,v] = \langle f, v \rangle_{L^2(\Omega)}$. Hence $u$ weakly solves (3.3). □

Remark 3.9. Note if all $b^i = 0$ and $c \ge 0$, for example if $L = -\Delta$, then coercivity of $B$ is automatic and we can take $\mu = 0$ above. So in this case the original problem (3.1) has a unique weak solution. If we additionally have $(a^{ij})$ symmetric, then the solution $u$ is characterized as the unique minimizer of $\frac{1}{2} B[v,v] - \langle f, v \rangle_{L^2(\Omega)}$. △

We can generalize from $f \in L^2(\Omega)$ to $f_0, \ldots, f_n \in H^{-1}(\Omega)$ in (3.3) and obtain the same result. We can nicely rephrase this as follows.

COROLLARY 3.10. *There is $\gamma \geq 0$ so that for $\mu \geq \gamma$, the map $L_\mu = L + \mu I : H_0^1(\Omega) \to H^{-1}(\Omega)$ is an isomorphism.*

**Definition 3.11**

If the $b^i \in C^1(\overline{\Omega})$, the *formal adjoint* of $L$ is,

$$L^* v = -\sum_{i,j} \frac{\partial}{\partial x_i}\left(a^{ij}\frac{\partial v}{\partial x_j}\right) - \sum_i b^i \frac{\partial v}{\partial x_i} + \left(c - \sum_i \frac{\partial b^i}{\partial x_i}\right) v.$$

The *adjoint* of $B$ is given by $B^*[v,u] = B[u,v]$. We say $u \in H_0^1(\Omega)$ is a *weak solution of the adjoint problem*,

$$\begin{cases} L^* u = f & \text{on } \Omega \\ u = 0 & \text{on } \partial\mathbb{Z}, \end{cases}$$

if $B^*[u,v] = \langle f, v\rangle_{L^2(\Omega)}$ for all $v \in H_0^1(\Omega)$.

**Theorem 3.12: Fredholm Alternative for Elliptic Operators**

(i) Exactly one of the following two statements holds:
  ($\alpha$) For each $f \in L^2(\Omega)$, there is a unique weak solution of the standard Dirichlet problem (3.1).
  ($\beta$) There is a non-trivial weak solution of the homogeneous problem (3.1) with $f = 0$.

(ii) If ($\beta$) holds, the space $N \subset H_0^1(\Omega)$ of weak solutions of the homogeneous problem is finite and equals the dimension of the weak solutions of the homogeneous problem for $L^*$.

(iii) ($\alpha$) holds and (3.1) has a weak solution if and only if $f \in \ker(L^*)^\perp$.

PROOF. By Theorem 3.8, $L_\gamma := L + \gamma I : H_0^1(\Omega) \to H^{-1}(\Omega)$ is an isomorphism. $u$ is a weak solution of (3.1) if and only if $u = L_\gamma^{-1}(\gamma u + f)$. Taking $K = \gamma L_\gamma^{-1}$, we see that $u$ solve (3.1) if and only if $u - Ku = h$ where $h = L_\gamma^{-1} f$.

By restricting the domain and applying Sobolev embedding, we have a linear operator $K : L^2(\Omega) \to L^2(\Omega)$. We claim this is bounded and compact. Take $u \in H_0^1(\Omega)$ with $g = L_\gamma u$. We have, from our energy estimates,

$$\beta \|u\|^2_{H_0^1(\Omega)} \leq B_\gamma[u,u] = \langle g, u\rangle_{L^2(\Omega)} \leq \|g\|_{L^2(\Omega)} \cdot \|u\|_{H_0^1(\Omega)}.$$

Hence,

$$\|Kg\|_{H_0^1(\Omega)} \leq C\|g\|_{L^2(\Omega)}.$$

Thus $K : L^2(\Omega) \to H_0^1(\Omega)$ is bounded and by Rellich-Kondrachov, after composing with the compact embedding $H_0^1(\Omega) \subset L^2(\Omega)$, $K$ is compact.

We can thus apply the Fredholm alternative to $I - K$. We conclude that $\dim(\ker(I-K)) = \dim(\ker(I-K^*))$, proving (ii). If the $\dim(\ker(I-K)) = 0$, then $(\alpha)$ holds, otherwise we obtain solutions of the homogeneous problem giving $(\beta)$; this proves (i). Recall also $h \in \operatorname{im}(I-K)$ iff $h \in \ker(I-K^*)^\perp$. If $v \in \ker(I-K^*)$ then,

$$\langle h, v\rangle_{L^2(\Omega)} = \frac{1}{\gamma}\langle Kf, v\rangle_{L^2(\Omega)} = \frac{1}{\gamma}\langle f, K^*v\rangle_{L^2(\Omega)} = \frac{1}{\gamma}\langle f, v\rangle_{L^2(\Omega)}.$$

Hence $h \in \ker(I-K^*)^\perp$ iff $f = L_\gamma h \in \ker(I-K^*)^\perp$. But note $v = K^*v$ iff $L_\gamma^* v = \gamma v$ iff $L^*v = 0$. Thus $u$ solves (3.1) if and only if $f \in \ker(L^*)^\perp$, proving (iii). □

Remark 3.13. I feel I have a little trouble remembering this result so here's my attempt at simplifying. Firstly, we can dispense with the "alternative" part of the statement. Simply put, the dimension of the space of solutions of the homogeneous problem $Lu = 0$ equals the dimension of the space of solutions of the adjoint homogeneous problem $L^*v = 0$ (and is finite); moreover, $Lu = f$ has a solution precisely when $f$ is orthogonal to the solutions $v$ of the homogeneous adjoint problem. In order to remember the proof, the key is just that $K = \gamma L_\gamma^{-1}$ is compact so we can apply the Fredholm alternative, and I should be able to reconstruct the rest of the argument from that.

**Theorem 3.14: Spectrum of Elliptic Operator** [Evans, Thm. 6.2.5]

There is a countable set $\Sigma \subset \mathbb{R}$ so that $\lambda \notin \Sigma$ when the boundary-value problem,

$$\begin{cases} Lu = \lambda u + f & \text{on } \Omega \\ u = 0 & \text{on } \partial\Omega, \end{cases}$$

has a unique weak solution for each $f \in L^2(\Omega)$. If $\Sigma$ is infinite, it is a sequence converging to infinity.

Proof. Without loss of generality we may increase $\gamma$ so it is positive. By the first existence for weak solutions we know $\lambda \leq -\gamma$ does not belong to $\Sigma$, so that $\lambda > -\gamma$.

By the Fredholm alternative, $\lambda \in \Sigma$ iff there is a non-zero weak solution of the homogeneous problem $Lu = \lambda u$. This is true iff there is a weak solution

of $Lu + \gamma u = (\gamma + \lambda)u$. This holds iff,

$$u = L_\gamma^{-1}(\gamma + \lambda)u = \frac{\gamma + \lambda}{\gamma} Ku.$$

Hence, $\lambda$ is in the spectrum of $L$ (i.e. in $\Sigma$) precisely when $\frac{\gamma}{\gamma+\lambda}$ is an eigenvalue of $K$.

By Proposition 1.13, the eigenvalues of $K$ are countable, and if infinite converge to zero. Hence the spectrum of $L$ is either finite or a sequence converging to infinity (since the $\lambda$ are bounded below). □

PROPOSITION 3.15. *If $\lambda \notin \Sigma$, there is a constant $C(\lambda, \Omega, L)$ so that $\|u\|_{L^2(\Omega)} \leq C\|f\|_{L^2(\Omega)}$ for $u$ a weak solution of $Lu = \lambda u + f$.*

PROOF. Suppose not. We can find, after rescaling, two sequence $u_k, f_k$ so that $\|u_k\|_{L^2(\Omega)} = 1$ and $\|f_k\|_{L^2(\Omega)} \to 0$. By Banach–Alaoglu and Rellich–Kondrachov, there is a subsequence so that $u_k$ converge to some $u$ weakly in $H_0^1(\Omega)$ and strongly in $L^2(\Omega)$. We conclude $u \neq 0$ is a weak solution of $Lu = \lambda u$. But this is not possible as $\lambda \notin \Sigma$. □

**3.3. Elliptic Regularity.** We now have some conditions that guarantee the existence of weak solutions of (3.1). We now prove results demonstrating that weak solutions are in fact smooth, and hence they are strong solutions. The model for this is the fact the Poisson representation formula and Cauchy integral formula imply solutions of Laplace's equation and the Cauchy–Riemann equations are smooth.

Note in the Poisson equation $\Delta u = f$, the $L^2$ norm of $D^2 u$ is equal to $L^2$ norm of $f$. Differentiating, the $L^2$ norm of $D^{2+n}u$ is controlled by the $L^2$ norm of $D^n f$. Generalizing, we may hope that if $f \in H^n$ for all $n$, then the same will hold for $u$ and hence $u$ is smooth. This analysis doesn't quite work if we don't already know we can differentiate $u$ and so we will have to work harder to prove results on the regularity of solutions of general elliptic PDEs avoiding circular arguments.

**Theorem 3.16: Interior $H^2$-regularity [Evans, Thm. 6.3.1]**

Assume the coefficients of $L$ satisfy $a^{ij} \in C^1(\Omega)$ and $b^i, c \in L^\infty(\Omega)$, and that $f \in L^2(\Omega)$. If $u \in H^1(\Omega)$ is a weak solution of $Lu = f$ then $u \in H^2_{\text{loc}}(\mathbb{Z})$. If $V \Subset \Omega$, then there is a constant $C(\Omega, V, L)$ so that,

$$\|u\|_{H^2(V)} \leq C(\|f\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)}).$$

Not we are not assuming anything on the boundary conditions. The conclusion of this theorem implies $u$ is a strong solution a.e.. Warning... long argument ahead!

PROOF. Fix $V \Subset W \Subset \Omega$. Let $\xi$ be a smooth bump function which is non-negative, obtains a maximum of one on $V$, and vanishes off of $W$.

Since $u$ is a weak solution of $Lu = f$, we have for all $v \in H_0^1(\Omega)$,

$$\sum_{i,j} \int_\Omega a^{ij} \partial_{x_i} u \, \partial_{x_j} v \, \mathrm{d}x = \int_{\mathbb{Z}} \widetilde{f} v \, \mathrm{d}x \quad \text{where} \quad \widetilde{f} = f - \sum_i b^i \partial_{x_i} u - cu.$$

Call the left side of this equality $A$ and the right side $B$; we will make estimates on both for $v$ of a specified form.

Take a small $|h| > 0$ and $k \in \{1, \dots, n\}$. Choose,

$$v = -D_k^{-h}(\xi^2 D_k^h u) \quad \text{where} \quad D_k^h u = \frac{u(x + he_k) - u(x)}{h},$$

and $e_k$ is the $k$th standard basis vector of $\mathbb{R}^n$. These difference quotients satisfy the integration by parts and product rule formulae,

$$\int_\Omega v D_k^{-h} w \, \mathrm{d}x = -\int_\Omega w D_k^h v \, \mathrm{d}x \quad \text{and} \quad D_k^h(vw) = v^h D_k^h w + w D_k^k v,$$

where $v^h(x) = v(x + he_k)$. Then,

$$\begin{aligned}
A &= -\sum_{i,j} \int_\Omega a^{ij} \frac{\partial u}{\partial x_i} \frac{\partial}{\partial x_j} \left[ D_k^{-h}(\xi^2 D_k^h u) \right] \mathrm{d}x \\
&= \sum_{i,j} \int_\Omega D_k^h \left( a^{ij} u_{x_i} \right) \frac{\partial}{\partial x_j} \left[ \xi^2 D_k^h u \right] \mathrm{d}x \\
&= \sum_{i,j} \int_\Omega a^{ij,h} D_k^h u_{x_i} \partial_{x_j} [\xi^2 D_k^h u] + (D_k^h a^{ij}) u_{x_i} \partial_{x_j} [\xi^2 D_k^h u] \, \mathrm{d}x \\
&= \sum_{i,j} \int_\Omega \Big[ a^{ij,h} D_k^h u_{x_i} \xi^2 D_k^h u_{x_j} + a^{ij,h} D_k^h u_{x_i} 2\xi \xi_{x_j} D_k^h u \\
&\qquad + (D_k^h a^{ij}) u_{x_i} \xi^2 D_k^h u_{x_j} + (D_k^h a^{ij}) u_{x_i} 2\xi \xi_{x_j} D_k^h u \Big] \, \mathrm{d}x \, .
\end{aligned}$$

Call the first term $A_1$ and the sum of the next three terms $A_2$, so that $A = A_1 + A_2$. We have by uniform ellipticity,

$$A_1 \geq \theta \int_\Omega \xi^2 |D_k^h Du|^2 \, \mathrm{d}x \, .$$

Since $a^{ij}$ is $C^1$ and $0 \le \xi \le 1$ smooth,

$$|A_2| \le C \int_\Omega \Big( \xi |D_k^h Du| \cdot |D_k^h u| + \xi |D_k^h Du| \cdot |Du| + \xi |D_k^h u| \cdot |Du| \Big) \, \mathrm{d}x \, .$$

Again using $ab \le \varepsilon a^2 + b^2/4\varepsilon$, and allowing $C$ to depend on $\varepsilon$ for a moment,

$$\le \varepsilon \int_\Omega \xi^2 |D_k^h Du|^2 \, \mathrm{d}x + C \int_W |D_k^h u|^2 + |Du|^2 \, \mathrm{d}x \, .$$

It's proved in [**Evans**, Thm 5.8.3] that $\|D_k^h u\|_{L^2(W)} \le C \|Du\|_{L^2(\Omega)}$. Taking $\varepsilon = \theta/2$,

$$\le \frac{\theta}{2} \int_\Omega \xi^2 |D_k^h Du|^2 \, \mathrm{d}x + C \int_\Omega |Du|^2 .$$

Combining estimates on $A_1$ and $A_2$,

$$A \ge \frac{\theta}{2} \int_\Omega \xi^2 |D_k^h Du|^2 \, \mathrm{d}x - C \int_\Omega |Du|^2 \, \mathrm{d}x \, .$$

For $B$ we have, since $b^i, c$ are bounded,

$$|B| \le C \int_\Omega (|f| + |Du| + |u|)|v| \, \mathrm{d}x) .$$

We also compute using [**Evans**, Thm 5.8.3],

$$\begin{aligned} \int_\Omega |v|^2 \, \mathrm{d}x &\le C \int_Z Z |D(\xi^2 D_k^h u)|^2 \, \mathrm{d}x \\ &\le C \int_W |D_k^h u|^2 + \xi^2 |D_k^h Du|^2 \, \mathrm{d}x \\ &\le C \int_\Omega |Du|^2 + \xi^2 |D_k^h Du|^2 \, \mathrm{d}x \, . \end{aligned}$$

Plugging this into our estimate for $|B|$ and using our AM-GM inequality,

$$|B| \le \varepsilon \int_\Omega \xi^2 |D_k^h Du|^2 \, \mathrm{d}x + \frac{C}{\varepsilon} \int_\Omega (f^2 + u^2 + |Du|^2) \, \mathrm{d}x \, .$$

Setting $\varepsilon = \theta/4$,

$$\le \frac{\theta}{4} \int_\Omega \xi^2 |D_k^h Du|^2 \, \mathrm{d}x + C \int_\Omega (f^2 + u^2 + |Du|^2) \, \mathrm{d}x \, .$$

We conclude combining our estimates on $A$ and $B$ that,

$$\int_V |D_k^h Du|^2 \, \mathrm{d}x \le \int_\Omega \xi^2 |D_k^h Du|^2 \, \mathrm{d}x \le C \int_\Omega (f^2 + u^2 + |Du|^2) \, \mathrm{d}x \, .$$

From [**Evans**, Thm 5.8.3], we conclude $Du \in H^1_{\mathrm{loc}}(\Omega; \mathbb{R}^n)$. Hence $u \in H^2_{\mathrm{loc}}(\Omega)$ and,

$$\|u\|_{H^2(V)} \le C(\|f\|_{L^2(\mathbb{Z})} + \|u\|_{H^1(\Omega)}) .$$

By the same argument, we can bound by the norms on $W$ instead of $\Omega$, where $C$ now depends on $W$. Pick a bump function $0 \le \zeta \le 1$ vanishing off of $\Omega$ and one on $W$. Then using $B[u,v] = \langle f, v\rangle$ with $v = \zeta^2 u$,

$$\int_\Omega \zeta^2 |Du|^2 \,\mathrm{d}x \le C \int_\Omega (f^2 + u^2)\,\mathrm{d}x\,.$$

Hence,

$$\|u\|_{H^1(W)} \le C(\|f\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)}).$$

Combined with the inequality above, this yields the estimate we were looking for. □

**Theorem 3.17: Higher Interior Regularity** [Evans, Thm. 6.3.2]

Let $m$ be a non-negative integer. Assume $a^{ij}, b^i, c \in C^{m+1}(\Omega)$ and $f \in H^m(\Omega)$. If $u \in H^1(\Omega)$ is a weak solution of $Lu = f$, then $u \in H^{m+2}_{\text{loc}}(\Omega)$. For each $V \Subset \Omega$, there is a constant $C(m, \Omega, V, L)$ so that,

$$\|u\|_{H^{m+2}(V)} \le C(\|f\|_{H^m(\Omega)} + \|u\|_{L^2(\Omega)}).$$

Proof. We work inductively. The case $m = 0$ is the $H^2$ regularity we just proved. Suppose our theorem hold for some $m \ge 0$ for all $\Omega$ and operators $L$. Now suppose $a^{ij}, b^i, c \in C^{m+2}(\Omega)$, $f \in H^{m+1}(\Omega)$ and $u \in H^1(\Omega)$ is a weak solution of $Lu = f$.

We know by inductive assumption that $u \in H^{m+2}_{\text{loc}}(\Omega)$ with,

$$\|u\|_{H^{m+2}(W)} \le C(\|f\|_{H^m(\Omega)} + \|u\|_{L^2(\Omega)}) \quad \text{for all} \quad W \Subset \Omega.$$

Now fix $V \Subset W \Subset \Omega$ and a multi-index $|\alpha| = m+1$. Choose a test function $\widetilde{w} \in C_c^\infty$ and let $w = (-1)^{|\alpha|} D^\alpha \widetilde{w}$. Let $\widetilde{u} = D^\alpha u \in H^1(W)$. We know that $B[u,w] = \langle f, w\rangle_{L^2(\Omega)}$. Integrating by parts, one obtains, $B[\widetilde{u}, \widetilde{w}] = \langle \widetilde{f}, \widetilde{w}\rangle_{L^2(W)}$ where,

$$\widetilde{f} = D^\alpha f - \sum_{\substack{\beta \le \alpha \\ \beta \ne \alpha}} \binom{\alpha}{\beta} \left[ -\sum_{i,j} (D^{\alpha-\beta} a^{ij} D^\beta u_{x_i})_{x_j} + \sum_i D^{\alpha-\beta} b^i D^\beta u_{x_i} + D^{\alpha-\beta} c D^\beta u \right].$$

Note $\widetilde{u}$ is a weak solution of $L\widetilde{u} = \widetilde{f}$ on $W$. Using our estimate for $\|u\|_{H^{m+2}(W)}$ and the regularity of $f, a^{ij}, b^i, c$, we conclude $\widetilde{f} \in L^2(W)$ with,

$$\|\widetilde{f}\|_{L^2(W)} \le C(\|f\|_{H^{m+1}(W)} + \|u\|_{L^2(\Omega)}).$$

By our $H^2$ regularity theorem, $\widetilde{u} \in H^2(V)$ with,

$$\|\widetilde{u}\|_{H^2(V)} \le C(\|\widetilde{f}\|_{L^2(W)} + \|\widetilde{u}\|_{L^2(W)}) \le C(\|f\|_{H^{m+1}(\Omega)} + \|u\|_{L^2(\Omega)}).$$

Since this is true for any multi-index $|\alpha| = m+1$, we conclude that $u \in H^{m+3}(V)$ with,

$$\|u\|_{H^{m+3}(V)} \leq C(\|f\|_{H^{m+1}(\Omega)} + \|u\|_{L^2(\Omega)}).$$

□

**Theorem 3.18: Smooth Interior Regularity** [Evans, Thm. 6.3.3]

If $a^{ij}, b^i, c$ and $f$ are smooth and $u \in H^1(\Omega)$ is a weak solution of $Lu = f$ on $\mathbb{Z}$, then $u \in C^\infty(\Omega)$.

PROOF. We know by the previous theorem that $u \in H^m_{\text{loc}}(\Omega)$ for all $m \in \mathbb{N}$. By the general Sobolev inequalities, $u \in C^\ell(\Omega)$ for each $\ell \in \mathbb{N}$. □

Note that we did not discuss boundary conditions, showing that any singularities of $u$ on the boundary do not propagate to the interior. We now return to the Dirichlet problem and show in this case under some assumptions we can improve from local to global $H^2$ regularity (and then $H^m$ regularity for all $m$).

**Theorem 3.19: Boundary $H^2$-regularity** [Evans, Thm. 6.3.4]

Suppose $\partial\Omega$ is $C^2$. Assume $a^{ij} \in C^1(\overline{\Omega})$, $b^i, c \in L^\infty(\Omega)$, and $f \in L^2(\Omega)$. If $u \in H^1_0(\Omega)$ is a weak solution of (3.1) then $u \in H^2(\Omega)$ with, for a constant $C = C(\Omega, L)$,

$$\|u\|_{H^2(\Omega)} \leq C(\|f\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)}).$$

PROOF. We first deal with the unit half-ball $H = D^n \cap \mathbb{R}^n_+$ and $V \subset H$ the half-ball of radius $\frac{1}{2}$, $B^n_{1/2}(0) \cap \mathbb{R}^n_+$. Pick a smooth bump function $0 \leq \xi \leq 1$ on $\mathbb{R}^n$ which is 1 on $B^n_{1/2}(0)$ and 0 outside of $D^n$.

$u$ is a weak solution of (3.1) and so $B[u, v] = \langle f, v\rangle_{L^2(H)}$ for all $v \in H^1_0(H)$. Hence,

$$\sum_{i,j} \int_H a^{ij} u_{x_i} v_{x_j} \, \mathrm{d}x = \int_H \left( f - \sum_i b^i u_{x_i} - cu \right) v \, \mathrm{d}x \, .$$

We denote $A$ and $B$ for the left and right side of this equality.

Take $h > 0$ small and any $k \in \{1, \ldots, n-1\}$. We choose,

$$v = -D_k^{-h}(\xi^2 D_k^h u).$$

We have, for $x \in H$,

$$\begin{aligned} v(x) &= -\frac{1}{h} D_k^{-h}\Big(\xi^2(x)[u(x+he_k) - u(x)]\Big) \\ &= \frac{1}{h^2}\Big(\xi^2(x-he_k)\big[u(x) - u(x-he_k)\big] - \xi^2(x)\big[u(x+he_k) - u(x)\big]\Big) \end{aligned}$$

Since $u = 0$ along $\mathbb{R}^n_+$ and $\xi = 0$ near $\partial D^n$, we have $v \in H^1_0(H)$.

Estimating $A$ and $B$ as in the proof of interior regularity one obtains,

$$A \geq \frac{\theta}{2} \int_H \xi^2 |D_k^h Du|^2 \, \mathrm{d}x - C \int_H |Du|^2 \, \mathrm{d}x$$

and,

$$|B| \leq \frac{\theta}{4} \int_H \xi^2 |D_k^h Du|^2 \, \mathrm{d}x + C \int_H (f^2 + u^2 + |Du|^2) \, \mathrm{d}x \, .$$

The equality $A = B$ yields, for $k = 1, \ldots, n-1$,

$$\int_V |D_k^h Du|^2 \, \mathrm{d}x \leq C \int_H (f^2 + u^2 + |Du|^2) \, \mathrm{d}x \, .$$

From [**Evans**, Thm. 5.8.3], $u_{x_k} \in H^1(V)$ with,

$$\sum_{\substack{k,\ell=1 \\ k+\ell<2n}}^{n} \|u_{x_k x_\ell}\|_{L^2(V)} \leq C(\|f\|_{L^2(H)} + \|u\|_{H^1(H)}).$$

We just need an estimate on $u_{x_n x_n}$. From interior regularity, we know that $Lu = f$ holds pointwise almost everywhere on $H$. Rewriting our PDE in non-divergence form, this implies,

$$-\sum_{i,j} a^{ij} u_{x_i} u_{x_j} + \sum_i \widehat{b}^i u_{x_i} + cu = f,$$

for $\widehat{b}^i = b^i - \sum_j a^{ij}_{x_j}$. Hence,

$$a^{nn} u_{x_n x_n} = -\sum_{\substack{i,j=1 \\ i+j<2n}}^{n} a^{ij} u_{x_i x_j} + \sum_i \widehat{b}^i u_{x_i} + cu - f.$$

Uniform ellipticity applied to the basis vector $e_n$ gives $a^{nn}(x) \geq \theta > 0$ uniformly in $x \in H$. Since our coefficients our bounded,

$$|u_{x_n x_n}| \leq C \left( \sum_{\substack{i,j=1 \\ i+j<2n}}^{n} |u_{x_i x_j}| + |Du| + |u| + |f| \right).$$

Hence, as with interior regularity, $\|u\|_{H^2(V)} \leq C(\|f\|_{L^2(H)} + \|u\|_{L^2(H)})$.

We now wish to extend to general $\Omega$. Pick a $C^2$ boundary chart $U$ for $\Omega$, $\Phi : U \to \mathbb{R}^n$ so that, $|\det \Phi| = 1$, the boundary maps to $\partial\mathbb{R}^n_+$, and $\Phi(\Omega \cap U) \supset H$, for $H$ some small half-ball centred at the origin.

Let $\widetilde{u}(x) = u(\Phi^{-1}(x))$ be a function on $H$. We have $u' \in H^1(H)$ and $u' = 0$ on $\mathbb{R}^n_+ \cap \partial H$ in the trace sense.

We define new coefficients after our change of variables,

$$\begin{aligned}
\widetilde{a}^{k\ell}(x) &= \sum_{r,s} a^{rs}(\Phi^{-1}(x))\, \Phi^k_{x_r}(\Phi^{-1}(x))\, \Phi^\ell_{x_2}(\Phi^{-1}(y)),\\
\widetilde{b}^k(x) &= \sum_r b^r(\Phi^{-1}(x))\, \Phi^k_{x_r}(\Phi^{-1}(x)),\\
\widetilde{c}(x) &= c(\Phi^{-1}(x)), \quad \text{and} \quad \widetilde{f}(x) = f(\Phi^{-1}(x)).
\end{aligned}$$

Defining the linear operator,

$$\widetilde{L}\widetilde{u} = -\sum_{k,\ell} (\widetilde{a}^{k\ell}\widetilde{u}_{x_k})_{x_\ell} + \sum_k \widetilde{b}^k \widetilde{u}_{x_k} + \widetilde{c}\widetilde{u},$$

we have that, $\widetilde{u}$ is a weak solution of $\widetilde{L}\widetilde{u} = \widetilde{f}$. Seeing this is just a matter of applying the change of variables theorem and using the unit determinant condition to conclude,

$$\widetilde{B}[\widetilde{u}, \widetilde{v}] = B[u, v] = \langle f, v\rangle_{L^2(\Phi^{-1}(H))} = \langle \widetilde{f}, \widetilde{v}\rangle_{L^2(H)}.$$

We claim $\widetilde{L}$ is also uniformly elliptic. Indeed one has,

$$\sum_{k,\ell} \widetilde{a}^{k,\ell}(x)\xi_k\xi_\ell = \sum_{r,s} a^{rs}(\Phi^{-1}(x))\eta_r\eta_s \geq \theta|\eta|^2,$$

where $\eta = \xi D\Phi$. We clearly have $|\xi| \leq C|\eta|$, from which we conclude ellipticity. Also the coefficients $\widetilde{a}^{k,\ell}$ are $C^1$, since $\Phi$ is $C^2$, and the $\widetilde{b}^k, \widetilde{c}$ are bounded.

We can thus apply the case of the half-ball above to conclude,

$$\|\widetilde{u}\|_{H^2(V)} \leq C(\|\widetilde{f}\|_{L^2(H)} + \|\widetilde{u}\|_{L^2(H)}).$$

Changing variables,

$$\|u\|_{H^2(\Phi^{-1}(V))} \leq C(\|f\|_{L^2(U)} + \|u\|_{L^2(U)}).$$

We may cover $\Omega$ with finitely many boundary charts and an interior chart and combine the above inequality with the interior regularity inequality to obtain the desired result and conclude $u \in H^2(\Omega)$. $\square$

**Theorem 3.20: Higher Boundary Regularity [Evans, Thm. 6.5.3]**

Let $m$ be a non-negative integer so that $a^{ij}, b^i, c \in C^{m+1}(\overline{\Omega})$, $f \in H^m(\Omega)$ and $\partial\mathbb{Z}$ is $C^{m+2}$. Suppose $u \in H_0^1(\Omega)$ is a weak solution of (3.1). Then, $u \in H^{m+2}(\Omega)$ and there is a constant $C = C(m, \Omega, L)$ so that,

$$\|u\|_{H^{m+2}(\Omega)} \leq C(\|f\|_{H^m(\Omega)} + \|u\|_{L^2(\Omega)}).$$

PROOF. Again we start with the case of a half ball $H$ and $V \subset H$ a smaller half-ball. We work by induction. The case $m = 0$ was proved in the last theorem. Suppose that $a^{ij}, b^i, c \in C^{m+2}(\overline{H})$, $f \in H^m(H)$ and $u$ is a weak solution of $Lu = f$ on $H$ vanishing on $\partial\mathbb{R}^n_+$ in the trace sense. Let $W = B_0(r) \cap \mathbb{R}^n_+$ be a half ball with $V \subset W \subset H$. By inductive assumption, $u \in H^{m+2}(W)$ with,

$$\|u\|_{H^{m+2}(W)} \leq C(\|f\|_{H^m(H)} + \|u\|_{L^2(H)}).$$

By interior regularity $u \in H^{m+3}_{\text{loc}}(H)$.

Now consider a multi-index $|\alpha| = m + 1$ so that $\alpha_n = 0$. We know that $\widetilde{u} = D^\alpha u \in H^1(H)$ and vanishes on $\partial\mathbb{R}^n_+$ in the trace sense. As in the proof of higher interior regularity, we find that $\widetilde{u}$ is a weak solution of $L\widetilde{u} = \widetilde{f}$ on $H$, where

$$\widetilde{f} = D^\alpha f - \sum_{\substack{\beta \leq \alpha \\ \beta \neq \alpha}} \binom{\alpha}{\beta} \left[ \sum_{i,j} -\left(D^{\alpha-\beta} a^{ij} D^\beta u_{x_i}\right)_{x_j} + \sum_i D^{\alpha-\beta} b^i D^\beta u_{x_i} + D^{\alpha-\beta} c D^\beta u \right].$$

Using the estimate for $\|u\|_{H^{m+2}(W)}$ and the regularity of $f, a^{ij}, b^i, c$,

$$\|\widetilde{f}\|_{L^2(W)} \leq C(\|f\|_{H^{m+1}(H)} + \|u\|_{L^2(H)}),$$

and so $\widetilde{f} \in L^2(W)$. By our $H^2$ boundary regularity theorem,

$$\|\widetilde{u}\|_{H^2(V)} \leq C(\|\widetilde{f}\|_{L^2(W)} + \|\widetilde{u}\|_{L^2(W)}) \leq C(\|f\|_{H^{m+1}(H)} + \|u\|_{L^2(H)}).$$

We conclude,

$$\|D^\beta u\|_{L^2(V)} \leq C(\|f\|_{H^{m+1}(H)} + \|u\|_{L^2(H)}),$$

for any $\beta$ with $|\beta| = m + 3$ and $\beta_n \leq 2$.

To remove the restriction on $\beta_n$, suppose by induction,

$$\|D^\beta u\|_{L^2(V)} \leq C(\|f\|_{H^{m+1}(H)} + \|u\|_{L^2(H)}),$$

for any $\beta$ with $|\beta| = m + 3$ and $\beta_n \leq j$. Now consider $\beta$ with $|\beta| = m + 3$ and $\beta_n = j + 1$. We may write $\beta = \gamma + \delta$, where $\delta = (0, \ldots, 0, 2)$. We have

$|\gamma| = m+1$ and since $u \in H^{m+3}_{\text{loc}}(H)$ and $Lu = f$ we have $D^\gamma Lu = D^\gamma f$ almost everywhere on $H$. Hence,

$$D^\gamma Lu = a^{nn} D^\beta u + \{\text{etc.}\},$$

where each term in "etc." has at most $m+3$ derivatives of $u$, at most $j$ of which are with respect to $x_n$.

Since $a^{nn} \geq \theta > 0$ by ellipticity and by the estimate on $\|D^\beta u\|_{L^2(V)}$ above,

$$\|D^\beta u\|_{L^2(V)} \leq C(\|f\|_{H^{m+1}(H)} + \|u\|_{L^2(H)})$$

for $|\beta| = m+3$ and $\beta_n = j+1$.

By induction, we get this equality for any $|\beta| = m+3$. Hence,

$$\|u\|_{H^{m+3}(H)} \leq C(\|f\|_{H^{m+1}(H)} + \|u\|_{L^2(H)}).$$

To extend to general domains $\Omega$, simply cover $\partial\Omega$ by $C^{m+2}$ boundary charts and apply the change of variables argument of the last proof, along with its conclusion for the interior of $\Omega$ to obtain the desired result. □

**Theorem 3.21: Smooth Boundary Regularity** [Evans, Thm. 6.3.6]

Suppose $a^{ij}, b^i, c, f \in C^\infty(\overline{\Omega})$ and $\partial\Omega$ is smooth. If $u \in H^1_0(\Omega)$ is a weak solution of (3.1), then $u \in C^\infty(\overline{\Omega})$.

Proof. By the previous theorem, $u \in H^m(\Omega)$ for all $m \in \mathbb{N}$. By the general Sobolev inequalities, $u \in C^k(\overline{\Omega})$ for each $k \in \mathbb{N}$. □

**3.4. Maximum Principles.** So far our results have come about as a result of so called *energy methods*, where we apply ellipticity to bound various integrals. Now we study *maximum principle methods* that come from pointwise constraints. We know that if $x_0$ is a local maximum of a $C^2$ function $u$ then $Du(x_0) = 0$ and $D^2u(x_0) \leq 0$. If $u$ is a $C^2$ solution of a homogeneous second order elliptic problem, then we can use the elliptic equation $Lu = 0$ to restrict the existence of such a point.

For our purposes, it is useful to put our elliptic operator $L$ in non-divergence form, and assume our coefficients are continuous with the $a^{ij}$ symmetric.

**Theorem 3.22: Weak Maximum Principle [Evans, Thm. 6.4.1]**

Let $u \in C^2(\Omega) \cap C(\overline{\Omega})$. Suppose the operator $L$ has zeroth order coefficient $c = 0$ on $\Omega$. If $u$ is a *subsolution*, i.e. $Lu \leq 0$, then it attains its maximum on the boundary,

$$\max_{\overline{\Omega}} u = \max_{\partial\Omega} u.$$

If $u$ is a *supersolution*, $Lu \geq 0$, it attains its minimum the boundary,

$$\min_{\overline{\Omega}} u = \min_{\partial\Omega} u.$$

PROOF. Suppose $Lu < 0$ on $\Omega$ with $x_0 \in \Omega$ obtaining the maximum of $u$ on $\overline{\Omega}$. We must have $Du(x_0) = 0$ and $D^2u(x_0) \leq 0$.

We can make an orthogonal change of basis $U$ so that $(a^{ij}(x_0))$ becomes diagonal with positive entries $\lambda_k$. If $y = x_0 + U(x - x_0$, after this change of basis we have,

$$\sum_{i,j} a^{ij}(x_0) u_{x_i x_j}(x_0) = \sum_k \lambda_k u_{y_k y_k}(x_0) \leq 0,$$

since $D^2u(x_0) \leq 0$. Hence,

$$Lu = -\sum_{i,j} a^{ij}(x_0) u_{x_i x_j}(x_0) + \sum_i b^i(x_0) \cancelto{0}{u_{x_i}(x_0)} \geq 0.$$

This gives a contradiction. To deal with the case $Lu \leq 0$, define $u^\varepsilon(x) = u(x) + \varepsilon e^{\lambda x_1}$ for $\varepsilon, \lambda > 0$ as yet unspecified. By uniform ellipticity, $a^{ii}(x) \geq \theta$ and so,

$$Lu^\varepsilon = Lu + \varepsilon L(e^{\lambda x_1}) \leq \varepsilon e^{\lambda x_1}\left[-\lambda^2 a^{11} + \lambda b^1\right] \leq \varepsilon e^{\lambda x_1}\left[-\lambda^2\theta + \lambda\|\mathbf{b}\|_{L^\infty}\right].$$

The last expression is negative if $\lambda > 0$ is large enough. We conclude by above that $\max_{\overline{\Omega}} u^\varepsilon = \max_{\partial\Omega} u^\varepsilon$. Taking $\varepsilon \to 0$ gives the conclusion. The supersolution case follows by changing the sign of $u$. □

Define the positive and negative parts of a function:

$$u^+(x) = \max\{u(x), 0\} \quad \text{and} \quad u^- = -\min\{u(x), 0\}.$$

**Theorem 3.23: Weak Maximum Principle with $c \geq 0$**

Let $u \in C^2(\Omega) \cap C(\overline{\Omega})$. Suppose the operator $L$ has zeroth order coefficient $c \geq 0$ on $\Omega$. If $Lu \leq 0$, then

$$\max_{\overline{\Omega}} u \leq \max_{\partial\Omega} u^+.$$

If $Lu \geq 0$, then

$$\min_{\overline{\Omega}} u \geq -\max_{\partial\Omega} u^-.$$

PROOF. Let $u$ be a subsolution. If $V = \{x \in \Omega : u(x) > 0\}$, we have $Ku := Lu - cu \leq -cu \leq 0$ on $V$. By the previous theorem, for $V \neq \varnothing$,

$$\max_{\overline{\Omega}} u = \max_{\overline{V}} u = \max_{\partial V} u = \max_{\partial\Omega} u^+,$$

giving the result. If $V = \varnothing$ the conclusion is immediate. For the supersolution case, we can again just change the sign of $u$. □

LEMMA 3.24 (Hopf's Lemma). *Let $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$. Suppose the operator $L$ has zeroth order coefficient $c = 0$ on $\Omega$. Suppose $u$ is a subsolution $Lu \leq 0$ on $\Omega$ and there is a point $x_0 \in \partial\Omega$ with $u(x_0) > u(x)$ for all $x \in \Omega$. Suppose also $x_0$ lies on the boundary of a ball $B \subset \Omega$. Then, $\frac{\partial u}{\partial \nu} > 0$ where $\nu$ is the outer unit normal to $B$ at $x_0$.*

*If $c \geq 0$ in $\Omega$ then the same conclusion holds assuming $u(x_0) \geq 0$.*

REMARK 3.25. The non-trivial part of this is the strictness of the inequality $\frac{\partial u}{\partial \nu} > 0$. △

PROOF. By translation $B$ is centred at the origin with radius $r > 0$. Define,

$$v(x) = e^{-\lambda|x|^2} - e^{-\lambda r^2}$$

on $\overline{B}$ with $\lambda > 0$ to be determined. By uniform ellipticity,

$$\begin{aligned} Lv &= -\sum_{i,j} a^{ij} v_{x_i x_j} + \sum_i b^i v_{x_i} + cv \\ &= e^{-\lambda|x|^2} \sum_{i,j} a^{ij}(-4\lambda^2 x_i x_j + 2\lambda\delta_{ij}) \\ &\qquad - e^{-\lambda|x|^2} \sum_i b^i 2\lambda x_i + c(e^{-\lambda|x|^2} - e^{-\lambda r^2}) \\ &\leq e^{-\lambda|x|^2}\big(-4\theta\lambda^2|x|^2 + 2\lambda \operatorname{tr} \mathbf{A} + 2\lambda|\mathbf{b}| \cdot |x| + c\big). \end{aligned}$$

Let $R \subset B$ be the annulus given by removing the ball of radius $r/2$. On $R$, for $\lambda > 0$ large enough,

$$Lv \leq e^{-\lambda|x|^2}\left(-\theta\lambda^2 r^2 + 2\lambda \operatorname{tr} \mathbf{A} + 2\lambda|\mathbf{b}|r + c\right) \leq 0.$$

Since $u(x_0) > \max_\Omega u$, there is $\varepsilon > 0$ so that $u(x_0) \geq u(x) + \varepsilon v(x)$ for all $x$ on the sphere $\partial B(r/2)$. Since $v = 0$ on $\partial B$, this inequality also holds on $\partial B$.

From the inequality on the annulus, we have on $R$,

$$L(u + \varepsilon v - u(x_0)) \leq -cu(x_0) \leq 0,$$

We also see that $u + \varepsilon v - u(x_0) \leq 0$ on $\partial R$.

By our weak maximum principle we know $u + \varepsilon v - u(x_0) \leq 0$ on $R$. Since $u(x_0) + \varepsilon v(x_0) - u(x_0) = 0$, we have,

$$\frac{\partial u}{\partial \nu} + \varepsilon \frac{\partial v}{\partial \nu} \geq 0.$$

Hence,

$$\frac{\partial u}{\partial \nu} \geq -\varepsilon \frac{\partial v}{\partial \nu} = -\frac{\varepsilon}{r} Dv(x_0) \cdot x_0 = 2\lambda\varepsilon r e^{-\lambda r^2} > 0.$$

□

**Theorem 3.26: Strong Maximum Principle** [Evans, Thm. 6.4.3]

Let $u \in C^2(\Omega) \cap C(\overline{\Omega})$. Suppose the operator $L$ has zeroth order coefficient $c = 0$ on $\Omega$. Further suppose $\Omega$ is connected, open, and bounded.

(1) If $Lu \leq 0$ on $\Omega$ and $u$ obtains its maximum on $\overline{\Omega}$ within the interior, then $u$ is constant.
(2) If $Lu \geq 0$ on $\Omega$ and $u$ obtains its minimum on $\overline{\Omega}$ within the interior, then $u$ is constant.
(3) If we instead just have $c \geq 0$, then same conclusions hold assuming in (i) the maximum is non-negative and in (ii) the minimum is non-positive.

Proof. Let $M = \max_\Omega u$ and $C = u|_\Omega^{-1}(M)$. Assuming $u$ is not constantly equal to $M$, let $V = \Omega \setminus C$. Now take $y \in V$ so that $d(y, C) < d(y, \partial\Omega)$ and let $B$ be the largest open ball centred at $y$ contained in $V$. We thus have $x_0 \in C \cap \partial B$. We clearly have $u(x_0) > u|_V$ and so Hopf's lemma implies $\frac{\partial u}{\partial \nu} > 0$. On the other hand, $u$ had a local maximum at $x_0 \in \Omega$ and so $Du(x_0) = 0$. We have a contradiction and so $u \equiv M$. The case (ii) follows after changing $u$ by a sign. And (iii) follows using the sam argument and the statement of Hopf's lemma for $c \geq 0$. □

**Theorem 3.27: Harnack's inequality** [Evans, Thm. 6.4.3]

Assume $u \geq 0$ is a $C^2$ solution of $Lu = 0$ on $\Omega$ and $V \Subset \Omega$ is connected. There is a constant $C = C(V, L)$ so that,

$$\sup_V u \leq C \inf_V u.$$

We only prove the case when $b^i = c = 0$ and the $a^{ij}$ are smooth.

PROOF. We assume $u > 0$, as otherwise we could prove the result for $u + \varepsilon$ and then take $\varepsilon$ to zero. Let $v = \log u$. We have,

$$\sum_{i,j} a^{ij} v_{x_i x_j} + \sum_{i,j} a^{ij} v_{x_i} v_{x_j} = 0.$$

Let $w$ be the second term (or equivalently minus the first). We have,

$$w_{x_k x_\ell} = \sum_{i,j} 2a^{ij} v_{x_i x_k x_\ell} v_{x_j} + 2a^{ij} v_{x_i x_k} v_{x_j x_\ell} + R.$$

Here $R$ consists of terms where we differentiate the coefficients, and satisfies an estimate,

$$|R| \leq \varepsilon |D^2 v|^2 + C(\varepsilon)|Dv|^2,$$

for each $\varepsilon > 0$. Then,

$$-\sum_{k,\ell} a^{k\ell} w_{x_k x_\ell} = 2 \sum_{i,j} a^{ij} v_{x_j} \left( -\sum_{k,\ell} a^{\ell k} v_{x_i x_k x_\ell} \right) - 2 \sum_{k,\ell} a^{ij} a^{k\ell} v_{x_i x_k} v_{x_j x_\ell} - R.$$

Differentiating the other formula for $w$, for $R_i$ satisfying the same estimate,

$$-\sum_{k,\ell} a^{\ell k} v_{x_i x_k x_\ell} = w_{x_i} + R_i.$$

By uniform ellipticity,

$$\sum_{k,\ell} a^{ij} a^{k\ell} v_{x_i x_k} v_{x_j x_\ell} \geq \theta^2 |D^2 v|^2.$$

Combining the last two formulae with the one above it, we obtain for $\varepsilon$ small enough,

$$-\sum_{k,\ell} a^{k\ell} w_{x_k x_\ell} + \sum_k b^k w_{x_k} \geq \theta^2 |D^2 v|^2 - C|Dv|^2, \tag{3.4}$$

where,

$$b^k = -2 \sum_\ell a^{k\ell} v_{x_\ell}.$$

Now let $V \subset\subset \Omega$ be an open ball of radius $r$. Let $0 \leq \xi \leq 1$ be a smooth bump function which is one on $V$ and vanishes on $\partial\Omega$. Set $z = \xi^4 w$.

Assume $z$ obtains a maximum at some $x_0 \in \Omega$. We thus have,

$$\xi w_{x_k} + 4\xi_{x_k} w = 0 \quad \text{for all} \quad k = 1, \ldots, n.$$

We also have the inequality at $x_0$,

$$0 \leq -\sum_{k,\ell} a^{k\ell} z_{x_k x_\ell} + \sum_k b^k z_{x_k}.$$

Hence,

$$0 \leq \xi^4 \left( -\sum_{k,\ell} a^{k\ell} w_{x_k x_\ell} + \sum_k b^k w_{x_k} \right) + \widehat{R}, \tag{3.5}$$

where $\widehat{R}$ consists of terms where we differentiate $\xi$ and satisfies,

$$|\widehat{R}| \leq C\xi^2 w + \widetilde{C}\xi^3 |Dw|.$$

Since $\xi w_{x_k} = -4\xi_{x_k} w$ as above, we may take $\widetilde{C} = 0$. Plugging this and (3.4) into (3.5) gives,

$$\xi^4 |D^2 v|^2 \leq C\xi^4 |Dv|^2 + C\xi^2 w.$$

We have $\theta |Dv|^2 \leq w$ and by definition of $w$, $w \leq C|D^2 v|$. Hence, $\xi^4 w^2 \leq C\xi^2 w$. So, $z = \xi^4 w \leq C$ at $x_0$. But $z$ had its maximum at $x_0$ and we know $z = w$ on $V$. Hence, $|Dv| \leq C$ on $V$.

Given any $x_1, x_2 \in V$ we conclude,

$$v(x_2) - v(x_1) \leq \sup_V |Dv| r \leq C.$$

Exponentiating, $u(x_2) \leq u(x_1) e^C$. This proves the inequality for $V$ a ball. For general $V \Subset \Omega$, we may cover $V$ by finitely many balls within $\Omega$. □

## 4. The Calculus of Variations

We now turn to studying variational problems. That is, PDEs which arise by looking for the critical points of some functional. This is a rich source of geometrically and physically motivated PDEs. For example, all the equations of motion for classical mechanical systems and classical field theories arise in this manner. The existence of some functional giving rise to the PDE also gives important technical tools for helping to solve these problems; for example minimizing the functional will lead to a solution of the associated PDE.

**4.1. Variational Problems and the Euler–Lagrange Equations.** We will consider PDEs $A[u] = 0$ which arise from a functional $I[u]$ in the sense that $A[u] = I'[u]$ and so solving the PDE is equivalent to finding critical points of the functional.

Most of the examples we consider take the following form. Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with smooth boundary, and consider some smooth function,

$$L : \mathbb{R}^n \times \mathbb{R} \times \Omega \to \mathbb{R},$$

which we call the *Lagrangian*. We will use as coordinates on $L$ the variables $(p, z, x) = (p_1, \dots, p_n, z, x_1, \dots, x_n)$. The Lagrangian then produces an *action functional*,

$$I[w] = \int_\Omega L(Dw(x), w(x), x)\,\mathrm{d}x\,, \tag{4.1}$$

which is defined on the space of smooth function $w : \Omega \to \mathbb{R}$ satisfying a Dirichlet boundary condition $w|_{\partial\Omega} = g$.

If $u$ is a minimizer (or stationary point) of (4.1) among functions $w$ as above, then $u$ will automatically solve a certain second-order non-linear PDE.

If $u$ is such a critical point, its directional derivatives must vanish. So let $v \in C_c^\infty(\Omega)$ be arbitrary and we must have $i'(0) = 0$, where $i$ is the function,

$$i(\tau) = I[u + tv].$$

We compute,

$$i(\tau) = \int_\Omega L(Du + \tau Dv, u + \tau v, x)\,\mathrm{d}x\,.$$

$$i'(\tau) = \int_U \sum_{i=1}^n L_{p_i}(Du + \tau Dv, u + \tau v, x)v_{x_i} + L_z(Du + \tau Dv, u + \tau v, x)v\,\mathrm{d}x\,.$$

At $\tau = 0$,

$$0 = \int_U \sum_{i=1}^n L_{p_i}(Du, u, x)v_{x_i} + L_z(Du, u, x)v\,\mathrm{d}x\,.$$

Integrating by parts since $v$ has compact support,

$$= \int_U \left( -\sum_{i=1}^n L_{p_i}(Du, u, x)x_i + L_z(Du, u, x) \right) v\,\mathrm{d}x\,.$$

Since this must vanish for any $v$, we must have,

$$\boxed{-\sum_{i=1}^{n}(L_{p_i}(Du,u,x))_{x_i}+L_z(Du,u,x)=0.} \tag{4.2}$$

This is called the *Euler–Lagrange equation* for $I$.

EXAMPLE 4.1. Let us see some examples of this.

**Dirichlet Principle:** Let $L(p,z,x)=\frac{1}{2}|p|^2$. We easily see that the associated action functional is the Dirichlet energy,

$$I[w]=\frac{1}{2}\int_\Omega |Dw|^2\,\mathrm{d}x\,,$$

and the corresponding Euler–Lagrange equation is Laplace's equation $\Delta u=0$.

**Generalized Dirichlet:** Now consider a more general quadratic Lagrangian,

$$L(p,z,x)=\frac{1}{2}\sum_{i,j=1}^{n}a^{ij}(x)p_ip_j-zf(x),$$

where $a^{ij}=a^{ji}$. The action functional is then,

$$I[w]=\int_\Omega \frac{1}{2}\sum_{i,j=1}^{n}a^{ij}w_{x_i}w_{x_j}-wf\,\mathrm{d}x\,,$$

and the Euler–Lagrange equations are,

$$-\sum_{i,j}\partial_{x_i}(a^{ij}u_{x_j}).$$

Under a uniform ellipticity condition on the $a^{ij}$, we will have an existence result for minimizers of $I$. Hence calculus of variations naturally includes the study of second order elliptic PDEs in divergence form, as we studied previously.

**Nonlinear Poisson Equation:** Let $f:\mathbb{R}\to\mathbb{R}$ be smooth with antiderivative $F(z)$ satisfying $F(0)=0$. The functional,

$$I[w]=\int_\Omega \frac{1}{2}|Dw|^2-F(w)\,\mathrm{d}x\,,$$

gives rise to Euler–Lagrange equation $-\Delta u=f(u)$, i.e. the nonlinear Poisson equation.

**Minimal Surfaces:** Consider the Lagrangian $L(p, z, x) = \sqrt{1 + |p|^2}$ with action functional,

$$I[w] = \int_\Omega \sqrt{1 + |Dw|^2}\, \mathrm{d}x\,.$$

This measures the volume of the graph of $w$ as a hypersurface in $\Omega \times \mathbb{R}$. We obtain an Euler–Lagrange equation,

$$\sum_{i=1}^{n} \partial_{x_i} \left( \frac{u_{x_i}}{\sqrt{1 + |Du|^2}} \right) = 0.$$

This is the minimal surface equation in Euclidean space. It asks that the mean curvature of the graph be zero.

**Classical Mechanics:** Let $V(z)$ be some smooth function, which we think of as describing some energy potential. The standard Lagrangian in one-dimensional classical mechanics is given by the difference of kinetic and potential energy,

$$L(v, x, t) = \frac{1}{2} m v^2 - V(x),$$

and the associated action is,

$$I[w] = \int \frac{1}{2} m \dot{w}^2 - V(w)\, \mathrm{d}t\,.$$

The associated Euler–Lagrange equation is,

$$m\ddot{u} = -V'(u).$$

Since force is just minus the derivative of the potential, this is just Newton's law. The generalization to higher dimensions can be obtained by allowing our action functional $I$ to take as inputs multi-component functions and considering the corresponding generalization of the Euler–Lagrange equation; we will discuss this case shortly.

**Yang–Mills Equations:** This example is in a more general setting than the above, when we allow our spaces to be manifolds. Let $E \to M$ be a principal $G$-bundle, for $G$ a compact Lie group, over a smooth manifold. We may consider a $G$-connection $A$ on $E$, locally described by a Lie algebra valued one form, and its associated curvature $F_A = \mathrm{d}A + [A \wedge A]$, locally a Lie algebra valued two form. Choosing a metric on $E$ and the Lie algebra of $G$, we can define a functional on the space of connections,

$$I[A] = \int_M |F_A|^2\, \mathrm{dvol}\,.$$

The associated Euler–Lagrange equation is,

$$\mathrm{d}_A * F_A = 0.$$

Here $*$ is the Hodge star operator and $\mathrm{d}_A$ is the exterior covariant derivative determined by $A$. This is called the *Yang–Mills equations*. It is the classical equation for a gauge field in quantum field theory. If $G = U(1)$, then the Yang–Mills equations along with the Bianchi identity $\mathrm{d}_A F_A = 0$, which always holds, are equivalent to the vacuum Maxwell's equations for the electromagnetic potential $A$. ▲

*Second Variation.* Now suppose $u$ is a local minimum of the functional $I$. Hence for any $v \in C_c^\infty(\Omega)$, the function $i(\tau) = I[u + \tau v]$ will satisfy $i''(0) \geq 0$. We compute,

$$\begin{aligned} i''(\tau) = \int_\Omega \sum_{i,j=1}^n & L_{p_i p_j}(Du + \tau Dv, u + \tau v, x) v_{x_i} v_{x_j} \\ & + 2 \sum_{i=1}^n L_{p_i z}(Du + \tau Dv, u + \tau v, x) v_{x_i} v + L_{zz}(Du + \tau Dv, u + \tau v, x) v^2 \, \mathrm{d}x \,. \end{aligned}$$

At $\tau = 0$,

$$0 \leq \int_\Omega \sum_{i,j=1}^n L_{p_i p_j}(Du, u, x) v_{x_i} v_{x_j} + 2 \sum_{i=1}^n L_{p_i z}(Du, u, x) v_{x_i} v + L_{zz}(Du, u, x) v^2 \, \mathrm{d}x$$

By approximation, this inequality still holds whenever $v$ is Lipschitz continuous and vanishes on $\partial\Omega$. Let $\rho$ be the 1-periodic function given by a symmetric height $1/2$ triangle on $[0, 1]$. For $\xi \in \mathbb{R}^n$, $\varepsilon > 0$, and $\zeta \in C_c^\infty(\Omega)$, set $v(x) = \varepsilon \rho(x \cdot \xi / \varepsilon) \zeta(x)$.

Sending $\varepsilon \to 0$, one has $v_{x_i}(x) = \rho'(x \cdot \xi / \varepsilon) \xi_i \zeta + O(\varepsilon)$. We obtain,

$$0 \leq \int_\Omega \sum_{i,j=1}^n L_{p_i p_j}(Du, u, x)(\rho')^2 \xi_i \xi_j \zeta^2 \, \mathrm{d}x + O)\varepsilon).$$

Since $|\rho'| = 1$ almost everywhere, as $\varepsilon \to 0$, we obtain,

$$0 \leq \int_\Omega \sum_{i,j=1}^n L_{p_i p_j}(Du, u, x) \xi_i \xi_j \zeta^2 \, \mathrm{d}z \,.$$

Hence,

$$\sum_{i,j=1}^n L_{p_i p_j}(Du, u, x) \xi_i \xi_j \geq 0.$$

This requirement relates to a convexity assumption on $L$ which guarantees existence of solutions to our variational problem.

*Higher Dimensions.* We now generalize to systems of equations. Consider a Lagrangian,

$$L(p_1^1, \dots, p_n^m, z^1, \dots, z^m, x_1, \dots, x_n) : \mathbb{R}^{m\times n} \times \mathbb{R}^m \times \Omega \to \mathbb{R}.$$

Here we think of the variables of $\mathbb{R}^{m\times n}$ as defining a $m$ by $n$ matrix with entries $p_j^i$ in the $i$th row and $j$th column. We have a corresponding functional,

$$I[\mathbf{w}] = \int_\Omega L(D\mathbf{w}(x), \mathbf{w}(x), x)\, \mathrm{d}x\,,$$

defined on the space of smooth functions $\mathbf{w} : \Omega \to \mathbb{R}^m$ with fixed boundary conditions.

By the same computation as above, we see that the stationary points $u$ of $I$ must satisfy a quasilinear system of second order PDEs,

$$-\sum_{i=1}^n \partial_{x_i} L_{p_i^k}(D\mathbf{u}, \mathbf{u}, x) + L_{z^k}(D\mathbf{u}, \mathbf{u}, x) \quad \text{for} \quad k = 1, \dots, m, \tag{4.3}$$

called the *Euler–Lagrange equations* for $I$.

*Null Lagrangians.* We will be interested in a certain class of Lagrangians whose action functionals are "topological," i.e. they depend purely on boundary values.

**Definition 4.2**

We say $L$ is a *null Lagrangian* if the corresponding Euler–Lagrange equations (4.3) are satisfied by all smooth functions $u : \Omega \to \mathbb{R}^m$.

**Theorem 4.3**

If $L$ is a null Lagrangian giving an action functional $I[w]$, then for any $C^2$ functions $u$, $\widetilde{u}$ which agree on $\partial\Omega$, one has $I[u] = I[\widetilde{u}]$.

Proof. Let $i(\tau) = I[\tau u + (1 - \tau\widetilde{u})]$. We compute,

$$\begin{aligned} i'(\tau) = \sum_{k=1}^m \int_\Omega \Bigg( &-\sum_{i=1}^n \partial_{x_i} L_{p_i^k}(\tau D\mathbf{u} + (1-\tau)D\widetilde{\mathbf{u}}, \tau\mathbf{u} + (1-\tau)\widetilde{\mathbf{u}}, x) \\ &+ L_{z^k}(\tau D\mathbf{u} + (1-\tau)D\widetilde{\mathbf{u}}, \tau\mathbf{u} + (1-\tau)\widetilde{\mathbf{u}}, x)\Bigg)(u^k - \widetilde{u}^k)\, \mathrm{d}x\,. \end{aligned}$$

This must vanish since the Euler–Lagrange equations hold for $\tau\mathbf{u}+(1-\tau)\widetilde{\mathbf{u}}$. □

In the one-dimensional case ($m = 1$), the only examples of null Lagrangians are functions which are linear in $p$. For $m > 1$, there are other interesting examples. For example the function $L(P) = \det(P)$ for $P \in \mathbb{R}^{n\times n}$ is a null Lagrangian. The short computational proof can be found in [**Evans**, p. 436].

**4.2. Existence.** We now want to show under certain conditions that solutions to our variational problems exist. Note that even in the simplest possible case, plenty of $\mathbb{R} \to \mathbb{R}$ functions do not obtain their infimum. In order that $I[w]$ obtain its minima, we thus need to control the behaviour of $I$ as $|w| \to \infty$, in particular by making sure it grows fast enough.

We assume for some $q > 1$ that,

$$L(p, z, x) \geq \alpha|p|^2 - \beta, \tag{4.4}$$

for some constants $\alpha > 0, \beta \geq 0$ and any $p, z, x$. This will imply,

$$I[w] \geq \alpha\|Dw\|_{L^q}^q - \gamma,$$

where $\gamma = \beta\text{vol}(\Omega)$. We refer to this as a *coercivity condition* on $I$; it implies in particular that $I$ grows to infinity whenever $\|Dw\|_{L^q}$ does.

To expand our space of solutions, we set $\mathcal{A}$ to be the set of functions in $W^{1,q}(\Omega)$ which agree with $g$ on $\partial\Omega$ in the trace sense. Assuming coercivity (4.4) of $L$, the functional $I[w]$ is defined on all of $\mathcal{A}$ (but may be infinite).

Let $m$ be the infimum of $I[w]$ over all $w \in \mathcal{A}$. We can construct a sequence $u_k \in \mathcal{A}$ so that $I[u_k]$ converges to $m$. We hope that on a subsequence $u_k$ will converge to a minimizer, but this requires some compactness. From standard functional analysis, we are more likely to succeed if we work in the weak topology.

Indeed, coercivity implies our sequence lies in a bounded subset of $W^{1,q}(\Omega)$. Since this space is reflexive, Banach–Alaoglu implies there is a subsequence which weakly converges to $u$ in $W^{1,q}(\Omega)$ for some $u \in \mathcal{A}$.

Unfortunately, $I$ is unlikely to be continuous in the weak topology and so we cannot conclude $u$ is a minimizer. We will however be able to conclude this with the following assumption on $I$.

**Definition 4.4**

A functional $I[w]$ on $W^{1,q}(\Omega)$ is *weakly lower semicontinuous* if,

$$I[u] \leq \liminf_{k\to\infty} I[u_k],$$

whenever $u_k$ is a sequence weakly converging to $u$ in $W^{1,q}(\Omega)$.

Thus existence in our variational problem will hold for a coercive Lagrangian if we can show the associated functional is weakly lower semicontinuous.

Recall we had a necessary inequality for the existence of a minimizer,

$$\sum_{i,j=1}^{n} L_{p_i p_j}(Du, u, x)\xi_i\xi_j \geq 0.$$

This motivates an assumption that $L$ be convex in its first entry. Indeed this will suffice to give weak lower semicontinuity.

**Theorem 4.5**

Let $L(p, z, x)$ be a smooth Lagrangian which is bounded below and convex in $p$ for all $z$ and $x$. Then $I[w]$ is weakly lower semicontinuous on $W^{1,q}(\Omega)$.

Proof. Consider a sequence $\{u_k\}$ in $W^{1,q}(\Omega)$ which converges weakly to $u$. Let $\ell$ be the lim inf of $I[u_k]$; we need to show $I[u] \leq \ell$.

We may pass to a subsequence so that $\ell = \lim_{k\to\infty} I[u_k]$. By Rellich–Kondrachov, $u_k$ converge strongly to $u$ in $L^q(\Omega)$ and hence on a subsequence $u_k \to u$ pointwise almost everywhere.

Take $\varepsilon > 0$. Since $u_k \to u$ almost everywhere, Egoroff's theorem implies $u_k \to u$ uniformly on a measurable set $E_\varepsilon$ so that $\mathrm{vol}(\Omega \setminus E_\varepsilon) \leq \varepsilon$. We can choose these compatibly for different $\varepsilon$ so that $E_\delta \subset E_\varepsilon$ when $\delta < \varepsilon$. Now take,

$$F_\varepsilon = \{x \in \Omega : |u(z)| + |Du(z)| \leq \frac{1}{\varepsilon}\}.$$

One will have these sets exhaust $\Omega$ as $\varepsilon \to 0$. Take $G_\varepsilon = E_\varepsilon \cap F_\varepsilon$; these will also exhaust $\Omega$.

By adding a constant, $L \geq 0$. We then have,

$$\begin{aligned} I[u_k] &= \int_\Omega L(Du_k, u_k, x)\,\mathrm{d}x \\ &\geq \int_{G_\varepsilon} L(Du_k, u_k, x)\,\mathrm{d}x\,. \end{aligned}$$

By convexity of $L$ in $p$,

$$\geq \int_{G_\varepsilon} L(Du, u_k, x)\,\mathrm{d}x + \int_{G_\varepsilon} D_p L(Du, u_k, x) \cdots (Du_k - Du)\,\mathrm{d}x\,.$$

Using the uniform converge of $u_k$ on $E_\varepsilon$,

$$\lim_{k\to\infty} \int_{G_\varepsilon} L(Du, u_k, x)\,\mathrm{d}x = \int_{G_\varepsilon} L(Du, u, x)\,\mathrm{d}x\,.$$

We know $D_p L(Du, u_k, x) \to D_p L(Du, u, x)$ uniformly on $G_\varepsilon$ and $Du_k \to Du$ weakly in $L^q$. Hence,

$$0 = \lim_{k\to\infty} \int_{G_\varepsilon} D_p L(Du, u_k, x) \cdots (Du_k - Du)\,\mathrm{d}x\,.$$

The above three equations combine to give,

$$\ell \geq \int_{G_\varepsilon} L(Du, u_k, x)\,\mathrm{d}x\,.$$

Taking $\varepsilon \to 0$ and applying the monotone converge theorem yields our result. □

We now expand on the argument we sketched above to give a complete proof of existence.

**Theorem 4.6: Existence of Minimizers**

Let $L(p, z, x)$ be a Lagrangian which is convex in $p$ and satisfies the coercivity condition (4.4). Then there is at least one $u \in \mathcal{A}$ which obtains the minimum of the action functional $I$.

Proof. Let $m$ be the infimum of $I$, which we can clearly take to be finite. Consider a sequence $u_k$ so that $I[u_k]$ converges to $m$. We can shift $L$ by a constant so that $L \geq \alpha|p|^q$ by coercivity. Hence we always have,

$$I[w] \geq \alpha \int_\Omega |Dw|^q\,\mathrm{d}x\,.$$

We must thus have $\sup_k \|Du_k\|_{L^q(\Omega)}$ is finite.

Fix any $w \in \mathcal{A}$. By the boundary conditions of $\mathcal{A}$, $u_k - w \in W_0^{1,q}(\Omega)$ for each $k$. By Poincaré's inequality, we have,

$$\|u_k\|_{L^q(\Omega)} \leq \|u_k - w\|_{L^q(\Omega)} + \|w\|_{L^q(\Omega)} \leq C\|Du_k - Dw\|_{L^q(\Omega)} + C \leq C.$$

Thus the $u_k$ are uniformly bounded in $L^q(\Omega)$, and hence in $W^{1,q}(\Omega)$.

By Banach–Alaoglu, after passing to a subsequence, $u_k$ converge weakly to some $u$ in $W^{1,q}(\Omega)$. We claim $u \in \mathcal{A}$. Since $u_k - w \in W_0^{1,q}(\Omega)$ and $W_0^{1,q}(\Omega)$ is closed and convex, Mazur's theorem implies $W_0^{1,q}(\Omega)$ is weakly closed Thus $u - w \in W_0^{1,q}(\Omega)$ and so $u|\partial\Omega = g$ in the trace sense. By theorem 4.5, we conclude $I[u] \leq \lim_{k\to\infty} I[u_k] = m$. Thus $I[u] = m$ and $u$ is a minimizer. □

*Uniqueness.* If we wish to require that the minimizer be unique, then we will need additional assumptions on $L$. One way to achieve this is to ask that $L(p, x)$ be independent of $z$ and that $L(p, z)$ is *uniformly convex* in $p$, i.e. there exists $\theta > 0$ with,

$$\sum_{i,j=1}^{n} L_{p_i p_j}(p, x)\xi_i\xi_j \geq \theta|\xi|^2.$$

**Theorem 4.7**

Suppose $L(p, x)$ is a Lagrangian independent of $z$ and uniformly convex in $p$. Then a minimizer $u \in \mathcal{A}$ of $I$ so that $I[u] < \infty$ is unique.

Proof. Suppose $u$ and $\widetilde{u}$ are minimizers of $I$. Let $v$ be their average. We show that $I[v] \leq \frac{1}{2}(I[u] + I[\widetilde{u}])$ and the inequality is strict unless $u = \widetilde{u}$ almost everywhere.

By uniform convexity, we have for any $p, q \in \mathbb{R}^n$ and $x \in \Omega$,

$$L(p, x) \geq L(q, x) + D_p L(q, x)(p - q) + \frac{\theta}{2}|p - q|^2.$$

Taking $p = Du, q = \frac{1}{2}(Du + D\widetilde{u})$ and integrating,

$$I[v] + \int_\Omega D_p L\left(\frac{Du + D\widetilde{u}}{2}, x\right) \cdot \left(\frac{Du - D\widetilde{u}}{2}\right) \mathrm{d}x + \frac{\theta}{8}\int_\Omega |Du - D\widetilde{u}|^2 \,\mathrm{d}x \leq I[u].$$

If we instead take $p = D\widetilde{u}$, we get,

$$I[v] + \int_\Omega D_p L\left(\frac{Du + D\widetilde{u}}{2}, x\right) \cdot \left(\frac{D\widetilde{u} - Du}{2}\right) \mathrm{d}x + \frac{\theta}{8}\int_\Omega |D\widetilde{u} - Du|^2 \,\mathrm{d}x \leq I[\widetilde{u}].$$

Adding these together yields,

$$I[v] + \frac{\theta}{8}\int_\Omega |Du - D\widetilde{u}|^2 \,\mathrm{d}x \leq \frac{1}{2}(I[u] + I[\widetilde{u}]).$$

This gives our claimed inequality. For equality, we need $Du = D\widetilde{u}$ almost everywhere and hence $u = \widetilde{u}$ almost everywhere since $u - \widetilde{u} \in W_0^{1,q}(\Omega)$. Since $u$ and $\widetilde{u}$ are minimizers, so is $v$ and hence $u = \widetilde{u}$ almost everywhere. □

*Weak Solutions.* Since our minimizer $u$ is only in $W^{1,q}(\Omega)$, it is not clear that it satisfies the Euler–Lagrange equations in any sense. In order to make things work, we will make some growth conditions on $L$. For some $C > 0$ and any $p, z, x$ we ask:

(4.5)

$$|L(p,z,x)| \le C(|p|^q + |z|^q + 1) \quad \text{and} \quad \begin{cases} |D_p L(p,z,x)| \le C(|p|^{q-1} + |z|^{q-1} + 1) \\ |L_z(p,z,x)| \le C(|p|^{q-1} + |z|^{q-1} + 1). \end{cases}$$

Suppose $u \in W^{1,q}(\Omega)$. These conditions imply $|D_p L(Du,u,x)| \in L^{q'}(\Omega)$ for $1 = q^{-1} + (q')^{-1}$ and also $|L_z(Du,u,x)| \le L^{q'}(\Omega)$.

**Definition 4.8**

We say $u \in \mathcal{A}$ is a *weak solution* of the Euler–Lagrange equation (4.2) with boundary condition $u|_{\partial\Omega} = g$ if for all $v \in W_0^{1,q}(\Omega)$,

$$\int_\Omega \sum_{i=1}^n L_{p_i}(Du,u,z) v_{x_i} + L_z(Du,u,x)\,\mathrm{d}x = 0.$$

**Theorem 4.9**

Suppose $L$ satisfies (4.5) and $u \in \mathcal{A}$ minimizes $I$. Then $u$ is a weak solution of (4.2).

Proof. We just need to follow through our derivation before taking more care with regularity. Consider $v \in W^{1,q}(\Omega)$ and let $i(\tau) = I[u+\tau v]$. Because of (4.5), this is always finite. For $\tau \neq 0$, we can consider the difference quotient,

$$\frac{1}{\tau}(i(\tau) - i(0)) = \tau^{-1} \int_\Omega L(Du+\tau Dv, u+\tau v, x) - L(Du,u,x)\,\mathrm{d}x = \int_\Omega L^\tau(x)\,\mathrm{d}x\,,$$

where,

$$L^\tau(x) = \tau^{-1}\Big(L(Du(x) + \tau Dv(x), u(x) + \tau v(x), x) - L(Du(x), u(x), x)\Big).$$

Now $L^\tau(x)$ will converge to $\sum_{i=1}^n L_{p_i}(Du,u,x)v_{x_i} + L_z(Du,u,x)v$ almost everywhere as $\tau \to 0$. We then have,

$$\begin{aligned} L^\tau(x) &= \frac{1}{\tau}\int_0^\tau \frac{\mathrm{d}}{\mathrm{d}s} L(Du + sDv, u + sv, x)\,\mathrm{d}s \\ &= \frac{1}{\tau}\int_0^\tau \sum_{i=1}^n L_{p_i}(Du + sDv, u + sv, x) v_{x_i} + L_z(Du + sDv, u + sv, x) v\,\mathrm{d}s\,. \end{aligned}$$

Since $c, v \in W^{1,q}(\Omega)$, Young's inequality and (4.5) implies,

$$|L^\tau(x)| \leq C(|Du|^q + |u|^q + |Dv|^q + |v|^q + 1) \in L^1(\Omega).$$

By the dominated convergence theorem we may take the limit of our difference quotient expression as $\tau \to 0$ to conclude,

$$i'(0) = \int_\Omega \sum_{i=1}^n L_{p_i}(Du, u, x) v_{x_i} + L(Du, u, x) v \, \mathrm{d}x \, .$$

We know $i$ has a minimum at $\tau = 0$. so $i'(0) = 0$ and $u$ is a weak solution of (4.2). □

Remark 4.10. Note that the Euler–Lagrange equations may have other solutions which are not minima of $I$. But if $(p, z) \mapsto L(p, z, x)$, then any solution will be a minimizer. △

We may easily generalize our above results to systems. If we again assume $L(P, z, x)$ is convex in $P$ and coercive, then a minimizer will exist. If $L$ is independent of $z$ and uniformly convex in $P$, a finite action minimizer will be unique. When $L$ satisfies the system analogue of (4.5), any minimizer will be a weak solution of the Euler–Lagrange equations (4.3), i.e. it satisfies,

$$\sum_{k=1}^m \int_\Omega \sum_{i=1}^n L_{p_i^k}(D\mathbf{u}, \mathbf{u}, x) w^k_{x_i} + L_{z^k}(D\mathbf{u}, \mathbf{u}, x) w^k \, \mathrm{d}x = 0.$$

*Local Minima.* Given some critical point of $I$, we need some conditions in order that it be a local (or global) minimum. Suppose $u$ is a smooth solution of the Euler Lagrange equation (4.2) with boundary condition $u|_{\partial\Omega} = g$; as we know this implies $u$ is a critical point of $I$, but it may be a saddle point or maximum rather than a minimum.

As before take $L(p, z, x)$ to be convex in $p$. Let $0 \in I$ be an open interval of $\mathbb{R}$ and $\{u(\cdot, \lambda) : \lambda \in I\}$ be a smooth one-parameter family of solutions of (4.2) so that $u(x) = u(x, 0)$ is a solution of our boundary value problem which we wish to analyze. For some smooth function $\theta : \Omega \to I$ satisfying $\theta = 0$ on $\partial\Omega$, we can define,

$$w(x) = u(x, \theta(x)).$$

Note that $w|_{\partial\Omega} = u|_{\partial\Omega} = g$.

**Theorem 4.11**

For $u, w$ as above, $I[u] \leq I[w]$. Hence if a solution $u$ of the Euler–Lagrange boundary value problem occurs within a one parameter family

$u_\lambda$ of critical points of $I$, then it is a minimizer amongst functions in $\mathcal{A}$ taking values within the range of all $u_\lambda$.

PROOF. By convexity in $p$ of $L$,

$$\begin{aligned} I[w] &= \int_\Omega L(Dw, w, x)\,\mathrm{d}x \\ &= \int_\Omega L(Du + u_\lambda D\theta, w, x)\,\mathrm{d}x \\ &\geq \int_\Omega L(Du, w, x) + u_\lambda D_p L(Du, w, x)\cdot D\theta\,\mathrm{d}x\,. \end{aligned}$$

Now set $\mathbf{B}$ to be the vector field with components,

$$b^i = \int_0^{\theta(x)} u_\lambda(x,\lambda) L_{p_i}(Du, u, x)\,\mathrm{d}\lambda\,.$$

We compute,

$$\begin{aligned} \operatorname{div}\mathbf{b} = &\sum_{i=1}^n \theta_{x_i} u_\lambda(x,\theta(x)) L_{p_i}(Du(x,\theta), u(x,\theta), x) \\ &+ \sum_{i=1}^n \int_0^{\theta(x)} u_{\lambda x_i}(x,\lambda) L_{p_i}(Du,u,x) + u_\lambda(x,\lambda)\partial_{x_i} L_{p_i}(Du,u,x)\,\mathrm{d}\lambda\,. \end{aligned}$$

By the Euler–Lagrange equation applied to the last term,

$$\begin{aligned} &= u_\lambda(x,\theta) D_p L(Du(x,\theta), u(x,\theta), x)\cdot D\theta \\ &\quad + \int_0^{\theta(x)} \sum_{i=1}^n u_{\lambda x_i}(x,\lambda) L_{p_i}(Du,u,x) + u_\lambda(x,\lambda) L_z(Du,u,x)\,\mathrm{d}\lambda\,. \end{aligned}$$

Note also that,

$$\partial_\lambda L(Du(x,\lambda), u(x,\lambda), x) = \sum_{i=1}^n L_{p_i}(Du,u,x) u_{x_i\lambda} + L_z(Du,u,x) u_\lambda.$$

Thus the above simplifies to,

$$\begin{aligned} \operatorname{div}\mathbf{b} &= u_\lambda(x,\theta) D_p L(Du,w,x)\cdot D\theta + \int_0^{\theta(x)} \partial_\lambda L(Du,u,x)\,\mathrm{d}\lambda \\ &= u_\lambda(x,\theta) D_p L(Du,w,x)\cdot D\theta + L(Du(x,\theta(x)),w,x) - L(Du,u,x). \end{aligned}$$

From the inequality obtained at the start of the proof,

$$\begin{aligned} I[w] &\geq \int_\Omega L(Du, w, x) + u_\lambda D_p L(Du, w, x) \cdot D\theta \, \mathrm{d}x \\ &= \int_\Omega L(Du, u, x) + \operatorname{div} \mathbf{b} \, \mathrm{d}x \\ &= I[u] + \int_{\partial\Omega} \mathbf{b} \cdot \nu \, \mathrm{d}S \,. \end{aligned}$$

Since $\theta = 0$ on $\partial\Omega$, $\mathbf{b}$ vanishes on the boundary too,

$$= I[u].$$

Thus the result is proved. □

**4.3. Regularity.** We now look to obtaining higher regularity on minimizers of the action functional. In general, this is challenging and so we suppose $I$ has the form,

$$I[w] = \int_\Omega L(Dw) - wf \, \mathrm{d}x \,,$$

for some $f \in L^2(\Omega)$. We also suppose our growth condition (4.5) holds for $q = 2$ so that we have,

$$|L(p)| \leq C(|p|^2 + 1) \quad \text{and} \quad |D_p L(p)| \leq C(|p| + 1). \tag{4.6}$$

The Euler–Lagrange equation simplifies to,

$$-\sum_{i=1}^n \partial_{x_i} L_{p_i}(Du) = f. \tag{4.7}$$

A weak solution $u$ will then be one which satisfies for all $v \in H_0^1(\Omega)$,

$$\int_\Omega \sum_{i=1}^n L_{p_i}(Du) v_{x_i} \, \mathrm{d}x = \int_\Omega fv \, \mathrm{d}x \,.$$

Our main goal will be to show a weak solution $u \in H^1(\Omega)$ is actually in $H^2_{\mathrm{loc}}(ZZ)$. To this end, we need further assumptions on $L$. We suppose the growth condition,

$$|D^2 L(p)| \leq C,$$

and assume uniform convexity,

$$\sum_{i,j}^n L_{p_i p_j}(p) \xi_i \xi_j \geq \theta |\xi|^2,$$

for some $\theta > 0$. This provides some non-linear analogue of uniform ellipticity, which we will try to leverage in a similar way to the regularity theory for 2nd order linear elliptic PDEs.

**Theorem 4.12**

Suppose $u \in H^1(\Omega)$ is a weak solution of (4.7) where $L$ satisfies the growth condition and uniform ellipticity as described above. Then $u \in H^2_{\text{loc}}(\Omega)$.

If $u \in H^1_0(\Omega)$ and $\Omega$ has $C^2$ boundary then $u \in H^2(\Omega)$ with,

$$\|u\|_{H^2(\Omega)} \leq C\|f\|_{L^2(\Omega)}.$$

PROOF. Pick open sets $V \Subset W \Subset \Omega$. Consider a smooth bump function $\zeta$ which is one on $V$ and vanishes off of $W$. Recall the difference quotient defined by,

$$D^h_k u(x) = \frac{u(x + he_k) - u(x)}{h}.$$

For $|h| > 0$ small and some $k \in \{1, \dots, n\}$, we set,

$$v = -D^{-h}_k(\zeta^2 D^h_k u).$$

Using this as a test function in the definition of weak solutions and integrating by parts with the difference quotient,

$$\sum_{i=1}^n \int_\Omega D^h_k(L_{p_i}(Du))\partial_{x_i}(\zeta^2 D^h_k u)\,\mathrm{d}x = -\int_\Omega f D^{-h}_k(\zeta^2 D^h_k u)\,\mathrm{d}x\,. \tag{4.8}$$

We compute,

$$\begin{aligned} D^h_k L_{p_i}(Du(x)) &= \frac{1}{h}(L_{p_i}(Du(x + he_k)) - L_{p_i}(Du(x))) \\ &= \frac{1}{h}\int_0^1 \frac{\mathrm{d}}{\mathrm{d}s} L_{p_i}(sDu(x + he_k) + (1 - s)Du(x))\,\mathrm{d}s \\ &= \frac{1}{h}\int_0^1 \sum_{j=1}^n L_{p_i p_j}(sDu(x + he_k) + (1 - s)Du(x)) \\ &\qquad \cdot (u_{x_j}(x + he_k) - u_{x_j}(x))\,\mathrm{d}s\,. \end{aligned}$$

We can rewrite this as,

$$= \sum_{j=1}^n a^{ij,h}(x) D^h_k u_{x_j}(x)$$

where,

$$a^{ij,h}(x) = \int_0^1 L_{p_i p_j}(sDu(x + he_k) + (1 - s)Du(x))\,\mathrm{d}s\,.$$

Plugging this into (4.8) yields,

$$\sum_{i,j=1}^{n} \int_\Omega \zeta^2 a^{ij,h} D_k^h u_{x_j} D_k^h u_{x_i} \, \mathrm{d}x + \sum_{i,j=1}^{n} \int_\Omega 2\zeta \zeta_{x_i} a^{ij,h} D_k^h u_{x_j} D_k^h u \, \mathrm{d}x \\ = - \int_\Omega f D_k^{-h} (\zeta^2 D_k^h u) \, \mathrm{d}x \, .$$

Let us call the terms on the left $A_1$ and $A_2$ and the term on the right $B$. By uniform convexity of $L$,

$$A_1 \geq \theta \int_\Omega \zeta^2 |D_k^h Du|^2 \, \mathrm{d}x \, .$$

Since we assumed $|D^2 L(p)| \leq C$,

$$|A_2| \leq C \int_W \zeta |D_k^h Du| |D_k^h u| \, \mathrm{d}x \, .$$

By Young's inequality,

$$\leq \varepsilon \int_W \zeta^2 |D_k^h Du|^2 \, \mathrm{d}x + \frac{C}{\varepsilon} \int_\Omega |D_k^h u|^2 \, \mathrm{d}x \, .$$

We can also find, as in the proof of elliptic regularity,

$$|B| \leq \varepsilon \int_\Omega \zeta^2 |D_k^h Du|^2 \, \mathrm{d}x + \frac{C}{\varepsilon} \int_\Omega f^2 + |Du|^2 \, \mathrm{d}x \, .$$

Now take $\varepsilon = \theta/4$. We obtain from the bounds on $A_1, A_2, B$,

$$\int_\Omega \zeta^2 |D_k^h Du|^2 \, \mathrm{d}x \leq C \int_W f^2 + |D_k^h u|^2 \, \mathrm{d}x \leq C \int_\Omega f^2 + |Du|^2 \, \mathrm{d}x \, .$$

Here we use the $L^p$ norm of the difference quotient is bounded by the $L^p$ norm of the derivative. Because $\zeta = 1$ on $V$,

$$\int_V |D_k^h Du|^2 \, \mathrm{d}x \leq C \int_\Omega f^2 + |Du|^2 \, \mathrm{d}x \, .$$

We conclude that $D^h Du \in L^2(V)$ and hence $Du \in H^1(V)$ and $u \in H^2(V)$. Thus $u \in H^2_{\mathrm{loc}}(\Omega)$. The fact that $u \in H^2(\Omega)$ if $u \in H^1_0(\Omega)$ and $\Omega$ has $C^2$ boundary just extends the boundary regularity theorem from elliptic PDEs. □

Ideally we would continue to derive higher order differentiability of $u$. In general this is not possible since $L$ is non-linear and attempting such a bound will become complicated as the derivatives of $L$ become more complicated. Here is a sketch of ideas that can be used for higher order regularity.

For simplicity we take $f = 0$. Let $w \in C_c^\infty(\Omega)$ and use $v = -w_{x_k}$ as a test function in the weak solution definition. As $u \in H^2_{\text{loc}}(\Omega)$ we obtain,

$$\int_\Omega \sum_{i,j=1}^n L_{p_i p_j}(Du) u_{x_k x_j} w_{x_i} \, \mathrm{d}x = 0.$$

We set $\widetilde{u} = u_{x_k}$ and $a^{ij} = L_{p_i p_j}(Du)$. Consider $V \Subset \Omega$. By smooth approximation, we have for all $w \in H_0^1(V)$,

$$\int_\Omega \sum_{i,j=1}^n a^{ij}(z) \widetilde{u}_{x_j} w_{x_i} \, \mathrm{d}x = 0.$$

Hence $\widetilde{u} \in H^1(V)$ is a weak solution of the elliptic PDE on $V$,

$$-\sum_{i,j=1}^n (a^{ij} \widetilde{u}_{x_j})_{x_1} = 0.$$

Since we only know that $a^{ij} \in L^\infty(V)$, we cannot use elliptic regularity to deduce regularity for $\widetilde{u}$.

Nevertheless, DeGiorgi, Moser, and Nash show any weak solution of this PDE is Hölder continuous for some exponent $\gamma > 0$. Hence $u \in C^{1,\gamma}_{\text{loc}}(\Omega)$. If $L$ is smooth, this then implies $a^{ij} \in C^{0,\gamma}_{\text{loc}}(\Omega)$ and so by Schauder estimates $u \in C^{2,\gamma}_{\text{loc}}(\Omega)$. We can then continue to bootstrap in this way applying Schauder estimates to show $u \in C^\infty(\Omega)$.

**4.4. Constraints and Lagrange Multipliers.** We now study variational problems subject to some constraints. The result will in certain cases be to introduce Lagrange multipliers into our Euler–Lagrange equations.

Consider the action functional,

$$I[w] = \frac{1}{2} \int_\Omega |Dw|^2 \, \mathrm{d}x \, .$$

Suppose we want to minimize $I$ over all functions $w$ which vanish on $\partial\Omega$ and satisfy the constraint,

$$J[w] := \int_\Omega G(w) \, \mathrm{d}z = 0,$$

for some smooth function $G : \mathbb{R} \to \mathbb{R}$. Let $G' = g$ and suppose a growth conidition,

$$|g(z)| \le C(|z| + 1) \quad \text{so} \quad |G(z)| \le C(|z|^2 + 1).$$

We now set $\mathcal{A}$ to be the set of functions $w \in H_0^1(\Omega)$ so that $J[w] = 0$. Suppose $\Omega$ is bounded, connected, and has smooth boundary.

**Theorem 4.13: Existence of Contrained Minima**

Suppose $\mathcal{A}$ is non-empty. Then there is $u \in \mathcal{A}$ which achieves the minimum of the functional $I$ over all functions in $\mathcal{A}$.

PROOF. Suppose $u_k$ is a sequence in $\mathcal{A}$ so that $I[u_k]$ converges to $m$, the infimum of $I$ over $\mathcal{A}$. After passing to a subsequence, $u_k$ converges weakly in $H_0^1(\Omega)$ to some $u \in H_0^1(\Omega)$, which by convexity satisfies $I[u] \leq m$. We just need $J[u] = 0$.

Since $H_0^1(\Omega)$ compactly embeds in $L^2(\Omega)$, we have that $u_k \to u$ in $L^2$. Thus,

$$\begin{aligned} |J(u)| &= |J(u) - J(u_k)| \\ &\leq \int_\Omega |G(u) - G(u_k)| \, \mathrm{d}x \, . \end{aligned}$$

By the growth condition,

$$\leq C \int_\Omega |u - u_k|(1 + |u| + |u_k|) \, \mathrm{d}x \, .$$

This converges to zero as $k \to \infty$. □

**Theorem 4.14: Lagrange Multipliers**

Suppose $u \in \mathcal{A}$ minimizes $I$ over $\mathcal{A}$. Then there is $\lambda \in \mathbb{R}$ so that,

$$\int_\Omega Du \cdot Dv \, \mathrm{d}x = \lambda \int_\Omega g(u)v \, \mathrm{d}x \quad \text{for all} \quad v \in H_0^1(\Omega).$$

Note this means $u$ is a weak solution of the non-linear eigenvalue problem,

$$\begin{cases} -\Delta u = \lambda g(u) \\ u|_{\partial\Omega} = 0, \end{cases}$$

for some $\lambda \in \mathbb{R}$. $\lambda$ is called a *Lagrange multiplier* corresponding to our integral constraint $J$.

PROOF. Fix $v \in H_0^1(\Omega)$. For now assume $g(u)$ is not zero almost everywhere. This implies we can find $w \in H_0^1(\Omega)$ so that,

$$\int_\Omega g(u)w \, \mathrm{d}x \neq 0.$$

Set,

$$j(\tau, \sigma) = J[u + \tau v + \sigma w] = \int_\Omega G(u + \tau v + \sigma w) \, \mathrm{d}x \, .$$

We have $j(0,0) = \int_\Omega G(u)\,\mathrm{d}x = 0$. $j$ is $C^1$ with derivatives,

$$\frac{\partial j}{\partial \tau} = \int_\Omega g(u+\tau v+\sigma w)v\,\mathrm{d}x$$
$$\frac{\partial j}{\partial \sigma} = \int_\Omega g(u+\tau v+\sigma w)w\,\mathrm{d}x\,.$$

By our choice of $w$, $\frac{\partial g}{\partial \sigma} \neq 0$.

Using the implicit function theorem, we may find a $C^1$ function $\varphi : \mathbb{R} \to \mathbb{R}$ such that $\varphi(0) = 0$ and $j(\tau, \varphi(\tau)) = 0$ for small enough $\tau$. Differentiating,

$$\frac{\partial j}{\partial \tau} + \frac{\partial j}{\partial \sigma}\varphi'(\tau) = 0.$$

From the computation of $j$'s derivatives, the above yields,

$$\varphi'(0) = -\frac{\int_\Omega g(u)v\,\mathrm{d}x}{\int_\Omega g(u)w\,\mathrm{d}x}.$$

We can define $w(\tau) = \tau v + \varphi(\tau)w$ and $i(\tau) = I[u+w(\tau)]$. Because $j(\tau, \varphi(\tau) = 0$, we know $J[u+w(\tau)] = 0$ and so $u_w(\tau) \in \mathcal{A}$. Thus $i$ is minimized at 0:

$$\begin{aligned} 0 = i'(0) &= \int_\Omega (Du + \tau Dv + \varphi(\tau)Dw)\cdot(Dv + \varphi'(\tau)Dw)\,\mathrm{d}x\,\big|_{\tau=0} \\ &= \int_\Omega Du\cdot(Dv+\varphi'(0)Dw)\,\mathrm{d}x\,. \end{aligned}$$

If we set,

$$\lambda = \frac{\int_\Omega Du\cdot Dw\,\mathrm{d}x}{\int_\Omega g(u)w\,\mathrm{d}x},$$

then the computation of $\varphi'(0)$ yields for all $v \in H_0^1(\Omega)$,

$$\int_\Omega Du\cdot Dv\,\mathrm{d}x = \lambda\int_\Omega g(u)v\,\mathrm{d}x\,.$$

This shows $u$ is a weak solution. In the other case where $g = 0$ almost everywhere, we may approximate $g$ by bounded functions to see $DG(u) = g(u)Du = 0$ almost everywhere. So $G$ is constant almost everywhere and thus is zero almost everywhere since $\int_\Omega G(u)\,\mathrm{d}x = 0$. Since $u|_{\partial\Omega} = 0$, $G(0) = 0$. Then we conclude $u = 0$ almost everywhere in order to be a minimizer. But then $u = 0$ is trivially a weak solution. □

Now we look at a different kind of constraint coming from a pointwise inequality. Consider an action functional,

$$I[w] = \int_\Omega \frac{1}{2}|Dw|^2 - fw\,\mathrm{d}x\,.$$

We can minimize over the set,

$$\mathcal{A} = \{w \in H_0^1(\Omega) : w \geq h \text{a.e. in } \mathbb{Z}\}.Z$$

Here $j : \mathbb{Z} \to \mathbb{R}$ is a smooth function called the *obstacle*.

**Theorem 4.15**

Suppose $\mathcal{A}$ is non-empty. Then there is a unique $u \in \mathcal{A}$ which minimizes $I$ over all functions in $\mathcal{A}$.

Proof. The existence of a minimizer follows as before. We find a minimizing sequence $u_k$ converging to $u$ weakly in $H_0^1(\Omega)$ and thus strongly in $L^2(\Omega)$. As $u_k \geq h$, the same holds for $u$ in the limit and so $u \in \mathcal{A}$ is our minimizer.

To see uniqueness, suppose $u, \widetilde{u}$ are distinct minimizers in $\mathcal{A}$ and let $w \in \mathcal{A}$ be their average. We compute,

$$\begin{aligned}
I[w] &= \int_\Omega \frac{1}{2}\left|\frac{1}{2}(Du + D\widetilde{u})\right|^2 - f\left(\frac{u+\widetilde{u}}{2}\right) \mathrm{d}x \\
&= \int_\Omega \frac{1}{8}(|Du|^2 + 2Du \cdot D\widetilde{u} + |D\widetilde{u}|^2) - f\left(\frac{u+\widetilde{u}}{2}\right) \mathrm{d}x \\
&= \int_\Omega \frac{1}{8}(2|Du|^2 + 2|D\widetilde{u}|^2 - |Du - D\widetilde{u}|^2) - f\left(\frac{u+\widetilde{u}}{2}\right) \mathrm{d}x \\
&< \frac{1}{2}\int_\Omega \frac{1}{2}|Du|^2 - fu \,\mathrm{d}x + \frac{1}{2}\int_\mathbb{Z} \frac{1}{2}|D\widetilde{u}|^2 - f\widetilde{u} \,\mathrm{d}x \\
&= \frac{1}{2}I[u] + \frac{1}{2}I[\widetilde{u}].
\end{aligned}$$

We thus obtain a contradiction. □

With this choice of constraint, the Euler–Lagrange equations become an inequality.

**Theorem 4.16**

Suppose $u \in \mathcal{A}$ is the unique minimizer of $I$. Then $u$ satisfies a *variational inequality*,

$$\int_\Omega Du \cdot D(w-u) \,\mathrm{d}x \geq \int_\Omega f(w-u) \,\mathrm{d}x \quad \text{for all} \quad w \in \mathcal{A}.$$

PROOF. Fix $w \in \mathcal{A}$. For $\tau \in [0,1]$, $u + \tau(w-u) \in \mathcal{A}$ since $\mathcal{A}$ is convex. Now set,

$$i(\tau) = I[u + \tau(w-u)].$$

Note $i(0)$ is a minimum of $i$ for $\tau \in [0,1]$. Thus $i'(0) \geq 0$. We compute,

$$\begin{aligned}\frac{1}{\tau}(i(\tau) - i(0)) &= \frac{1}{\tau}\int_\Omega \frac{1}{2}(|u + \tau D(w-u)|^2 - |Du|^2) - f(u + \tau(w-u) - u)\,\mathrm{d}x \\ &= \int_\Omega Du \cdot D(w-u) + \frac{\tau}{2}|D(w-u)|^2 - f(w-u)\,\mathrm{d}x\,.\end{aligned}$$

Hence,

$$0 \geq i'(0) = \int_\Omega Du \cdot D(w-u) - f(w-u)\,\mathrm{d}x\,.$$

As desired. □

Now we consider pointwise equality constraints. As an example we consider harmonic maps into spheres. That is, we study the functional,

$$I[\mathbf{w}] = \frac{1}{2}\int_\Omega |D\mathbf{w}|^2\,\mathrm{d}x\,,$$

and its minima on the set,

$$\mathcal{A} = \{\mathbf{w} \int H^1(\Omega;\mathbb{R}^m) : \mathbf{w} = \mathbf{g} \text{ on } \partial\Omega,\ |\mathbf{w}| = 1 \text{ a.e.}\}.$$

The existence of a minimizer is clear from our existing theory.

**Theorem 4.17**

Suppose $\mathbf{u} \in \mathcal{A}$ is a minimizer of $I$. Then $\mathbf{u}$ satisfies,

$$\int_\Omega D\mathbf{u} : D\mathbf{v}\,\mathrm{d}x = \int_\Omega |D\mathbf{u}|^2\mathbf{u}\cdot\mathbf{v}\,\mathrm{d}x\,,$$

for all $\mathbf{v} \in H^1_0(\Omega;\mathbb{R}^m) \cap L^\infty(\Omega;\mathbb{R}^m)$. Here : denotes the double dot product of matrices; $A : B$ is $A_{11}B_{11} + A_{12}B_{12} + \cdots + A_{nm}B_{nm}$.

This says that $\mathbf{u}$ is a weak solution of,

$$\begin{cases} -\Delta\mathbf{u} = |D\mathbf{u}|^2\mathbf{u} & \text{in } \Omega \\ \mathbf{u} = \mathbf{g} & \text{on } \partial\Omega. \end{cases}$$

We interpret $\lambda = |D\mathbf{u}|^2$ as a Lagrange multiplier. Because our constraint is pointwise instead of integral, the multiplier is a function rather than a number.

PROOF. Fix $\mathbf{v} \in H^1_0(\Omega;\mathbb{R}^m) \cap L^\infty(\Omega;\mathbb{R}^m)$. For small enough $\tau$, $|\mathbf{u}+\tau\mathbf{v}|$ is non-vanishing almost everywhere. Thus,

$$\tau = \frac{\mathbf{u}+\tau\mathbf{v}}{|\mathbf{u}+\tau\mathbf{v}|} \in \mathcal{A}.$$

Setting $i(\tau) = I[\mathbf{v}(\tau)]$, this will have a minimum at 0.

$$0 = i'(0) = \int_\Omega D\mathbf{u} : D\dot{\mathbf{v}}(0)\,\mathrm{d}x\,,$$

where we set,

$$\dot{\mathbf{v}}(\tau) = \frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{v}(\tau) = \frac{\mathbf{v}}{|\mathbf{u}+\tau\mathbf{v}|} - \frac{[(\mathbf{u}+\tau\mathbf{v})\cdot\mathbf{v}](\mathbf{u}+\tau\mathbf{v})}{|\mathbf{u}+\tau\mathbf{v}|^3}.$$

Thus $\dot{\mathbf{v}}(0) = \mathbf{v} - (\mathbf{u}\cdot\mathbf{v})\mathbf{u}$. Plugging this into the above,

$$0 = \int_\Omega D\mathbf{u} : D\mathbf{v} - D\mathbf{u} : D((\mathbf{u}\cdot\mathbf{v})\mathbf{u})\,\mathrm{d}x\,.$$

Note $|\mathbf{u}|^2 = 1$ implies $(D\mathbf{u})^\intercal\mathbf{u} = 0$. This implies almost everywhere,

$$D\mathbf{u} : D((\mathbf{u}\cdot\mathbf{v})\mathbf{u}) = |D\mathbf{u}|^2\mathbf{u}\cdot\mathbf{v},$$

which gives out result. □

We explain one final kind of constraint coming from fluid mechanics. Let $\Omega \subset \mathbb{R}^3$ be bounded and simply connected. Consider the action functional,

$$I[\mathbf{w}] = \int_\Omega \frac{1}{2}|D\mathbf{w}|^2 - \mathbf{f}\cdots\mathbf{w}\,\mathrm{d}x\,,$$

for given $\mathbf{f} \in L^2(\Omega;\mathbb{R}^3)$. We attempt to minimize this functional over the domain of divergenceless vector functions,

$$\mathcal{A} = \{\mathbf{w} \in H^1_0(\Omega;\mathbb{R}^3) : \operatorname{div}\mathbf{w} = 0 \text{ in } \Omega\}.$$

As usual, we obtain a unique minimizer of this constrained variational problem. We may think of $\mathbf{u}$ as a velocity field of fluid flow in $\Omega$ subject to a force $\mathbf{f}$. The constraint $\operatorname{div}\mathbf{u}$ expresses that the fluid is incompressible. We state without proof the Euler–Lagrange equation corresponding to such a constraint.

PROPOSITION 4.18. *Suppose* $\mathbf{u}$ *is a minimizer within* $\mathcal{A}$ *of the aciton functional* $I$. *Then there is a function* $p \in L^2_{\mathrm{loc}}(\Omega)$ *so that for every* $\mathbf{v} \in H^1(\Omega;\mathbb{R}^3)$ *with compact support in* $\Omega$ *one has,*

$$\int_\Omega D\mathbf{u} : D\mathbf{v}\,\mathrm{d}x = \int_\Omega p\operatorname{div}\mathbf{v} + \mathbf{f}\cdot\mathbf{v}\,\mathrm{d}x\,.$$

PROOF. See [**Evans**, §8.4.4]. □

This proposition says that $\mathbf{u}$ is a weak solution of *Stokes' problem*,

$$\begin{cases} -\Delta \mathbf{u} = \mathbf{f} - Dp & \text{in } \Omega \\ \operatorname{div} \mathbf{u} = 0 & \text{in } \Omega \\ \quad \mathbf{u} = 0 & \text{on } \partial\Omega \end{cases}$$

The function $p$ is the Lagrange multiplier corresponding to the incompressibility condition. Physically, $p$ corresponds to pressure.

# Bibliography


[Ad] R. Adams and J. Fournier, *Sobolev Spaces 2 ed.*, Academic Press, Pure and Applied Mathematics Series, 2003.

[Adh] M. R. Adhikari, *Basic Algebraic Topology and its Applications*, Springer, 2016.

[BGV] N. Berline, E. Getzler, M. Verge, *Heat Kernels and Dirac Operators*, Springer Grundlehren Text Editions, 2004.

[Br] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer Universitext, 2011.

[Brown] E. H. Brown Jr., "The Cohomology of $BSO_n$ and $BO_n$ with Integer Coefficients." Proceedings of the American Mathematical Society **85**, no. 2 (1982): 283-–88.

[CE] L. Cieliebak and Y. Eliashberg, *From Stein to Weinstein and Back: Symplectic Geometry of Affine Complex Manifolds*, American Mathematical Society, Colloquium Publications, Vol. 59, 2012.

[CF] P. Conner and E. Floyd, *Differentiable Periodic Maps*, Springer–Verlag, 1964.

[daS] A. C. da Silva, *Lectures on Symplectic Geometry*, Springer, Lecture Notes in Mathematics, 2008.

[DHJV] S. Dwivedi, J. Herman, L. Jeffrey, and T. van den Hurk, *Hamiltonian Groups Actions and Equivariant Cohomology*, Springer, Briefs in Mathematics, 2019.

[Evans] L. Evans, *Partial Differential Equations 2 ed.*, American Mathematical Society, Graduate Studies in Mathematics, 2010.

[FF] A. Fomenko and D. Fuchs, *Homotopical Topology 2 ed.*, Springer, Graduate Texts in Mathematics, 2016.

[Hatcher] A. Hatcher, *Algebraic Topology*, Cambridge University Press, 2002.

[HatcherSS] A. Hatcher, *Spectral Sequences*, unpublished fifth chapter of [**Hatcher**], retrieved from https://pi.math.cornell.edu/ hatcher/AT/SSpage.html, 2004.

[HS] G. Hochschild and J–P. Serre, "Cohomology of Lie Algebras," Annals of Mathematics **57**, no. 3 (1953): 591–603.

[Hutchings] M. Hutchings, *Lecture notes on Morse homology (with an eye towards Floer theory and pseudoholomorphic curves)*, lecture notes retrieved from `https://math.berkeley.edu/~hutching/`, 2002.

[Huy] D. Huybrechts, *Complex Geomtry: An Introduction*, Springer Universitext, 2005.

[Jost] J. Jost, *Partial Differential Equations*, Springer, Graduate Texts in Mathematics, 2002.

[Koch] S. Kochman, *Stable Homotopy Groups of Spheres*, Springer, Lecture Notes in Mathematics, 1990.

[May] P. May, *A Concise Course in Algebraic Topology*, University of Chicago Press, Chicago Lectures in Mathematics, 1999.

[McS1] D. McDuff and D. Salamon, *Introduction to Symplectic Topology 3 ed.*, Oxford University Press, Oxford Graduate Texts in Mathematics, 2016.

[McS2] D. McDuff and D. Salamon, *J-holomorphic Curves and Symplectic Topology: 2 ed.*, American Mathematical Society, Colloquium Publications Vol. 52, 2012.

[Mein] E. Meinrenken, *Symplectic Geometry*, lecture notes retrieved from https://www.math.toronto.edu/mein/teaching/LectureNotes/sympl.pdf, 2000.

[MS] J. Milnor and J. Stasheff, *Characteristic Classes*, Princeton University Press, Annals of Mathematics Studies, 1974.

[Rud] Y. Rudyak, *On Thom Spectra, Orientability, and Cobordism*, Springer, Monographs in Mathematics, 2008.

[Str] N. Strickland, *Bott Periodicity and Hopf Rings*, University of Manchester PhD thesis, retrieved from https://strickland1.org/research/thesis.pdf, 1992.

[Thom] R. Thom, "Quelques propriétés globales des variéés différentiables," Commentarii Mathematici Helvetici **28**, (1954): 17–86.