MATH 8230, SPRING 2010 LECTURE NOTES

MIKE USHER

Contents

1. Introductory remarks

Recall that a symplectic manifold is a pair (M, ω) consisting of a smooth manifold M and a 2-form $\omega \in \Omega^2(M)$ which is closed ($d\omega = 0$) and *non-degenerate* (*i.e.*, whenever $v \in TM$ is a nonzero tangent vector there is some other tangent vector w such that $\omega(v, w) \neq 0$). The standard example is $M = \mathbb{R}^{2n} = \{(x_1, \ldots, x_n, y_1, \ldots, y_n)\}$ equipped with the symplectic form $\omega_0 = \sum_{i=1}^n dx^i \wedge dy^i$. A basic result asserts that, *locally*, all symplectic manifolds are equivalent to this one:

Theorem 1.1. *(Darboux's Theorem) If (M, ω) is a symplectic manifold and $p \in M$, there is a neighborhood U of p and a diffeomorphism $\phi\colon U \to V \subset \mathbb{R}^{2n}$ onto an open subset $V \subset \mathbb{R}^{2n}$ such that $\phi^*\omega_0 = \omega$.*

Thus in order to construct invariants of symplectic manifolds (or, more broadly, to study properties that may hold in some symplectic manifolds of a given dimension but not in all of them), one has to go beyond local considerations. *Pseudoholomorphic curves* have emerged as powerful tools in extracting such global information.

Recall that an *almost complex structure* J on a smooth manifold M is a map $J: TM \to TM$ whose restriction to each tangent space T_pM is a linear map whose square is $-I$ (where I is the identity). We make the following definitions:

Definition 1.2. *Let (M, ω) be a symplectic manifold, and let $J: TM \to TM$ be an almost complex structure.*

- *J is called* ω-tame *if, for all nonzero $v \in TM$, we have*
$$\omega(v, Jv) > 0.$$
- *J is called* ω-compatible *if J is ω-tame and moreover, for all $p \in M$ and $v, w \in T_pM$ we have*
$$\omega(Jv, Jw) = \omega(v, w).$$

Let
$$\mathcal{J}(M, \omega) = \{J: TM \to TM | J \text{ is } \omega\text{-compatible}\}$$
and
$$\mathcal{J}_\tau(M, \omega) = \{J: TM \to TM | J \text{ is } \omega\text{-tame}\}.$$
We showed last semester that $\mathcal{J}(M, \omega)$ is always contractible (in particular it's nonempty), and it's also true that $\mathcal{J}_\tau(M\omega)$ is contractible (the proof is marginally harder, and can be found in [MS1, Chapter 4]).

If $J \in \mathcal{J}_\tau(M, \omega)$ and $v, w \in T_pM$ for some $p \in M$, define
$$g_J(v, w) = \frac{1}{2}\left(\omega(v, Jw) + \omega(w, Jv)\right).$$
A routine exercise shows that g_J is an inner product on T_pM, and so g_J defines a Riemannian metric on M. In particular we have $g_J(v, v) = \omega(v, Jv)$, which is positive for nonzero v by the definition of tameness. If J is ω-compatible and not just ω-tame then the above formula simplifies to $g_J(v, w) = \omega(v, Jw)$.

Recall that a *complex manifold* is by definition a smooth manifold M which can be covered by coordinate charts $\phi_\alpha: U_\alpha \to \phi_\alpha(U_\alpha) \subset \mathbb{C}^n$ with the property that the transition functions $\phi_\beta \circ \phi_\alpha^{-1}: \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$ are *holomorphic*. In other words, where $i: T\mathbb{C}^n \to T\mathbb{C}^n$ is the endomorphism which multiplies tangent vectors by $\sqrt{-1}$, we require that the linearizations of the coordinate charts obey
$$\phi_{\beta*} \circ \phi_{\alpha*}^{-1} \circ i = i \circ \phi_{\beta*} \circ \phi_{\alpha*}^{-1},$$
or equivalently
$$\phi_{\alpha*}^{-1} \circ i \circ \phi_{\alpha*} = \phi_{\beta*}^{-1} \circ i \circ \phi_{\beta*}. \tag{1}$$

If M is a complex manifold it has a natural almost complex structure $J: TM \to TM$ defined by requiring that in every chart $\phi_\alpha: U_\alpha \to \mathbb{C}^n$ we should have $\phi_{\alpha*} \circ J = i \circ \phi_{\alpha*}$. Thus $J = \phi_{\alpha*}^{-1} \circ i \circ \phi_{\alpha*}$, so (1) makes clear that this definition is consistent over all the charts and so indeed does define a map on all of TM, which is smooth and obeys $J^2 = -I$.

Thus we have a fairly broad class of examples of almost complex structures, arising from complex manifolds. Almost complex structures that arise in the way described in the above paragraph are called *integrable*. If $\omega \in \Omega^2(M)$ is a symplectic form with which some integrable almost complex structure is compatible then ω is called a *Kähler* form. Last semester we showed that there are symplectic manifolds which are not Kähler (the odd-dimensional Betti numbers of a Kähler manifold are even, and one can construct symplectic manifolds with b_1 odd)—from this it should be rather obvious that not all almost complex structures are integrable, since a non-Kähler symplectic manifold still has compatible almost complex structures. The fact that some almost complex structures aren't integrable can be seen more directly: there is a particular tensor called the Nijenhuis tensor N_J which can be constructed from an almost complex structure J (see [MS1, Chapter 4], or any number of other references), and it's not hard to show that for J to be integrable this tensor has to vanish. N_J can be written out in local coordinates, and if $\dim M \geq 4$ it's routine to construct almost complex structures (just in a local coordinate chart) for which N_J is nonzero, so such J must be non-integrable. A much harder theorem, due to Newlander and Nirenberg, shows that if $N_J = 0$ then J is integrable. Now N_J is always zero when $\dim M = 2$, so it follows then that every almost complex structure on a 2-manifold *is* integrable. There's a proof of the two-dimensional case of the Newlander-Nirenberg theorem in [MS2, Appendix E].

Definition 1.3. *If (M, J) is an almost complex manifold, a (parametrized) J-holomorphic curve in M is a map $u\colon \Sigma \to M$ where (Σ, j) is an almost complex 2-manifold such that*

$$u_* \circ j = J \circ u_*.$$

Convention has it that a pseudoholomorphic curve in a symplectic manifold (M, ω) is a map from a surface to M which is J-holomorphic for some $J \in \mathcal{J}_\tau(M, \omega)$.

As partial justification for the terminology, note that since the almost complex structure j on Σ is integrable Σ admits the structure of a complex manifold of dimension 1, *i.e.*, a complex curve.

If $p \in \Sigma$, we can work in a local chart around p, so we identify a neighborhood of p with an open set in $\mathbb{C}$ parametrized by the complex coordinate $z = s + it$. Thus $j\partial_s = \partial_t$ and $j\partial_t = -\partial_s$.

The equation $u_* \circ j = J \circ u_*$ (the *Cauchy-Riemann equation* associated to j and J) can be rearranged as

$$u_* + J \circ u_* \circ j = 0.$$

Applying this to ∂_s gives

$$\frac{\partial u}{\partial s} + J(u(s,t))\frac{\partial u}{\partial t} = 0.$$

(Applying it to ∂_t instead gives an equivalent equation and hence no new information). If J were integrable then it would be possible to choose (local) coordinates on M in terms of which J is just the constant endomorphism i, but if J is not integrable then this is not the case, and so the Cauchy-Riemann equation is nonlinear, even if we're only interested in its behavior in local coordinate charts.

2. Energy and area

Consider a compact almost complex 2-manifold (Σ, j) (possibly with boundary), a Riemannian manifold (M, g), and a map $u\colon \Sigma \to M$ (say u is continuously differentiable, though slightly less would suffice). From these data we will presently define the (Dirichlet) energy $E(u)$, and then we will make some observations about the energies of J-holomorphic curves (with respect to the metric $g = g_J$).

First of all, note that the almost complex structure j on Σ determines a *conformal class* $\mathcal{H}_j$ of Riemannian metrics h on Σ: namely, $h \in \mathcal{H}_j$ if and only if $h(v, jv) = 0$ and $h(jv, jv) = h(v, v)$ whenever $v \in T\Sigma$. Because the conditions on h for membership in $\mathcal{H}_j$ are preserved under convex combinations, it's easy to see that $\mathcal{H}_j \neq \emptyset$: take an atlas of coordinate charts $\{U_\alpha\}$ and a subordinate partition of unity $\{\phi_\alpha\}$, let h_α be metrics in U_α which satisfy the appropriate conditions, and then $h = \sum \phi_\alpha h_\alpha \in \mathcal{H}_j$. If $h_1, h_2 \in \mathcal{H}_j$ then it's not hard to see that h_1 and h_2 are *conformally equivalent*, which is to say that there is a function $f\colon \Sigma \to (0, \infty)$ such that

$$h_2(v, w) = f(p)h_1(v, w) \text{ for } v, w \in T_p\Sigma.$$

Now choose an arbitrary $h \in \mathcal{H}_j$. Like any metric, h determines a volume form vol_h which, in local coordinates, is given by $vol_h = e^1 \wedge e^2$ where $\{e_1, e_2\}$ is a local oriented h-orthonormal frame for $T\Sigma$ and e^i are locally defined 1-forms obeying $e^i(e_j) = \delta_{ij}$. The fact that $h \in \mathcal{H}_j$ means that we will have $e_2 = je_1$. If $h' = fh$ is conformally equivalent to h then one gets an h'-orthonormal basis by multiplying an h-orthonormal basis by $f^{-1/2}$, so the corresponding e^1 and e^2 covectors get multiplied by $f^{1/2}$ and so $vol_{fh} = f \cdot vol_h$.

Given a map $u\colon \Sigma \to M$ (and also the metric g on M) choose $h \in \mathcal{H}_j$ and, for $p \in \Sigma$ and a nonzero $v \in T_p\Sigma$ consider the quantity

$$|du(p)|^2 = \frac{g(u_*v, u_*v) + g(u_*jv, u_*jv)}{h(v, v)}.$$

Now a general nonzero vector in $T_p\Sigma$ has form $w = av + bjv$ (so $jw = -bv + ajv$) where a and b aren't both zero, and it's easy to see that replacing v by w leaves the above formula for $|du(p)|^2$ unchanged. Thus $|du(p)|^2$ depends just on the point p and the metrics h and g.

Now our assumption wasn't that h was given, but rather that j was. If we replace h by another metric $h' \in \mathcal{H}_j$, so $h' = fh$ for some function f, then obviously $|du(p)|^2$ gets multiplied by f^{-1}. On the other hand the associated volume form, as mentioned above, gets multiplied by f when we replace h by h'. Hence the *energy density 2-form*

$$|du(\cdot)|^2 vol_h$$

is independent of the representative h of $\mathcal{H}_j$. The Dirichlet energy is then half the integral of this energy density 2-form over Σ:

$$E(u) = \frac{1}{2}\int_\Sigma |du(\cdot)|^2 vol_h.$$

As advertised, $E(u)$ depends only on $u\colon \Sigma \to M$, on the almost complex structure j on Σ, and on the Riemannian metric g on M.

We now specialize to the case that (M,ω) is a symplectic manifold with $J \in \mathcal{J}_\tau(M,\omega)$ an ω-tame almost complex structure, inducing a metric g_J on M. Recall in particular that $g_J(v,v) = \omega(v, Jv)$. The metric g_J will be the one used in the definitions of all energies below.

Proposition 2.1. *Let $u\colon \Sigma \to M$ be a J-holomorphic curve, where (Σ, j) is a compact almost complex manifold. Then*

$$E(u) = \int_\Sigma u^*\omega.$$

Proof. $E(u)$ is half the integral of the energy density 2-form $|du(\cdot)|^2 vol_h$, so we need to show that this 2-form is equal to $2u^*\omega$.

So choose $p \in \Sigma$ and let $z = s + it$ be a local complex coordinate around p; in particular $j\partial_s = \partial_t$. There is then a metric $h \in \mathcal{H}_j$ with $h(\partial_s, \partial_s) = 1$ at p, so on $T_p\Sigma$ the energy density 2-form is

$$|du(p)|^2 vol_h = (g_J(u_*\partial_s, u_*\partial_s) + g_J(u_*\partial_t, u_*\partial_t))ds \wedge dt.$$

But the Cauchy-Riemann equation says that $Ju_*\partial_s = u_*\partial_t$ and $Ju_*\partial_t = -u_*\partial_s$, so by the definition of g_J we get (at p)

$$|du(p)|^2 vol_h = 2\omega(u_*\partial_s, u_*\partial_t)ds \wedge dt = 2u^*\omega,$$

as desired. □

In particular, this shows that if Σ is closed (i.e. compact without boundary) the Dirichlet energy of a J-holomorphic map is a *topological quantity*:

Corollary 2.2. *Let (Σ, j) be a closed almost complex 2-manifold and let $u\colon \Sigma \to M$ be a J-holomorphic map to a symplectic manifold (M,ω) with $J \in \mathcal{J}_\tau(M,\omega)$. Writing $[\omega] \in H^2(M,\mathbb{R})$ for the de Rham cohomology class of ω, if $u_*[\Sigma] = A \in H_2(M;\mathbb{Z})$ then we have*

$$\langle [\omega], A\rangle = E(u).$$

In particular $\langle [\omega], A\rangle \geq 0$, with equality only if u is constant.

In particular it follows that no nonconstant pseudoholomorphic curves exist within a coordinate chart—rather they are a more global phenomenon as was suggested earlier.

We will also have occasion to consider the case where Σ is compact with boundary, and $u\colon \Sigma \to M$ is J-holomorphic, with boundary mapping to a *Lagrangian submanifold* $L \subset M$ (i.e., $\dim L = \frac{1}{2}\dim M$ and $\omega|_L = 0$). In this case it again holds that the energy of the map is a topological quantity:

Corollary 2.3. *Let (M,ω) be a symplectic manifold and let $L \subset M$ be a Lagrangian submanifold. There is a homomorphism $I_\omega\colon H_2(M,L;\mathbb{Z}) \to \mathbb{R}$ such that whenever $J \in \mathcal{J}_\tau(M,\omega)$ and $u\colon \Sigma \to M$ is a J-holomorphic map where Σ is a compact 2-manifold with boundary $\partial\Sigma$ such that $U(\partial\Sigma) \subset L$, we have*

$$E(u) = I_\omega(A),$$

where $A \in H_2(M,L;\mathbb{Z})$ is the relative homology class represented by u.

Proof. If $u\colon (\Sigma, \partial\Sigma) \to (M, L)$ is such a map we have $E(u) = \int_\Sigma u^*\omega$; what needs to be shown is that this quantity depends only on the relative homology class of u. To see this, suppose that $u_i\colon (\Sigma_i, \partial\Sigma_i) \to (M, L)$ are two (smooth, but not necessarily pseudoholomorphic) maps representing the same relative homology class $A \in H_2(M, L; \mathbb{Z})$; we'll show that $\int_{\Sigma_1} u_1^*\omega = \int_{\Sigma_2} u_2^*\omega$. Denote by P_i the relative chains represented by u_i; the fact that the P_i are homologous amounts to the statement that there is a 3-dimensional chain $c \in C_3(M)$ such that $\partial c = P_1 - P_2$ modulo $C_2(L)$. Thus for some $v\colon S \to L$ we have (using Stokes' theorem)

$$0 = \int_c d\omega = \int_{\partial c} \omega = \int_{\Sigma_1} u_1^*\omega - \int_{\Sigma_2} u_2^*\omega + \int_S v^*\omega.$$

But the last term is zero since v has image contained in L and since $\omega|_L = 0$. This proves that $\int_{\Sigma_1} u_1^*\omega = \int_{\Sigma_2} u_2^*\omega$, and so $I_\omega(A)$ may be set equal to the common value. This defines I_ω, and it's straightforward to see that it is a homomorphism. □

All of the above is valid as long as J is ω-tame. If additionally J is ω-compatible one gets additional conclusions about the energy and, relatedly, the area. So let us fix a symplectic manifold (M, ω) and $J \in \mathcal{J}(M, \omega)$ and consider a C^1 (and not necessarily J-holomorphic) map $u\colon \Sigma \to M$ where Σ is a compact surface (perhaps with boundary).

The *area* of u (as measured by the metric g_J) is defined as follows. If $m \in M$ and $v, w \in T_mM$ write $|v \wedge w|_{g_J} = \sqrt{g_J(v, v)g_J(w, w) - g_J(v, w)^2}$ (thus $|v \wedge w|_{g_J}$ is the area of the parallelogram spanned by v and w). At each $p \in \Sigma$ choose an oriented basis $\{e_1, e_2\}$ for $T_p\Sigma$, with dual basis $\{e^1, e^2\}$ for $T_p^*\Sigma$. Then it's easy to see that the element $\alpha(p) \in \Lambda^2 T_p^*\Sigma$ defined by

$$\alpha(p) = |(u_*e_1) \wedge (u_*e_2)|_{g_J} e^1 \wedge e^2$$

is independent of the chosen basis. Thus we get an area form $\alpha \in \Omega^2(\Sigma)$, and we define

$$Area(u) = \int_\Sigma \alpha.$$

Of course this definition can be made given any metric on M in place of g_J and depends on this metric; meanwhile, unlike the Dirichlet energy, it doesn't depend on a metric (or even on a conformal class of metrics) on Σ.

Proposition 2.4. *If J is ω-compatible, a C^1 map $u\colon \Sigma \to M$ obeys*

$$Area(u) \geq \int_\Sigma u^*\omega, \tag{2}$$

with equality if u is J-holomorphic (for some almost complex structure j on Σ).

Proof. It's enough to show that, if $p \in \Sigma$, $\omega(u_*e_1, u_*e_2) \leq |(u_*e_1) \wedge (u_*e_2)|_{g_J}$ for some (and hence any) oriented basis for $T_p\Sigma$, with equality if u is J-holomorphic. By replacing e_2 with $e_2 - ae_1$ for a suitable $a \in \mathbb{R}$ if necessary we may assume that $g_J(u_*e_1, u_*e_2) = 0$. Then $|(u_*e_1) \wedge (u_*e_2)|_{g_J} = (g_J(u_*e_1, u_*e_1)g_J(u_*e_2, u_*e_2))^{1/2}$, while (since by compatibility for $v, w \in T_{u(p)}M$ we have $g_J(Jv, w) = \omega(Jv, Jw) = \omega(v, w)$)

$$\omega(u_*e_1, u_*e_2) = g_J(Ju_*e_1, u_*e_2) \leq (g_J(Ju_*e_1, Ju_*e_1)g_J(u_*e_2, u_*e_2))^{1/2} = |(u_*e_1) \wedge (u_*e_2)|_{g_J},$$

where the inequality is the Cauchy–Schwarz inequality for g_J and the final equality uses that $g_J(Jv, Jw) = g_J(v, w)$ (as follows directly from the definition of g_J). This proves (2) (by integrating over Σ), and shows that equality holds (2) exactly when equality holds in the Cauchy–Schwarz inequality

$$g_J(Ju_*e_1, u_*e_2) \leq (g_J(Ju_*e_1, Ju_*e_1)g_J(u_*e_2, u_*e_2))^{1/2}$$

at each point $p \in \Sigma$ (where $e_1, e_2 \in T_p\Sigma$ are chosen so that $g_J(u_*e_1, u_*e_2) = 0$). If u is J-holomorphic (with respect to the almost complex structure j on Σ), then if $e_1, e_2 \in T_p\Sigma$ with $je_1 = e_2$ we will have $Ju_*e_1 = u_*e_2$, and hence $g_J(u_*e_1, u_*e_2) = \omega(u_*e_1, J^2u_*e_1) = 0$ (so e_1, e_2 can be used in the above computation), and since

$Ju_*e_1 = u_*e_2$ equality does indeed hold in the above Cauchy–Schwarz inequality. Thus $Area(u) = \int_\Sigma u^*\omega$ when u is J-holomorphic. ◻

It follows from this that, when J is ω-compatible, J-holomorphic curves are *minimal surfaces*, *i.e.*, that if $u_t \colon \Sigma \to M$ is a smooth family of maps with $u_0 = u$ such that u_t coincides with u outside a compact subset of the interior of Σ, then $t \mapsto Area(u_t)$ has a critical point (indeed, a minimum, namely $\int_\Sigma u^*\omega$) at $t = 0$. Minimal surfaces have a rich theory going back as far as Lagrange (one reference is [Law]), and this allows results from that theory to be brought to bear on pseudoholomorphic curves.

Remark 2.5. If J is just ω-tame rather than ω-compatible then it's still true that a J-holomorphic map $u \colon \Sigma \to M$ will have $Area(u) = \int_\Sigma u^*\omega$. Indeed if $e_1 \in T_pM$ is nonzero (so that $\{e_1, je_1\}$ is an oriented basis for $T_p\Sigma$), then

$$g_J(u_*e_1, u_*je_1) = g_J(u_*e_1, Ju_*e_1) = \frac{1}{2}\left(\omega(u_*e_1, -u_*e_1) + \omega(Ju_*e_1, Ju_*e_1)\right) = 0$$

while

$$g_J(u_*je_1, u_*je_1) = g_J(Ju_*e_1, Ju_*e_1) = \omega(Ju_*e_1, -u_*e_1) = g_J(u_*e_1, u_*e_1).$$

Thus

$$\begin{aligned} |(u_*e_1) \wedge (u_*je_1)|_{g_J} &= \sqrt{g_J(u_*e_1, u_*e_1)g_J(u_*je_1, u_*je_1) - g_J(u_*e_1, u_*je_1)^2} = g_J(u_*e_1, u_*e_1) \\ &= \omega(u_*e_1, u_*je_1) = u^*\omega(e_1, je_1). \end{aligned}$$

Thus the area 2-form α coincides with $u_*\omega$, so $Area(u) = \int_\Sigma u^*\omega$. However, when J is just ω-tame it's no longer true that an *arbitrary* map $u \colon \Sigma \to M$ always has area at least $\int_\Sigma u^*\omega$.

When J is ω-compatible it is also true that J-holomorphic curves minimize the Dirichlet energy $E(u)$ among curves coinciding with them outside a compact subset of the interior, as the following shows:[1]

Proposition 2.6. *If (Σ, j) is a compact almost complex 2-manifold and $u \colon \Sigma \to M$ is C^1 then*

$$E(u) := \frac{1}{2}\int_\Sigma |du|^2 vol \geq \int_\Sigma u^*\omega,$$

with equality if and only if u is J-holomorphic.

Proof. Choose a metric h on Σ in the conformal class $\mathcal{H}_j$. Define

$$\bar\partial_J u = \frac{1}{2}(du + J \circ du \circ j)$$

and, for $p \in \Sigma$, choose an h-orthonormal basis $\{e_1, e_2\}$ for $T_p\Sigma$ with $e_2 = je_1$ and define

$$|\bar\partial_J u(p)|^2 = g_J(\bar\partial_J u(e_1), \bar\partial_J u(e_1)) + g_J(\bar\partial_J u(e_2), \bar\partial_J u(e_2)).$$

Now

$$\bar\partial_J u(e_1) = \frac{1}{2}(u_*e_1 + Ju_*e_2)$$

and

$$\bar\partial_J u(e_2) = \frac{1}{2}(u_*e_2 - Ju_*e_1).$$

[1] Actually, the fact that they minimize the Dirichlet energy implies that they're minimal surfaces, as explained in [Law, p. 61]; however it's quite a bit harder to prove this implication than it is to just prove area-minimization directly as we did above.

Hence

$$\begin{aligned}|\bar{\partial}_J u(p)|^2 &= \omega(\bar{\partial}_J u(e_1), J\bar{\partial}_J u(e_1)) + \omega(\bar{\partial}_J u(e_2), J\bar{\partial}_J u(e_2)) \\ &= \frac{1}{4}\left(\omega(u_*e_1 + Ju_*e_2, Ju_*e_1 - u_*e_2) + \omega(-Ju_*e_1 + u_*e_2, u_*e_1 + Ju_*e_2)\right) \\ &= \frac{1}{2}\left(\omega(u_*e_1, Ju_*e_1) + \omega(u_*e_2, Ju_*e_2) - \omega(u_*e_1, u_*e_2) - \omega(Ju_*e_1, Ju_*e_2)\right) \\ &= \frac{1}{2}|du(p)|^2 - u^*\omega(e_1, e_2)\end{aligned}$$

where we've used compatibility of J with ω in the last equation. Multiplying by the volume form of h and integrating over Σ then gives, in obvious notation

$$E(u) - \int_\Sigma u^*\omega = \int_\Sigma |\bar{\partial}_J u|^2 vol.$$

Since the right hand side is nonnegative and is zero iff u is J-holomorphic the result follows. □

3. Exact Lagrangian submanifolds and the non-squeezing theorem

A substantial amount of machinery needs to be set up before we can begin to prove significant applications of the theory of pseudoholomorphic curves; before setting up the machinery I'd like to discuss two sample applications of the theory to indicate how pseudoholomorphic curves can be used in practice. Both of these results were proven in Gromov's paper [Gr], which initiated the whole theory. Proofs will be outlined in sufficient detail as to make clear that, in both cases, the pivotal point is the existence of a pseudoholomorphic curve with certain properties—the proof that this curve exists (or indeed any indication of why it would be reasonable to expect it to exist) will have to wait until later.

The first result deals with the properties of Lagrangian submanifolds in $\mathbb{R}^{2n}$. Consider a bit more generally a symplectic manifold $(M, d\lambda)$ where the symplectic form is an exact 2-form. If $L \subset M$ is Lagrangian, *i.e.* $\dim L = \frac{1}{2}\dim M$ and $d\lambda|_L = 0$, then (since d commutes with restriction) the 1-form $\lambda|_L \in \Omega^1(L)$ is closed. The Lagrangian submanifold L is then called *exact* if $\lambda|_L$ is exact, *i.e.* if there is a function $f\colon L \to \mathbb{R}$ such that $df = \lambda|_L$. Note that in principle the notion of exactness depends on the primitive λ, since λ could be replaced by $\lambda + \theta$ for any closed 1-form θ on M to get the same symplectic manifold $(M, d\lambda)$. Clearly L is automatically exact if all closed 1-forms on L are exact, *i.e.* if it has first Betti number $b_1(L) = 0$.

The standard example of an exact Lagrangian submanifold occurs when $M = T^*N$ is the cotangent bundle of a smooth manifold N, and $\lambda \in \Omega^1(T^*N)$ is the Liouville 1-form $\lambda_{(x,p)}(v) = p(\pi_* v)$ where $\pi\colon T^*N \to N$ is the bundle projection. Then where L is the zero-section of T^*N we in fact have $\lambda|_L = 0$, so certainly L is exact. You might also recall that a standard set of Lagrangian submanifolds in $(T^*N, d\lambda)$ is given by the images L_α of closed 1-forms $\alpha \in \Omega^1(N)$, and one can show that L_α is exact if and only if α is exact as a 1-form.

The standard symplectic structure $\omega_0 = \sum_{i=1}^n dx^i \wedge dy^i$ is certainly exact, for example $\omega_0 = d\lambda$ where $\lambda = \sum_{i=1}^n x_i dy^i$.

Theorem 3.1. *In the symplectic manifold $(\mathbb{R}^{2n}, \omega_0)$, there are no embedded closed exact Lagrangian submanifolds.*

You might first try proving this when $2n = 2$. When $2n = 4$, note that it in particular follows that neither S^2 nor $\mathbb{R}P^2$ can be embedded into $(\mathbb{R}^4, \omega_0)$ as a Lagrangian submanifold, since these manifolds have $b_1 = 0$.

Sketch of proof. The key is the following lemma of Gromov:

Lemma 3.2. *Where J_0 is the standard almost complex structure on $\mathbb{R}^{2n}$ compatible with ω_0, if L is any embedded closed Lagrangian submanifold there is a nonconstant J_0-holomorphic curve $u\colon D^2 \to \mathbb{R}^{2n}$ with $u(\partial D^2) \subset L$*

Note that when $2n = 2$ the Riemann mapping theorem produces such a curve. I hope to prove this lemma by the end of the course.

Given the lemma, suppose for contradiction that L were an exact Lagrangian submanifold. With $u\colon (D^2, \partial D^2) \to (M, L)$ as in the lemma, Stokes' theorem, Proposition 2.1 and the fact that u is nonconstant give

$$0 < E(u) = \int_{D^2} u^*\omega_0 = \int_{D^2} d(u^*\lambda) = \int_{\partial D^2} (u|_{\partial D^2})^*\lambda.$$

Hence $u|_{\partial D^2}\colon S^1 \to L$ is a loop in L over which λ has nontrivial integral. But if $\lambda|_L$ is an exact one-form no such loop exists (by another application of Stokes' theorem). So $\lambda|_L$ must not be exact, *i.e.*, L is not exact. □

Interestingly, when $2n \geq 4$ it's possible to construct a symplectic form $\omega' = d\lambda'$ on $\mathbb{R}^{2n}$ with the property that $(\mathbb{R}^{2n}, \omega')$ *does* have closed exact Lagrangian submanifolds. It follows that $(\mathbb{R}^{2n}, \omega')$ is not symplectomorphic to $(\mathbb{R}^{2n}, \omega_0)$ (or indeed to any subset thereof)! Some details of this are explained at the end of [AL, Chapter X]. On the other hand, with a bit of work one can show that any symplectic form on $\mathbb{R}^2$ is symplectomorphic to ω_0.

The other result that I'll mention is the famous *nonsqueezing theorem*. Consider $\mathbb{R}^{2n}$ with its standard symplectic structure $\omega_0 = \sum_{i=1}^n dx^i \wedge dy^i$. For $r > 0$ let

$$B^{2n}(r) = \left\{ (\vec{x}, \vec{y}) \in \mathbb{R}^{2n} | \sum_{i=1}^n (x_i^2 + y_i^2) < r^2 \right\}$$

and

$$Z^{2n}(r) = \left\{ (\vec{x}, \vec{y}) \in \mathbb{R}^{2n} | x_1^2 + y_1^2 < r^2 \right\}$$

So $B^{2n}(r)$ is a standard ball of radius r and $Z^{2n}(r)$ is a cylinder of radius r (over a circle in the x_1y_1 plane.

Theorem 3.3. *If $r, R > 0$ and if there is a symplectic embedding $\phi\colon B^{2n}(r) \hookrightarrow Z^{2n}(R)$ (i.e. ϕ is an embedding and $\phi^*\omega_0 = \omega_0$), then $r \leq R$.*

Note that this is sensitive to how exactly the cylinder was defined—if instead we put $Y^{2n}(r) = \{x_1^2 + x_2^2 < r^2\}$ then for any $\epsilon > 0$ the symplectomorphism $(\vec{x}, \vec{y}) \mapsto (\epsilon\vec{x}, \epsilon^{-1}\vec{y})$ maps $B^{2n}(1)$ to $Y^{2n}(\epsilon)$. Prior to this theorem rather little was known about how the properties of symplectomorphisms differ from those of volume-preserving diffeomorphisms, but this is obviously a strong constraint that goes well beyond volume preservation.

Sketch of proof. We first "compactify" the problem. Let ϵ be any number with $0 < \epsilon < r/2$. Then ϕ restricts to the closed ball $\overline{B^{2n}(r-\epsilon)}$ as a continuous map on a compact set, which therefore has image contained in $B^2(R) \times \{(\vec{x}', \vec{y}') \in \mathbb{R}^{2n-2} | |\vec{x}'|^2 + |\vec{y}'|^2 < M\}$. Write $S^2(R+\epsilon)$ for the sphere endowed with a symplectic form of area $\pi(R+\epsilon)^2$ (this contains $B^2(R)$ as an open subset) and write $T^{2n-2}(M) = \mathbb{R}^{2n-2}/2M\mathbb{Z}^2$ endowed with the symplectic form pushed down from the standard one on $\mathbb{R}^{2n-2}$. Let $g\colon Z^{2n}(R) = B^2(R) \times \mathbb{R}^{2n-2} \to S^2(R+\epsilon) \times T^{2n-2}(M)$ be the map given by inclusion on the first factor and projection on the second; then

$$\psi := g \circ \phi\colon B^{2n}(r-\epsilon) \to S^2(R+\epsilon) \times T^{2n-2}(M)$$

is a symplectic embedding (where we use the product symplectic form Ω on the codomain). The key lemma is:

Lemma 3.4. *Let $J \in \mathcal{J}(S^2(R+\epsilon) \times T^{2n-2}(M), \Omega)$ be any compatible almost complex structure and let $p \in S^2(R-\epsilon) \times T^{2n-2}(M)$. Then there is a J-holomorphic sphere $u\colon S^2 \to S^2(R+\epsilon) \times T^{2n-2}(M)$ such that $u(\vec{0}) = p$ which represents the homology class $[S^2 \times (pt)] \in H_2(S^2 \times T^{2n-2}; \mathbb{Z})$.*

We should be able to prove this soon after we start discussing Gromov–Witten invariants in the middle of the course.

Since ψ is an embedding, we can put an almost complex structure on its image by pushing forward the standard almost complex structure J_0 on $B^{2n}(r)$ (i.e. this almost complex structure is $\psi_* \circ J_0 \circ \psi_*^{-1}$). The same argument that proves the contractibility of $\mathcal{J}(S^2(R+\epsilon) \times T^{2n-2}(M), \Omega)$ can easily be adapted to produce $J \in \mathcal{J}(S^2(R+\epsilon) \times T^{2n-2}(M), \Omega)$ which restricts to any preassigned compact subset of $Im(\psi)$ (say $\overline{\psi(B^{2n}(r-2\epsilon))}$) as the pushed-forward almost complex structure $\psi_* \circ J_0 \circ \psi_*^{-1}$. Now let $u\colon S^2 \to S^2(R+\epsilon) \times T^{2n-2}(M)$ be the pseudoholomorphic curve produced by the lemma. We have $Area(u) = \pi(R+\epsilon)^2$ by construction and Proposition 2.4.

Consider $S := u^{-1}(\psi(\overline{B^{2n}(r-2\epsilon)}))$. For simplicity assume that this is a submanifold with boundary of $\overline{B^{2n}(r-2\epsilon)}$ with boundary contained in $\partial(B^{2n}(r-2\epsilon))$ (it's not hard to use Sard's theorem to show that this will at least hold with 2ϵ replaced by $2\epsilon+\delta$ for arbitrarily small δ). By the construction of J, S is the image of a J_0-holomorphic curve (potentially with multiple connected components); moreover by construction S both contains the origin and intersects $\partial B^{2n}(r-2\epsilon)$ (the latter is true because of the condition on the homology class represented by u). It turns out to be possible to show that this implies that S has area at least $\pi(r-2\epsilon)^2$. (At least when [Gr] was written, the easiest way to show this was to note that, as mentioned after Proposition 2.4, S is a minimal surface, and then to appeal to a "monotonicity" result from minimal surface theory which shows that a minimal surface in a ball of radius ρ in $\mathbb{R}^{2n}$ with its standard metric that passes through the origin and is compact with boundary on the boundary of the ball has at least $\pi\rho^2$. This is proven in [AL, Chapter 3, Theorem 3.2.4]; Google also turns up a very-detailed and locally-produced proof at [Ra]. Later in these notes we will prove the relevant area estimate directly from the J_0-holomorphicity of S using just the isoperimetric inequality without appealing to anything from minimal surface theory—see Remark 8.8.) Since the almost complex structures have been set up so that $\psi|_{\psi^{-1}(\overline{B^{2n}(r-2\epsilon)})}$ is an isometric embedding, this shows that $u|_{\psi^{-1}(\overline{B^{2n}(r-2\epsilon)})}$ has area at least $\pi(r-2\epsilon)^2$. But since $Area(u) = \pi(R+\epsilon)^2$ we get

$$\pi(R+\epsilon)^2 \geq \pi(r-2\epsilon)^2.$$

ϵ was arbitrary, so $r \leq R$.

□

4. Sobolev spaces

For a natural number d and for $1 \leq p < \infty$, let $L^p(\mathbb{R}^d)$ denote as usual the space of Lebesgue measurable functions $f\colon \mathbb{R}^d \to \mathbb{R}$ such that

$$\int_{\mathbb{R}^d} |f(x)|^p d^d x < \infty,$$

equipped with the L^p norm

$$\|f\|_p = \left(\int_{\mathbb{R}^d} |f(x)|^p d^d x \right)^{1/p}.$$

If we regard two functions as being the same if they coincide Lebesgue-almost-everywhere,[2] the Minkowski inequality shows that $(L^p(\mathbb{R}^d), \|\cdot\|_p)$ is a normed vector space. Moreover, an important standard theorem in real analysis shows that this normed vector space is a *Banach space*, *i.e.* a normed vector space whose associated metric is *complete* (every Cauchy sequence has a limit in the space).

While $(L^p(\mathbb{R}^d), \|\cdot\|_p)$ contains some functions that are rather poorly behaved, it has dense subspaces consisting of much better-behaved functions. Notably:

Proposition 4.1. *The set*

$$C_0^\infty(\mathbb{R}^d) = \left\{ f\colon \mathbb{R}^d \to \mathbb{R} \,\middle|\, \begin{array}{l} \textit{support}(f) \textit{ is compact and all partial derivatives} \\ \textit{of } f \textit{ of all orders exist and are continuous} \end{array} \right\}$$

is dense in $(L^p(\mathbb{R}^d), \|\cdot\|_p)$.

[2] As one learns to do in a real analysis course, we'll generally neglect to include annoying qualifiers like "modulo equivalence almost everywhere." Note that any two *continuous* functions that coincide almost everywhere in fact coincide everywhere, since the set on which they disagree is open and any nonempty open set has positive measure. An element of $L^p(\mathbb{R}^d)$ which has a continuous function in its almost-everywhere-equivalence class will always be identified with this continuous function (which is unique when it exists by the previous sentence). Thus, when we say something like "$f \in L^p(\mathbb{R}^d)$ is differentiable," what strictly speaking is meant is "there is a continuous function equal to f almost everywhere, and this continuous function is differentiable."

In other words, for any $f \in L^p(\mathbb{R}^d)$ there is a sequence $\{f_n\}_{n=1}^{\infty}$ of compactly supported, smooth functions such that

$$\lim_{n\to\infty} \|f_n - f\|_p = 0.$$

Since we'll have occasion to use not just the statement of Proposition 4.1 but also some facts that arise in the course of (one approach to) proving it, let us give a detailed proof, which involves the properties of a procedure called *mollification*.

Choose an arbitrary smooth function $\beta\colon \mathbb{R}^d \to [0,1]$ such that $\beta(x) = \beta(-x)$ for all x, $\beta(x) = 0$ whenever $\|x\| \geq 1$ and such that $\int_{\mathbb{R}^d} \beta(x) d^d x = 1$. Now set $\beta_n(x) = n^d \beta(nx)$; thus β_n has support in $\{\|x\| \leq 1/n\}$ and still has integral 1. Now if $f \in L^p(\mathbb{R}^d)$ define

$$f_n(x) = (\beta_n * f)(x) := \int_{\mathbb{R}^d} \beta_n(x-y) f(y) d^d y.$$

We will show that the f_n are smooth functions which belong to $L^p(\mathbb{R}^d)$ and have $\|f_n - f\|_p$. (If f is compactly supported it's easy to see that the f_n are as well.)

More generally, if u and v are two functions on $\mathbb{R}^d$ (sufficiently regular for the following to be defined) the *convolution* of u and v is the function

$$(u * v)(x) = \int_{\mathbb{R}^d} u(x-y) v(y) d^d y.$$

Note that a simple change of variables shows that $u * v = v * u$ when both are defined. Here are some basic properties of convolutions:

Theorem 4.2. *Let $u \in C_0^\infty(\mathbb{R}^d)$ and $f \in L^p(\mathbb{R}^d)$ with $1 \leq p < \infty$. Then:*

(i) *The integral $(u * f)(x) = \int_{\mathbb{R}^d} u(x-y) f(y) d^d y$ exists for every $x \in \mathbb{R}^d$.*

(ii) *The function $u * f\colon \mathbb{R}^d \to \mathbb{R}$ is differentiable, with partial derivatives*

$$\frac{\partial}{\partial x_i}(u * f) = \left(\frac{\partial u}{\partial x_i}\right) * f.$$

Consequently $u * f \in C^\infty(\mathbb{R}^d)$.

(iii) *(Young's inequality) $u * f \in L^p(\mathbb{R}^d)$, with*

$$\|u * f\|_p \leq \|u\|_1 \|f\|_p.$$

Proof. For any given x, since u is compactly supported there is a bounded set Ω_x such that $u(x-y) = 0$ unless $y \in \Omega_x$. Moreover the continuity and compact support of u ensure that it is bounded; say $|u(z)| \leq M$ for all z. Hence where for a set A we denote by χ_A the indicator function of A (equal to 1 on A and 0 away from A) we have $|u(x-y)f(y)| \leq M|f(y)|\chi_{\Omega_x}(y)$ for all $y \in \mathbb{R}^d$. Hence

$$\int_{\mathbb{R}^d} |u(x-y)f(y)| d^d y \leq M \int_{\mathbb{R}^d} |f(y)|\chi_{\Omega_x}(y) d^d y \leq M \|f\|_p \|\chi_{\Omega_x}\|_q$$

where $\frac{1}{p} + \frac{1}{q} = 1$ by Hölder's inequality. Of course $\|\chi_{\Omega_x}\|_q = vol(\Omega_x)^{1/q} < \infty$, so this shows that the integral defining $(u * f)(x)$ is always absolutely convergent, so $u * f$ is a well-defined function on all of $\mathbb{R}^d$, proving (i).

As for (ii), we have (where e_i is the standard basis vector in the x_i direction)

$$\frac{(u * f)(x + he_i) - (u * f)(x)}{h} = \int_{\mathbb{R}^d} \left(\frac{u(x - y + he_i) - u(x-y)}{h}\right) f(y) d^d y. \tag{3}$$

Now where M' is the maximal value of $\left|\frac{\partial u}{\partial x_i}\right|$, the integrand in (3) is bounded in absolute value by $M'|f(y)|$ as a result of the mean value theorem. Moreover, for a given $x \in \mathbb{R}^d$, there is a bounded set Ω'_x such that whenever $|h| < 1$ the integrand vanishes unless $y \in \Omega'_x$. So in fact the integrand is bounded above (independently of h with

$|h| < 1$) in absolute value by $M'|f(y)|\chi_{\Omega'_x}(y)$, which is integrable by the same Hölder inequality argument as we used in (i). Hence the Lebesgue dominated convergence theorem shows that the limit as $h \to 0$ of (3) is equal to

$$\int_{\mathbb{R}^d} \lim_{h\to 0}\left(\frac{u(x-y+he_i)-u(x-y)}{h}\right) f(y)d^d y = \int_{\mathbb{R}^d} \frac{\partial u}{\partial x_i}(x-y)h(y)d^d y,$$

proving the formula for the derivative in (ii). The statement that $u * f$ is C^∞ follows immediately by induction, since if u belongs to $C_0^\infty(\mathbb{R}^d)$ then so do all derivatives of u.

As for (iii), where $\frac{1}{p}+\frac{1}{q}=1$, we have

$$\begin{aligned}|(u*f)(x)| &\le \int_{\mathbb{R}^d} |f(y)||u(x-y)|d^d y = \int_{\mathbb{R}^d} (|f(y)||u(x-y)|^{1/p})|u(x-y)|^{1/q} d^d y \\ &\le \left(\int_{\mathbb{R}^d} |f(y)|^p |u(x-y)| d^d y\right)^{1/p} \left(\int_{\mathbb{R}^d} |u(x-y)|d^d y\right)^{1/q}.\end{aligned}$$

Thus, since $p/q = p-1$,

$$|(u*f)(x)|^p \le \|u\|_1^{p-1}\left(\int_{\mathbb{R}^d} |f(y)|^p|u(x-y)|d^d y\right),$$

so that

$$\begin{aligned}\int_{\mathbb{R}^d} |(u*f)(x)|^p d^d x &\le \|u\|_1^{p-1}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |f(y)|^p|u(x-y)|d^d y d^d x \\ &\le \|u\|_1^{p-1}\int_{\mathbb{R}^d} |f(y)|^p\left(\int_{\mathbb{R}^d}|u(x-y)|d^d x\right)d^d y = \|u\|_1^p \int_{\mathbb{R}^d}|f(y)|^p d^d y \\ &= \|u\|_1^p\|f\|_p^p,\end{aligned}$$

proving (iii). □

Lemma 4.3. *Suppose that $g\colon \mathbb{R}^d \to \mathbb{R}$ is a continuous, compactly supported function. Then $\beta_n * g \to g$ uniformly. Hence also $\beta_n * g \to g$ in L^p.*

Proof. Recalling that $\beta_n \ge 0$, $supp(\beta_n) \subset B_{1/n}(0)$, and $\int_{\mathbb{R}^d}\beta_n(x)d^d x = 1$, we have

$$\begin{aligned}|\beta_n * g(x) - g(x)| &\le \left|\int_{\mathbb{R}^d} g(x-y)\beta_n(y)d^d y - g(x)\right| = \left|\int_{\mathbb{R}^d}(g(x-y)-g(x))\beta_n(y)d^d y\right| \\ &\le \int_{B_{1/n}(0)} |g(x-y)-g(x)|\beta_n(y)d^d y.\end{aligned}$$

Since g is continuous and compactly supported it is uniformly continuous. Hence if $\epsilon > 0$ there is a natural number N so that $|g(x-y)-g(x)| < \epsilon$ whenever $y \in B_{1/N}(0)$. So once $n \ge N$, again using that $\int_{B_{1/n}(0)}\beta_n(y)d^d y = 1$, the above shows that $|\beta_n * g(x) - g(x)| \le \epsilon$ for all x. This proves that $\beta_n * g \to g$ uniformly.

Convergence in L^p then follows quickly since g has compact support: notice that if Ω is a bounded set which contains every point having distance one or less from a point of the support of g, then both $\beta_n * g$ and g vanish outside Ω. So

$$\|\beta_n * g - g\|_p^p \le \int_\Omega |\beta_n * g(x) - g(x)|^p d^d x,$$

which clearly tends to zero by the uniform convergence of $\beta_n * g$ to g since Ω has finite measure. □

Corollary 4.4. *If $f \in L^p(\mathbb{R}^d)$ then $\|\beta_n * f - f\|_p \to 0$ as $n \to \infty$.*

Proof. Let $\epsilon > 0$. It is then a standard measure theory exercise to find a compactly supported continuous $g\colon \mathbb{R}^d \to \mathbb{R}$ such that $\|g-f\|_p < \epsilon/3$ (First approximate f by simple functions $\sum_{i=1}^m a_i\chi_{A_i}$, then approximate the χ_{A_i} by characteristic functions of finite unions of d-dimensional rectangles, and then approximate these characteristic functions of unions of rectangles by continuous piecewise linear functions). Let N be so large that $\|\beta_n * g - g\|_p < \epsilon/3$ when $n \geq N$, as can be achieved by the preceding lemma. We then have, for $n \geq N$,

$$\begin{aligned}\|\beta_n * f - f\|_p &\leq \|\beta_n * (f - g)\|_p + \|\beta_n * g - g\|_p + \|g - f\|_p \\ &\leq \|\beta_n\|_1 \|f - g\|_p + \frac{2\epsilon}{3} < \epsilon,\end{aligned}$$

where we've used Young's inequality. □

Proof of Proposition 4.1. Let $f \in L^p(\mathbb{R}^d)$ and $\epsilon > 0$. The Lebesgue dominated convergence theorem shows that $\chi_{B_n(0)} f \to f$ in L^p as $n \to \infty$, so let N be large enough that

$$\|\chi_{B_N(0)} f - f\|_p < \epsilon/2.$$

Note that, since $\chi_{B_N(0)} f$ is compactly supported, so are the functions $\beta_m * (\chi_{B_N(0)} f)$ for each m (specifically they are supported in $B_{N+1/m}(0)$. Using Corollary 4.4, let M be so large that $\|\beta_M * (\chi_{B_N(0)} f) - \chi_{B_N(0)} f\|_p < \epsilon/2$. Thus $\beta_M * (\chi_{B_N(0)} f) \in C_0^\infty(\mathbb{R}^d)$, and $\|\beta_M * (\chi_{B_N(0)} f) - f\|_p < \epsilon$, as desired. □

Proposition 4.1 implies that the Banach space $(L_p(\mathbb{R}^d), \|\cdot\|_p)$ can (at least up to isomorphism of Banach spaces) be equivalently characterized as the *completion of* $C_0^\infty(\mathbb{R}^d)$ *with respect to the norm* $\|\cdot\|_p$. *Sobolev spaces* can be characterized as completions of $C_0^\infty(\mathbb{R}^d)$ with respect to certain norms which are stronger than $\|\cdot\|_p$ by virtue of taking into account the behavior of derivatives.

We first introduce standard "multi-index" notation that is used in PDE theory. The letter α will generally denote a *multi-index*, *i.e.* a d-tuple $(\alpha_1, \ldots, \alpha_d)$ of natural numbers. Then for a (sufficiently-differentiable) function $f\colon U \to \mathbb{R}$ where $U \subset \mathbb{R}^d$ is open, we define

$$D^\alpha f = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_d}}{\partial x_d^{\alpha_d}} f$$

whenever the derivative on the right hand side exists. In particular for $\alpha = \vec{0}$ we have $D^\alpha f = f$. Also, for a multi-index α, define

$$|\alpha| = \alpha_1 + \cdots + \alpha_d.$$

Thus the derivative $D^\alpha f$ is a certain partial derivative of order $|\alpha|$.

Definition 4.5. *The* Sobolev space $W^{k,p}(\mathbb{R}^d)$ *is the (Cauchy-)completion of the normed vector space* $C_0^\infty(\mathbb{R}^d)$ *equipped with the* Sobolev norm

$$\|f\|_{k,p} = \sum_{0 \leq |\alpha| \leq k} \|D^\alpha f\|_p.$$

Thus $W^{0,p}(\mathbb{R}^d) = L^p(\mathbb{R}^d)$. As a vector space, I claim that we can also describe $W^{k,p}(\mathbb{R}^d)$ as

$$W^{k,p}(\mathbb{R}^d) = \left\{ f \in L^p(\mathbb{R}^d) \;\middle|\; \begin{array}{c} (\exists \{f_n\}_{n=1}^\infty \subset C_0^\infty(\mathbb{R}^d))(f_n \to f \text{ in } L^p \text{ and for all } \alpha \text{ with } |\alpha| < k, \\ \{D^\alpha f_n\}_{n=1}^\infty \text{ is a Cauchy sequence in } L^p) \end{array} \right\}. \tag{4}$$

Indeed, $W^{k,p}(\mathbb{R}^d)$, as initially defined, consists of equivalence classes of Cauchy sequences of functions in C_0^∞ with respect to $\|\cdot\|_{k,p}$, where $\{f_n\}_{n=1}^\infty$ and $\{g_n\}_{n=1}^\infty$ are equivalent provided that $\|f_n - g_n\|_{k,p} \to 0$. Any such Cauchy sequence $\{f_n\}_{n=1}^\infty$ has $D^\alpha f_n$ Cauchy in L^p for all $|\alpha| \leq k$; in particular f_n is Cauchy and so converges to some $f \in L^p(\mathbb{R}^d)$ which belongs to the set on the right hand side of (4), and if $\{f_n\}_{n=1}^\infty$ and $\{g_n\}_{n=1}^\infty$ are equivalent they have the same L^p limit and so determine the same f. This provides a surjective map from $W^{k,p}$ to the set on the right hand side of (4); to show that the map is injective we need that if both $\{f_n\}_{n=1}^\infty$ and $\{g_n\}_{n=1}^\infty$ are Cauchy with respect to $\|\cdot\|_{k,p}$ and have the same L^p limit f then they are equivalent as Cauchy sequences in $W^{k,p}$ (i.e. $\|D^\alpha f_n - D^\alpha g_n\|_p \to 0$ for $|\alpha| \leq k$). This is not particularly obvious, but follows from:

Lemma 4.6. *(i) If $f, h \in C_0^\infty(\mathbb{R}^d)$ then*

$$\int_{\mathbb{R}^d} (D^\alpha f(x))h(x)d^d x = (-1)^{|\alpha|} \int_{\mathbb{R}^d} f(x) D^\alpha h(x) d^d x.$$

(ii) Let $f \in W^{k,p}(\mathbb{R}^d)$ where $1 < p < \infty$, and choose a multi-index α with $|\alpha| \le k$. Then there is a unique[3] *function $D^\alpha f \in L^p(\mathbb{R}^d)$ obeying, for all $f, h \in C_0^\infty(\mathbb{R}^d)$,*

$$\int_{\mathbb{R}^d} (D^\alpha f(x))h(x)d^d x = (-1)^{|\alpha|} \int_{\mathbb{R}^d} f(x) D^\alpha h(x) d^d x. \tag{5}$$

Specifically, if $\{f_n\}_{n=1}^\infty$ is any sequence in $C_0^\infty(\mathbb{R}^d)$ which is Cauchy in the norm $\|\cdot\|_{k,p}$ and such that $\|f_n - f\|_p = 0$, then $D^\alpha f$ is the L^p limit of $D_\alpha f_n$.

(iii) If $\{f_n\}_{n=1}^\infty$ and $\{g_n\}_{n=1}^\infty$ are both sequences in $C_0^\infty(\mathbb{R}^d)$ which are Cauchy with respect to $\|\cdot\|_{k,p}$ and have the same L^p limit f, then $\|D^\alpha f_n - D^\alpha g_n\|_p \to 0$ when $|\alpha| \le k$. Consequently (4) holds.

Proof. (i) follows by integration by parts (and the fact that the involved functions have compact support): for instance choosing R such that the supports of f and h are contained in $[-R, R]^d$ we have

$$\int_{\mathbb{R}^d} \frac{\partial f}{\partial x_1}(x)h(x)d^d x = \int_{-R}^R \int_{\{s\}\times[-R,R]^{d-1}} \frac{\partial f}{\partial x_1}(s, x')h(s, x')d^{d-1}x' ds$$
$$= -\int_{-R}^R \int_{\{s\}\times[-R,R]^{d-1}} f(s, x')\frac{\partial h}{\partial x_1}(s, x')d^{d-1}x' ds = -\int_{\mathbb{R}^d} f(x)\frac{\partial h}{\partial x_1} d^d x,$$

where we've used the assumption on the supports to see that the boundary term in the integration by parts is zero. The same argument applies to derivatives with respect to the other coordinates x_i (thus proving (i) for $|\alpha| = 1$), and then an induction on $|\alpha|$ proves (i) in general.

As for (ii), if $f \in W^{k,p}(\mathbb{R}^d)$ choose $f_n \in C_0^\infty(\mathbb{R}^d)$ as in (4) and define $D^\alpha f$ as the L^p limit of $D^\alpha f_n$. Let $\frac{1}{p} + \frac{1}{q} = 1$, so since $1 < p < \infty$ we also have $1 < q < \infty$. For $h \in C_0^\infty(\mathbb{R}^d)$ we have $|\int_{\mathbb{R}^d}(D^\alpha f_n - D^\alpha f)hd^d x| \le \|D^\alpha f_n - D^\alpha f\|_p \|h\|_q \to 0$ and $|\int_{\mathbb{R}^d}(f_n - f)D^\alpha h d^d x| \le \|f_n - f\|_p \|D^\alpha h\|_q \to 0$, so

$$\int_{\mathbb{R}^d}(D^\alpha f)hd^d x = \lim_{n\to\infty} \int_{\mathbb{R}^d}(D^\alpha f_n)hd^d x = (-1)^{|\alpha|} \lim_{n\to\infty} \int_{\mathbb{R}^d} f_n D^\alpha h d^d x = (-1)^{|\alpha|} \int_{\mathbb{R}^d} f D^\alpha h d^d x.$$

To see that $D^\alpha f$ is the unique L^p function satisfying (5) for all $h \in C_0^\infty(\mathbb{R}^d)$, suppose that $v \in L^p(\mathbb{R}^d)$ is some other such function, so for all $h \in C_0^\infty(\mathbb{R}^d)$ we have $\int_{\mathbb{R}^d} vhd^d x = (-1)^{|\alpha|} \int_{\mathbb{R}^d} f D^\alpha h d^d x$. But then for all $h \in C_0^\infty(\mathbb{R}^d)$ we have

$$\int_{\mathbb{R}^d}(D^\alpha f - v)hd^d x = 0.$$

Now since $q < \infty$ $C_0^\infty(\mathbb{R}^d)$ is dense in $L^q(\mathbb{R}^d)$, so if $h \in L^q(\mathbb{R}^d)$ we can find $h_n \in C_0^\infty(\mathbb{R}^d)$ such that $h_n \to h$ in L^q, and then

$$\left|\int_{\mathbb{R}^d}(D^\alpha f - v)hd^d x\right| = \lim_{n\to\infty}\left|\int_{\mathbb{R}^d}(D^\alpha f - v)(h - h_n)d^d x\right| \le \lim_{n\to\infty} \|D^\alpha f - v\|_p \|h - h_n\|_q = 0.$$

But for $\int_{\mathbb{R}^d}(D^\alpha f - v)hd^d x = 0$ for all $h \in L^q$ (given that $D^\alpha f - v \in L^p$) it must be that $D^\alpha f = v$ a.e. (For instance, set $h = sgn(D^\alpha f - v)|D^\alpha f - v|^{p/q}$, which belongs to L^q because $D^\alpha f - v \in L^p$.)

As for (iii), either f_n or g_n can be used to construct the function $D^\alpha f$ in (ii), and the uniqueness statement shows that the same function will result from the use of either sequence. Thus $D^\alpha f_n - D^\alpha g_n \to 0$. The discussion in the paragraph before the statement of the lemma implies that (4) follows from this.

□

[3] up to almost-everywhere-equivalence

In general, for any locally integrable function f, a function $D^\alpha f$ is called the *weak derivative* of f of order α if it obeys relation (5) for every compactly supported smooth h. The preceding lemma shows that if $f \in W^{k,p}(\mathbb{R}^d)$ then for all $|\alpha| \leq k$ the weak derivative $D^\alpha f$ exists and belongs to $L^p(\mathbb{R}^d)$. We will presently show that the converse holds.

Lemma 4.7. *Suppose that $f \in L^p(\mathbb{R}^d)$ where $1 < p < \infty$ and that the weak derivative $D^\alpha f$ exists and belongs to $L^p(\mathbb{R}^d)$ for some multi-index α. Then the functions $f_n = \beta_n * f$ from the proof of Proposition 4.1 obey*

$$D^\alpha(\beta_n * f) = \beta_n * D^\alpha f.$$

Proof. Where β_n are the functions described after Proposition 4.1 set $f_n = \beta_n * f$; recall in particular that $\beta_n(-x) = \beta_n(x)$ for all x and that $f_n \in C^\infty(\mathbb{R}^d)$ (specifically we've seen that $D^\alpha f_n = (D^\alpha \beta_n) * f$). For $h \in C_0^\infty(\mathbb{R}^d)$ we have

$$\begin{aligned}
\int_{\mathbb{R}^d} (\beta_n * D^\alpha f)(x) h(x) d^d x &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \beta_n(x-y) D^\alpha f(y) h(x) d^d y d^d x \\
&= \int_{\mathbb{R}^d} D^\alpha f(y) \int_{\mathbb{R}^d} \beta_n(y-x) h(x) d^d x d^d y = \int_{\mathbb{R}^d} D^\alpha f(y) \int_{\mathbb{R}^d} h(y-z)\beta_n(z) d^d z d^d y \\
&= (-1)^{|\alpha|} \int_{\mathbb{R}^d} f(y) D^\alpha (h * \beta_n)(y) d^d y = (-1)^{|\alpha|} \int_{\mathbb{R}^d} f(y) \int_{\mathbb{R}^d} D^\alpha h(y-z) \beta_n(z) d^d z d^d y \\
&= (-1)^{|\alpha|} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(y)\beta_n(y-x) D^\alpha h(x) d^d x \\
&= (-1)^{\alpha} \int_{\mathbb{R}^d} (\beta_n * f)(x) D^\alpha h(x) d^d x,
\end{aligned}$$

which proves that $D^\alpha(\beta_n * f) = \beta_n * D^\alpha f$ by the uniqueness statement in Lemma 4.6(ii). □

Corollary 4.8. *Suppose that $f \in L^p(\mathbb{R}^d)$ where $1 < p < \infty$ and that the weak derivative $D^\alpha f$ exists and belongs to $L^p(\mathbb{R}^d)$ for all $|\alpha| \leq k$. Then $f \in W^{k,p}(\mathbb{R}^d)$, $\beta_n * f \in W^{k,p}(\mathbb{R}^d)$, and*

$$\|\beta_n * f - f\|_{k,p} \to 0.$$

Proof. First let us show that $f_n := \beta_n * f \in W^{k,p}(\mathbb{R}^d)$. Using Theorem 4.2 and Lemma 4.7, we have

$$D^\alpha f_n = \beta_n * D^\alpha f \in C^\infty(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$$

for all n (note that the f_n may not be compactly supported if f is not compactly supported). For each natural number m let χ_m be a compactly supported smooth function equal to 1 on the ball $B_m(0)$ and with all partial derivatives obeying $|D^\alpha \chi_m(x)| \leq 1$ for all x. Then the product rule and the dominated convergence theorem readily show that, since $f_n \in C^\infty(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$, $\|D^\alpha(\chi_m f_n) - D^\alpha f_n\|_p \to 0$ as $m \to \infty$. Thus $\{\chi_m f_n\}_{m=1}^\infty$ is a Cauchy sequence in $C_0^\infty(\mathbb{R}^d) \cap W^{k,p}(\mathbb{R}^d)$, with limit f_n, proving that $f_n \in W^{k,p}(\mathbb{R}^d)$.

Given this, the result follows immediately, since for all $|\alpha| \leq k$, $\{D^\alpha f_n\}_{n=1}^\infty = \{\beta_n * D^\alpha f\}_{n=1}^\infty$ is Cauchy in L^p with limit $D^\alpha f$. Thus $\{f_n\}_{n=1}^\infty$ forms a Cauchy sequence in $W^{k,p}(\mathbb{R}^d)$, whose L^p-limit, namely f, therefore belongs to $W^{k,p}(\mathbb{R}^d)$ by Lemma 4.6 and the remarks preceding it. Moreover

$$\|f_n - f\|_{k,p} = \sum_{|\alpha| \leq k} \|D^\alpha f_n - D^\alpha f\|_p \to 0$$

as $n \to \infty$. □

The space $W^{k,p}(\mathbb{R}^d)$ is sometimes (for instance in [Ev]) instead defined as the space of functions with the property that the weak derivative $D^\alpha f$ exists and belongs to L^p for every multi-index α with $|\alpha| \leq k$, and Lemma 4.6 (ii) and the above corollary show that this is an equivalent definition.

While a function $f \in W^{k,p}(\mathbb{R}^d)$ has *weak* derivatives in the sense of (5), this does not imply that f is differentiable in the usual sense of difference quotients. However, recalling that f is approximated in the $W^{k,p}$

norm by C^∞ functions, we note that, if this approximation property holds in a stronger sense, then we do get differentiability:

Proposition 4.9. *Suppose that $\{f_n\}_{n=1}^\infty$ is a sequence of continuously differentiable functions such that $f_n \to f$ uniformly and $\nabla f_n \to G$ uniformly (where $f\colon \mathbb{R}^d \to \mathbb{R}$ and $G\colon \mathbb{R}^d \to \mathbb{R}^d$), then f is differentiable and $\nabla f = G$.*

(Recall the standard fact in undergraduate analysis that a the limit of a uniformly convergent sequence of continuous functions is continuous; hence f and G are automatically continuous under the assumption of the proposition).

Proof. For a general continuously differentiable function $g\colon \mathbb{R}^d \to \mathbb{R}$, if $x, h \in \mathbb{R}^d$ applying the fundamental theorem of calculus to the function $r\colon [0, \|h\|] \to \mathbb{R}$ defined by $r(t) = g(x + th/\|h\|)$ shows that

$$g(x+h) - g(x) = \int_0^{\|h\|} \nabla g(x + th/\|h\|) \cdot (h/\|h\|) dt.$$

So

$$f(x+h) - f(x) = \lim_{n\to\infty} (f_n(x+h) - f_n(x)) = \lim_{n\to\infty} \int_0^{\|h\|} (\nabla f_n(x + th/\|h\|) \cdot (h/\|h\|))\, dt.$$

Now if $\epsilon > 0$ then once n is so large that $\sup \|\nabla f_n - G\| < \frac{\epsilon}{\|h\|}$ the integral on right hand side above differs from $\int_0^{\|h\|} G(x + th/\|h\|) \cdot (h/\|h\|) dt$ by at most ϵ. Hence

$$f(x+h) - f(x) = \int_0^{\|h\|} G\left(x + t\frac{h}{\|h\|}\right) \cdot \frac{h}{\|h\|} dt = G(x) \cdot h + o(\|h\|)$$

as $h \to 0$ by the continuity of G. Thus f is differentiable and $\nabla f = G$. □

The behavior of general elements the Sobolev spaces $W^{k,p}(\mathbb{R}^d)$ is qualitatively different depending on whether $p < d$, $p = d$, or $p > d$. The nicest case, which is the one that we'll need most often, is where $p > d$, so we'll largely restrict attention to this case (see [Ev, Chapter 5] for a much broader introduction to Sobolev spaces which in particular covers the other cases). The role of the assumption $p > d$ boils down to the following:

Proposition 4.10. *Assume $d < p < \infty$. For $x \in \mathbb{R}^d$, where $B_r(x)$ is the open ball of radius r around x, there is a constant C depending only on d and p such that, where $\frac{1}{p} + \frac{1}{q} = 1$,*

$$\left(\int_{B_r(x)} \left(\frac{1}{\|y - x\|^{d-1}}\right)^q d^d y\right)^{1/q} = Cr^{1-d/p}.$$

A formula for C can be extracted from the proof below, but its precise form won't be important for us. Note that if $p = d$ the integral on the left hand side isn't even finite (despite what one might guess from the shape of the right hand side), which is part of the reason why the case $p = d$ differs in significant ways from $p > d$.

Proof. Where dS denotes the standard surface area measure on the unit sphere $\partial B_1(\vec{0})$ (so for a function f we have $\int_{B_r(\vec{0})} f(x) d^n x = \int_{\partial B_1(\vec{0})} \int_0^r f(sw) s^{d-1} ds dS(w)$) and where $\alpha_d = \int_{\partial B_1(\vec{0})} 1 dS(w)$ is the "surface area" of $\partial B_1(\vec{0})$, we have

$$\begin{aligned}\int_{B_r(x)} \left(\frac{1}{\|y - x\|^{d-1}}\right)^q d^d y &= \int_{\partial B_1(\vec{0})} \int_0^r \frac{s^{d-1}}{s^{(d-1)q}} ds dS(w) \\ &= \alpha_d \frac{r^{d-q(d-1)}}{d - q(d-1)}.\end{aligned}$$

The result will follow as soon as we show that $(d - q(d-1))/q = 1 - d/p$. We compute:

$$(d - q(d-1))/q = d\left(1 - \frac{1}{p}\right) - (d-1) = \left(d - \frac{d}{p}\right) - (d-1) = 1 - \frac{d}{p},$$

as desired. □

In general, if $U \subset \mathbb{R}^d$ is a bounded open subset and if $f \in L^1(U)$, write

$$\fint_U f(x)d^dx = \frac{\int_U f(x)d^dx}{\int_U 1d^dx} = \frac{1}{vol(U)}\int_U f(x)d^dx,$$

i.e., $\fint_U f(x)d^dx$ is the average value of f over U. Proposition 4.10 is relevant largely due to the following:

Lemma 4.11. *There is a constant C, depending only on d, with the following property. Let $f \in C_0^\infty(\mathbb{R}^d)$, $x \in \mathbb{R}^d$, and write $B_r(x)$ for the ball of radius r around x. Then*

$$\fint_{B_r(x)} |f(x) - f(y)|d^dy \le C\int_{B_r(x)} \frac{\|\nabla f(y)\|}{\|y - x\|^{d-1}}d^dy.$$

Proof. Recall that the fundamental theorem of calculus gives, for $w \in B_1(\vec{0})$ and $s > 0$,

$$f(x + sw) - f(x) = \int_0^s \nabla f(x + tw) \cdot wdt.$$

Hence, where we use the substitution $y = x + tw$ and convert from spherical to Euclidean coordinates,

$$\begin{aligned}\int_{\partial B_1(\vec{0})} |f(x + sw) - f(x)|dS(w) &\le \int_0^s \int_{\partial B_1(\vec{0})} \|\nabla f(x + tw)\|dS(w)dt \\ &= \int_{B_s(x)} \frac{\|\nabla f(y)\|}{\|y - x\|^{d-1}}d^dy.\end{aligned}$$

Thus

$$\begin{aligned}\int_{B_r(x)} |f(x) - f(y)|d^dy &= \int_0^r s^{n-1}\int_{\partial B_1(\vec{0})} |f(x) - f(x + sw)|dS(w)ds \\ &\le \int_0^r s^{d-1}\int_{B_s(x)} \frac{\|\nabla f(y)\|}{\|y - x\|^{d-1}}d^dyds \le \int_0^r s^{d-1}\int_{B_r(x)} \frac{\|\nabla f(y)\|}{\|y - x\|^{d-1}}d^dy \\ &= \frac{r^d}{d}\int_{B_r(x)} \frac{\|\nabla f(y)\|}{\|y - x\|^{d-1}}d^dy.\end{aligned}$$

Since the volume of $B_r(x)$ is a dimensional constant times r^d, dividing through by the volume of $B_r(x)$ then proves the lemma. □

Theorem 4.12 (Morrey's inequality)**.** *Let $p > d$. Then there are constants C_1 and C_2 such that for any $f \in C_0^\infty(\mathbb{R}^d)$ we have*

$$|f(x)| \le C_1\|f\|_{1,p} \text{ and } |f(x) - f(y)| \le C_2\|f\|_{1,p}\|x - y\|^{1-d/p}$$

for every $x, y \in \mathbb{R}^d$.

Proof. Let $\frac{1}{p} + \frac{1}{q} = 1$. If $x \in \mathbb{R}^d$ we have

$$\begin{aligned}|f(x)| &= \fint_{B_1(x)} |f(x)|d^dy \le \fint_{B_1(x)} (|f(y)| + |f(y) - f(x)|)d^dy \\ &\le \frac{1}{vol(B_1(x))}\int_{B_1(x)} |f(y)|d^dy + C\int_{B_1(x)} \frac{\|\nabla f(y)\|}{\|y - x\|^{d-1}}d^dy \\ &\le \frac{1}{vol(B_1(x))}\|f\|_{L^p(B_1(x))}\|1\|_{L^q(B_1(x))} + C\|\nabla f\|_{L^p(B_1(x))}\left\|\frac{1}{\|y - x\|^{d-1}}\right\|_{L^q(B_1(x))} \\ &\le C\|f\|_{1,p}\end{aligned}$$

for an appropriate constant C, where we've used Hölder's inequality and the fact that, by Proposition 4.10, $\left\|\frac{1}{\|y-x\|^{d-1}}\right\|_{L^q(B_1(x))}$ is some (finite) constant depending only on d and p when $p > d$.

For the second statement let $r = \|x - y\|$ and let $W = B_r(x) \cap B_r(y)$. Since W contains the ball of radius $r/2$ around the midpoint of x and y, W has volume at least 2^{-d} times the common volume of $B_r(x)$ and $B_r(y)$. Of course,

$$|f(x) - f(y)| = \fint_W |f(x) - f(y)| d^d z \leq \fint_W |f(z) - f(x)| d^d z + \fint_W |f(z) - f(y)| d^d z.$$

Now, where the M_j below denote constants depending only on d and p, since $W \subset B_r(x)$

$$\fint_W |f(z) - f(x)| d^d z \leq \frac{1}{vol(W)} \int_{B_r(x)} |f(z) - f(x)| d^d z \leq \frac{vol(B_r(x))}{vol(W)} \fint_{B_r(x)} |f(z) - f(x)| d^d z$$
$$\leq M_1 \fint_{B_r(x)} \frac{\|\nabla f(x)\|}{\|y - x\|^{d-1}} d^d y \leq M_2 \|\nabla f\|_{L^p(B_r(x))} r^{1-d/p} \leq M_2 \|\nabla f\|_p r^{1-d/p}.$$

By the same token

$$\fint_W |f(z) - f(y)| d^n z \leq M_2 \|\nabla f\|_p r^{1-d/p},$$

and so

$$|f(x) - f(y)| \leq 2M_2 \|\nabla f\|_p r^{1-d/p} = 2M_2 \|f\|_{k,p} \|x - y\|^{1-d/p}.$$

□

Corollary 4.13. *Let $f \in W^{k,p}(\mathbb{R}^d)$, with $k \geq 1$ and $p > d$. Then (some function equal almost everywhere to) f is $(k-1)$-times continuously differentiable, and there is a constant C depending on k, p, d but not on f such that for all $x, y \in \mathbb{R}^d$ and all β with $0 \leq |\beta| \leq k - 1$ we have*

$$|D^\beta f(x)| \leq C \|f\|_{k,p} \text{ and } |D^\beta f(x) - D^\beta f(y)| \leq C \|f\|_{k,p} \|x - y\|^{1-d/p}.$$

Proof. Where $\chi_n \colon \mathbb{R}^n \to [0, 1]$ belongs to $C_0^\infty(\mathbb{R}^d)$ and is equal to 1 on $B_n(\vec{0})$, if $f \in W^{k,p}(\mathbb{R}^d)$ then $\chi_n f \in W^{k,p}(\mathbb{R}^d)$, and f satisfies the conclusions of the corollary if and only if $\chi_n f$ does for every n (take n large enough that $\|x\|, \|y\| \leq n$). So there is no loss of generality in assuming that f vanishes identically outside some compact set.

Let $f_n \in C_0^\infty(\mathbb{R}^d)$ be a Cauchy sequence with respect to the norm $\|\cdot\|_{k,p}$ which converges in L^p (hence also in $W^{k,p}$ by what we've previously shown) to f. Thus for $|\beta| \leq k - 1$, writing $l = k - |\beta|$, we have $l \geq 1$ and $D^\beta f_n$ is Cauchy in $W^{l,p}$ and converges in L^p to $D^\beta f$. Then for $x \in \mathbb{R}^d$ we have $|D^\beta f_n(x) - D^\beta f_m(x)| \leq C_1 \|D^\beta f_n - D^\beta f_m\|_{1,p} \to 0$, so $\{D^\beta f_n\}_{n=1}^\infty$ is uniformly Cauchy and therefore has a pointwise limit to which it converges uniformly. But since $D^\beta f_n \to D^\beta f$ in L^p this forces $D^\beta f$ to be equal to this uniform limit (modulo the almost everywhere ambiguity in the definition of $D^\beta f$). In particular since uniform limits of continuous functions are continuous this proves that $D^\beta f$ is continuous whenever $|\beta| \leq k - 1$. Moreover, the fact that the covergence of the $D^\beta f_n$ to $D^\beta f$ is uniform shows, via Proposition 4.9 and induction on $|\beta|$, that $D^\beta f$ is in fact equal to the order-β derivative of f (in the difference quotient sense, not just the weak sense (5)).

Since $D^\beta f_n \to D^\beta f$ uniformly and since $\|D^\beta f_n\|_{1,p} \to \|D^\beta f\|_{1,p} \leq \|f\|_{k,p}$, the two inequalities in the statement of the corollary follow directly from Theorem 4.12. □

In view of this, when $k \geq 1$ and $p > d$, we can and hereinafter do identify $W^{k,p}(\mathbb{R}^d)$ with a subspace of the space of continuous functions on $\mathbb{R}^d$ (in particular identifying elements of $W^{k,p}(\mathbb{R}^d)$, which we previously defined as almost-everywhere-equivalence classes, with the unique continuous functions that represent them).

Corollary 4.14. *Let $k \geq 1$ and $p > d$. Then there is a constant C such that, if $f, g \in W^{k,p}(\mathbb{R}^d)$ then $fg \in W^{k,p}(\mathbb{R}^d)$ and*

$$\|fg\|_{k,p} \leq C \|f\|_{k,p} \|g\|_{k,p}.$$

Proof. We have (as can easily be verified using approximating sequences of smooth functions), for $|\alpha| \leq k$,

$$D^\alpha(fg) = \sum_{\beta+\gamma=\alpha} (D^\beta f) D^\gamma g.$$

Thus it suffices to show that whenever $|\beta| + |\gamma| \leq k$ there is a constant $C_{\beta\gamma}$ such that

$$\|D^\beta f D^\gamma g\|_p \leq C_{\beta\gamma}\|f\|_{k,p}\|g\|_{k,p}. \tag{6}$$

Since $k \geq 1$ if $|\beta| + |\gamma| \leq k$ then one or both of β and γ is strictly less than k; without loss of generality assume $|\gamma| < k$. Then by Corollary 4.13 we have a uniform bound $\|D^\gamma g\|_{C^0} \leq C'\|D^\gamma g\|_{1,p} \leq C'\|g\|_{k,p}$ (where $\|h\|_{C^0} = \sup_{\mathbb{R}^d}|h|$), so we get $\|D^\beta f D^\gamma g\|_p \leq \|D^\gamma g\|_{C^0}\|D^\beta f\|_p \leq C'\|f\|_{k,p}\|g\|_{k,p}$, confirming (6). □

The following is a standard definition in operator theory in Banach spaces:

Definition 4.15. *A linear operator $K\colon X \to Y$ where $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are Banach spaces is called* compact *if, whenever $\{x_n\}_{n=1}^\infty$ is a sequence in X with $\|x_n\|_X \leq M$ for some fixed $M \in \mathbb{R}$, the sequence $\{Kx_n\}_{n=1}^\infty$ has a convergent subsequence in Y.*

If $\Omega \subset \mathbb{R}^d$ is a bounded open subset, define $W_0^{k,p}(\Omega)$ to be the completion of the space of smooth functions with support compactly contained in Ω with respect to the norm $\|\cdot\|_{k,p}$. Obviously we have an inclusion $W_0^{k,p}(\Omega) \subset W^{k,p}(\mathbb{R}^d)$.

Theorem 4.16. *Let $\Omega \subset \mathbb{R}^d$ be a bounded open subset, let $k \geq 1$ and $p > d$. Then the inclusion $W_0^{k,p}(\Omega) \hookrightarrow W_0^{k-1,p}(\Omega)$ is a compact operator.*

Proof. Let $\{u_n\}_{n=1}^\infty$ be some sequence in $W_0^{k,p}(\Omega)$ with $\|u_n\|_{k,p} \leq M$ for all n. Then whenever $|\beta| \leq k-1$ each $D^\beta u_n$ is continuous and we have uniform estimates $|D^\beta u_n(x)| \leq CM$ and $|D^\beta u_n(x) - D^\beta u_n(y)| \leq CM\|x-y\|^{1-d/p}$. The first of these statements shows that the $D^\beta u_n$ are pointwise bounded, and the second shows that they are uniformly equicontinuous (for any $\epsilon > 0$ there is $\delta > 0$ (independent of n) such that if $|x-y| < \delta$ then $|D^\beta u_n(x) - D^\beta u_n(y)| < \epsilon$). Since the $D^\beta u_n$ all have support in the fixed compact set $\bar{\Omega}$, the Arzela-Ascoli theorem states that these two properties imply that $\{D^\beta u_n\}_{n=1}^\infty$ has a uniformly convergent subsequence, with limit some continuous function $D^\beta u$. (By applying Arzela-Ascoli to the sequence of $\mathbb{R}^N$-valued functions whose coordinates are the various derivatives $D^\beta u_n$ we may arrange that the same subsequence works for each derivative.) Since $\bar{\Omega}$ has finite volume and contains the supports of all functions involved, the uniform convergence to $D^\beta u$ (for $|\beta| \leq k-1$) implies L^p convergence, so our subsequence converges to u in $W_0^{k-1,p}(\Omega)$. □

As mentioned earlier, the situation is different for $p < d$. Here, although a function in $W^{1,p}(\mathbb{R}^d)$ is typically not continuous, it is at least contained in $L^{p^*}(\mathbb{R}^d)$ for a certain value $p^* > p$. The general formula for p^* works out to be $p^* = \frac{dp}{d-p}$. We'll prove this in the case $d = 2$ since the notation in this case is more easily digestible and since it's the only case that we'll need. See [Ev] or [MS2, Appendix B] for the proof for general d (which is conceptually very similar).

Theorem 4.17 (Gagliardo-Nirenberg-Sobolev inequality)**.** *Let $1 \leq p < 2$. There is a constant C, depending only on p, such that if $f \in C_0^\infty(\mathbb{R}^2)$ then*

$$\|f\|_{p^*} \leq C\|\nabla f\|_p,$$

where $p^* = \frac{2p}{2-p}$ Consequently there is a continuous inclusion map $W^{k+1,p}(\mathbb{R}^2) \hookrightarrow W^{k,p^*}(\mathbb{R}^2)$ for any $k \geq 0$.

Proof. Consider first the inequality for $p = 1$. If $f\colon \mathbb{R}^2 \to \mathbb{R}$ is a compactly supported, continuously differentiable function we have, by the Fundamental Theorem of Calculus,

$$|f(x,y)| \leq \int_{-\infty}^{x}\left|\frac{\partial f}{\partial x}(x',y)\right|dx' \leq \int_{-\infty}^{\infty}\left|\frac{\partial f}{\partial x}(x',y)\right|dx' =: v_1(y)$$

and likewise

$$|f(x,y)| \leq \int_{-\infty}^{\infty}\left|\frac{\partial f}{\partial y}(x,y')\right|dy' =: v_2(x).$$

Hence

$$\int_{\mathbb{R}^2} |f(x,y)|^2 dxdy \le \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} v_1(y)v_2(x)dxdy$$

$$= \left(\int_{-\infty}^{\infty} v_1(y)dy\right)\left(\int_{-\infty}^{\infty} v_2(x)dx\right) = \left(\int_{-\infty}^{\infty}\left|\frac{\partial f}{\partial x}\right| dydx\right)\left(\int_{-\infty}^{\infty}\left|\frac{\partial f}{\partial y}\right| dxdy\right) \le \left(\int_{\mathbb{R}^2} \|\nabla f\| dxdy\right)^2.$$

In other words, we have

$$\|f\|_2 \le \|\nabla f\|_1 \tag{7}$$

for f compactly supported and C^1, proving the inequality for $p = 1$.

We now attempt to use (7) to prove an inequality of the shape $\|f\|_{2\gamma} \le \|\nabla f\|_p$ for any given $p < n$, where $\gamma > 1$ is to be determined. In this direction, note first that, if $f \in C_0^\infty(\mathbb{R}^2)$ and $\gamma > 1$, then $|f|^\gamma$ is compactly supported and continuously differentiable with

$$\|\nabla(|f|^\gamma)\| = \gamma|f|^{\gamma-1}\|\nabla f\|$$

(the only places where differentiability might be a problem are those where $f = 0$, but there the derivative is easily seen to exist and be equal to zero because $\gamma > 1$). So we can apply (7) to $|f|^\gamma$ to obtain

$$\left(\int_{\mathbb{R}^2} |f|^{2\gamma} d^2x\right)^{1/2} = \||f|^{2\gamma}\|_2 \le \gamma \int_{\mathbb{R}^2} |f|^{\gamma-1}\|\nabla f\| \le \gamma\||f|^{\gamma-1}\|_q \|\nabla f\|_p \tag{8}$$

where $\frac{1}{p} + \frac{1}{q} = 1$ (*i.e.*, $q = \frac{p}{p-1}$). Of course

$$\||f|^{\gamma-1}\|_q = \left(\int_{\mathbb{R}^2} |f|^{\frac{(\gamma-1)p}{p-1}} d^2x\right)^{1-\frac{1}{p}}.$$

The appropriate choice of γ is then the one that causes $\frac{(\gamma-1)p}{p-1} = 2\gamma$, *i.e.*, $\gamma = \frac{p}{2-p}$. Then where $p^* = 2\gamma = \frac{2p}{2-p}$ (8) becomes

$$\left(\int_{\mathbb{R}^2} |f|^{p^*} dxdy\right)^{\frac{1}{p}-\frac{1}{2}} \le \gamma\|\nabla f\|_p.$$

Clearly $p^* = \frac{2p}{2-p}$ gives $\frac{1}{p^*} = \frac{1}{p} - \frac{1}{2}$, so we have proven the inequality $\|f\|_{p^*} \le C\|\nabla f\|_p$.

The last sentence of the theorem is an immediate consequence, since the inequality shows that if $\{f_n\}_{n=1}^\infty$ is a sequence of compactly supported smooth functions which is Cauchy with respect to the norm $\|\cdot\|_{k+1,p}$, so that whenever $|\alpha| \le k$ $\{D^\alpha f_n\}_{n=1}^\infty$ is Cauchy with respect to $\|\cdot\|_{1,p}$, then $\{D^\alpha f_n\}_{n=1}^\infty$ is also Cauchy with respect to $\|\cdot\|_{p^*}$ whenever $|\alpha| \le k$, whence $\{f_n\}_{n=1}^\infty$ is Cauchy with respect to $\|\cdot\|_{k,p^*}$. □

Corollary 4.18. *If* $1 < p \le 2$ *and if* $f \in W^{k+2,p}(\mathbb{R}^2)$ *then* f *is* k*-times continuously differentiable*

Proof. If $1 < p < 2$ we have an embedding $W^{k+2,p}(\mathbb{R}^2) \hookrightarrow W^{k+1,p^*}$ where $p^* = \frac{2p}{2-p}$. In particular $p^* > 2$. So if $f \in W^{k+2,p}(\mathbb{R}^2)$ then $f \in W^{k+1,p^*}(\mathbb{R}^2)$ and so f is k-times continuously differentiable by Corollary 4.13. If $p = 2$, note that since differentiability is a local property $f \in W^{k+2,2}(\mathbb{R}^d)$ is k-times continuously differentiable if and only if χf is k-times continuously differentiable for every compactly supported smooth function χ. Now if $\chi \in C_0^\infty(\mathbb{R}^2)$, then $\chi f \in W^{k+2,2}(\mathbb{R}^2)$ has support in a finite measure set, so $\chi f \in W^{k+2,p}(\mathbb{R}^2)$ for any $p < 2$. So χf is k-times continuously differentiable by what has already been shown. □

The correct generalization of this corollary to higher dimension is that (for $1 < p < \infty$) $f \in W^{k+l,p}(\mathbb{R}^d)$ is k-times continuously differentiable provided that $lp > d$.

In what follows we'll really be interested in functions $u\colon \mathbb{R}^2 \to \mathbb{R}^{2n}$. Such a function can equally well be viewed as a $2n$-tuple of functions $\mathbb{R}^2 \to \mathbb{R}$. As such, writing $\mathbb{W}^{k,p}(\mathbb{R}^d;\mathbb{R}^{2n})$ for the space of $2n$-tuples of functions in $W^{k,p}(\mathbb{R}^d;\mathbb{R})$, with norm $\|(f_1,\ldots,f_{2n})\|_{k,p} = \left(\sum_{i=1}^{2n}\|f_i\|_{k,p}^p\right)^{1/p}$, the above results (in particular Theorem 4.16, Corollary 4.13, and Theorem 4.17) extend trivially from $W^{k,p}(\mathbb{R}^d)$ to $W^{k,p}(\mathbb{R}^d;\mathbb{R}^{2n})$ (though perhaps with different constants).

5. The linear Cauchy–Riemann operator

Let $f\colon U \to \mathbb{C}$ be a (suitably-differentiable) complex-valued function on an open subset $U \subset \mathbb{C}$. Where a general element of $\mathbb{C}$ is written $z = x + iy$, recall the notation

$$\partial_{\bar{z}} f = \frac{\partial f}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}\right), \quad \partial_z f = \frac{\partial f}{\partial z} = \frac{1}{2}\left(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y}\right).$$

$\partial_{\bar{z}}, \partial_z$ are derivations on the space of complex-valued smooth functions and as such can be thought of as sections of the complexified tangent bundle $TU \otimes_{\mathbb{R}} \mathbb{C}$ (*i.e.* as complexified vector fields). Where $dz = dx + idy$ and $d\bar{z} = dx - idy$ form the standard basis of complexified 1-forms on U (*i.e.* sections of $T^*U \otimes_{\mathbb{R}} \mathbb{C}$), the definitions have been chosen so that at each point $p \in U$ $\{\partial_z, \partial_{\bar{z}}\} \subset T_pU \otimes_{\mathbb{R}} \mathbb{C}$ is a dual basis to the basis $\{dz, d\bar{z}\}$ of $T^*_pU \otimes_{\mathbb{R}} \mathbb{C}$.

Partly as a result of this, one has

$$df = \partial_z f dz + \partial_{\bar{z}} f d\bar{z}$$

if f is differentiable. Correspondingly, one has

$$f(z') - f(z) = \partial_z f(z)(z' - z) + \partial_{\bar{z}} f(z)(\bar{z}' - \bar{z}) + o(|z' - z|)$$

as $z' \to z$ provided that f is C^1, just as in the usual mutlivariable Taylor formula.

The *Cauchy-Riemann equation* (i.e. the equation that f must satisfy in order to be holomorphic with respect to the standard complex structure on $\mathbb{C}$) can then be written

$$\partial_{\bar{z}} f = 0.$$

This section is devoted to properties of solutions to the equation

$$\partial_{\bar{z}} f = g$$

which we will later leverage to learn more about solutions to the nonlinear Cauchy-Riemann equation that we encounter in the theory of J-holomorphic curves.

First we provide the following generalization of the Cauchy integral formula:

Theorem 5.1. *Let $f\colon \bar{U} \to \mathbb{C}$ be a continuous function which is continuously differentiable on the bounded open set $U \subset \mathbb{C}$ with smooth boundary ∂U. Then, for $z \in U$,*

$$f(z) = \frac{1}{2\pi i}\left(\int_{\partial U} \frac{f(w)}{w - z} dw + \int_U \frac{\partial_{\bar{z}} f}{w - z} dw \wedge d\bar{w}\right).$$

Note that $dw \wedge d\bar{w} = -2idx \wedge dy$ is a constant multiple of the standard volume form. Obviously the above reduces to the Cauchy integral formula when f is holomorphic.

Proof. Let $\epsilon > 0$ be any number which is small enough that the disc $B_\epsilon(z)$ of radius ϵ around z is contained in U. Let us apply Stokes' theorem to the 1-form $\alpha(w) = \frac{f(w)}{w-z} dw$ on the region $U \setminus B_\epsilon(z)$ (clearly α is continuously differentiable on this region). We have

$$\int_{U \setminus B_\epsilon(z)} d\alpha = \int_{\partial(U \setminus B_\epsilon(z))} \alpha = \int_{\partial U} \frac{f(w)}{w - z} dw - \int_{\partial B_\epsilon(z)} \frac{f(w)}{w - z} dw. \tag{9}$$

We have (using the product rule and the fact that $\frac{1}{w-z}$ is holomorphic in w where it is defined) $d\alpha = -\frac{\partial_{\bar{z}} f(w)}{w-z} dw \wedge d\bar{w}$. Choosing M large enough that $|\partial_{\bar{z}} f(w)| \leq M$ for all w in some compact set with nonempty interior containing z, for ϵ small enough we have

$$\begin{aligned}\left|\int_{B_\epsilon(z)} d\alpha\right| &= \left|\int_{B_\epsilon(z)} \frac{\partial_{\bar{z}} f(w)}{z - w} dw \wedge d\bar{w}\right| \\ &= \left|-2i \int_0^{2\pi} \int_0^\epsilon \frac{\partial_{\bar{z}} f(z + re^{i\theta})}{re^{i\theta}} r dr d\theta\right| \leq \int_0^{2\pi} \int_0^\epsilon M dr d\theta \to 0\end{aligned}$$

as $\epsilon \to 0$ (of course in the process of the above we've shown that $d\alpha$ is integrable over $B_\epsilon(z)$, though a reader who remembers the proof of Morrey's inequality would not have found this surprising given the form of $d\alpha$). Consequently as $\epsilon \to 0$ the left hand side of (9) tends to $-\int_U \frac{\partial_{\bar{z}} f(w)}{z-w} dw \wedge d\bar{w}$ (which in particular is well-defined).

Meanwhile for small ϵ,

$$\int_{\partial B_\epsilon(z)} \frac{f(w)}{w-z} dw = \int_0^{2\pi} \frac{f(z+\epsilon e^{i\theta})}{\epsilon e^{i\theta}} i\epsilon e^{i\theta} d\theta \to 2\pi i f(z)$$

by the continuity of f. Thus sending $\epsilon \to 0$ in (9) gives

$$-\int_U \frac{\partial_{\bar{z}} f(w)}{w-z} dw \wedge d\bar{w} = \int_{\partial U} \frac{f(w)}{z-w} dw - 2\pi i f(z),$$

which upon rearrangement gives the theorem. □

Corollary 5.2. *If $f\colon \mathbb{C} \to \mathbb{C}$ is a compactly supported C^1 function then*

$$f(z) = \frac{1}{2\pi i} \int_{\mathbb{C}} \frac{\partial_{\bar{z}} f(w)}{w-z} dw \wedge d\bar{w}.$$

Proof. Apply the theorem above with U equal to a disc around the origin of sufficiently large radius that it contains the support of f in its interior. □

Write $C_0^\infty(\mathbb{C};\mathbb{C})$ for the space of compactly supported smooth functions from $\mathbb{C}$ to itself.

Theorem 5.3. *If $f \in C_0^\infty(\mathbb{C};\mathbb{C})$ define*

$$Pf(z) = \int_{\mathbb{C}} \frac{f(w)}{w-z} dw \wedge d\bar{w}.$$

Then $Pf\colon \mathbb{C} \to \mathbb{C}$ is continuously differentiable, and we have

$$\partial_{\bar{z}}(Pf) = P(\partial_{\bar{z}} f) = f \text{ and } \partial_z(Pf) = P(\partial_z f).$$

Note that this theorem does not guarantee that Pf is compactly supported, and indeed one can check via examples that it typically is not.

Proof. Write $\Phi(w) = -\frac{1}{2\pi i w}$, so that by definition

$$Pf(z) = \int_{\mathbb{C}} \Phi(z-w) f(w) dw \wedge d\bar{w}.$$

A change of variables to $v = z - w$ shows that, similarly,

$$Pf(z) = \int_{\mathbb{C}} f(z-v)\Phi(v) dv \wedge d\bar{v}.$$

Hence

$$Pf(z+h) - Pf(z) = \int_{\mathbb{C}} (f(z-v+h) - f(z-v))\Phi(v) dv \wedge d\bar{v}. \tag{10}$$

Since f (and hence all of its derivatives) is compactly supported there is a number M such that, for any $w \in \mathbb{C}$ and any $h \in B_1(0)$ we have

$$|f(w+h) - f(w) - \partial_z f(w) h - \partial_{\bar{z}} f(w)\bar{h}| \le M|h|^2.$$

Moreover, assuming the support of f to be contained in $B_R(0)$, when $|h| \le 1$ the integrand on the right hand side of (10) is zero outside of $\{v : |v-z| \le R+1\}$. Hence for $|h| \le 1$ the integral on the right hand side of (10) differs from

$$\int_{\mathbb{C}} (\partial_z f(z-v) h + \partial_{\bar{z}} f(z-v)\bar{h})\Phi(v) dv \wedge d\bar{v} \tag{11}$$

by something of modulus at most

$$\int_{|v-z|<R+1} M|h|^2\Phi(v)dw\wedge d\bar{w}\le M'|h|^2$$

for a suitable (f-dependent) constant M'. But (11) above is (using again the identity $\int\Phi(z-w)g(w)=\int g(z-v)\Phi(v)$) just equal to

$$P\partial_z f(z)h + P\partial_{\bar{z}}f(z)\bar{h}.$$

Thus we have, for $|h|\le 1$,

$$Pf(z+h)-Pf(z)=P\partial_z f(z)h+P\partial_{\bar{z}}f(z)\bar{h}+O(|h|^2).$$

This in particular proves that Pf is continuous; applying the same argument to the compactly supported smooth functions $\partial_z f$ and $\partial_{\bar{z}}f$ shows that $P\partial_z f$ and $P\partial_{\bar{z}}f$ are also continuous. Thus Pf is continuously differentiable with derivatives $\partial_z Pf = P\partial_z f$ and $\partial_{\bar{z}}Pf = P\partial_{\bar{z}}f$. The fact that $P\partial_{\bar{z}}f = f$ is Corollary 5.2. □

Proposition 5.4. *Fix a bounded open set $\Omega\subset\mathbb{C}$. Then there is a constant M_Ω such that if $f\in L^p_0(\Omega;\mathbb{C})$ with $p>1$ then*

$$\|Pf\|_{L^p(\Omega;\mathbb{C})}:=\left(\int_\Omega |Pf|^p\right)^{1/p}\le M_\Omega\|f\|_p.$$

This is nearly a special case of Young's inequality, which states that $\|g*f\|_p\le\|g\|_1\|f\|_p$ for functions $f,g\colon\mathbb{R}^d\to\mathbb{R}$. Here we have $Pf=\Phi*f$ where Φ is integrable over Ω (whereas it's not integrable over $\mathbb{C}$, which accounts for the need for an Ω-dependent constant).

Proof. Let $\frac{1}{p}+\frac{1}{q}=1$. Note that if Ω is contained in a ball of radius R then for any $x\in\Omega$ we have

$$\int_\Omega|\Phi(x-y)|d^2y\le\int_{B_{2R(0)}}|\Phi(v)|d^2v.$$

Write $M_\Omega=\int_{B_{2R(0)}}|\Phi(v)|d^2v$.

We have

$$\begin{aligned}|Pf(x)|&\le\int_\Omega|\Phi(x-y)||f(y)|d^2y=\int_\Omega|f(y)||\Phi(x-y)|^{1/p}|\Phi(x-y)|^{1/q}d^2y\\&\le\left(\int_\Omega|f(y)|^p|\Phi(x-y)|d^2y\right)^{1/p}\left(\int_\Omega|\Phi(x-y)|d^2y\right)^{1/q}\le M_\Omega^{1/q}\left(\int_\Omega|f(y)|^p|\Phi(x-y)|d^2y\right)^{1/p}\end{aligned}$$

Hence

$$\begin{aligned}\int_\Omega|Pf(x)|^pd^2x&\le M_\Omega^{p/q}\int_\Omega\left(\int_\Omega|f(y)|^p|\Phi(x-y)|d^2y\right)d^2x\le M_\Omega^{p/q}\int_\Omega\int_{\{z=x-y|x\in\Omega\}}|f(y)|^p|\Phi(z)|d^2zd^2y\\&=M_\Omega^{1+p/q}\int_\Omega|f(y)|^pd^2y.\end{aligned}$$

Raising both sides to the power $\frac{1}{p}$ then gives the proposition. □

Proposition 5.5. *Define $T\colon C_0^\infty(\mathbb{C};\mathbb{C})\to C(\mathbb{C};\mathbb{C})$ by*

$$T=P\circ\partial_z.$$

Then, for $f\in C_0^\infty(\mathbb{C};\mathbb{C})$ we have

$$\partial_z f=T\partial_{\bar{z}}f=\partial_{\bar{z}}Tf\text{ and }\partial_zTf=T\partial_zf. \tag{12}$$

Furthermore,

$$Tf(z)=\frac{1}{2\pi i}\lim_{\epsilon\to0^+}\int_{|w-z|\ge\epsilon}\frac{f(w)}{(w-z)^2}dw\wedge d\bar{w} \tag{13}$$

for all z (in particular the limit on the right hand side exists everywhere[4]).

Proof. If $f \in C_0^\infty(\mathbb{C};\mathbb{C})$, repeated application of Theorem 5.3 gives

$$\partial_z f = \partial_z \circ P \circ \partial_{\bar{z}} f = (P \circ \partial_z) \circ \partial_{\bar{z}} f = T\partial_{\bar{z}} f,$$

where the first equation is justified since $\partial_z f \in C_0^\infty(\mathbb{C};\mathbb{C})$ and the second is justified since $\partial_{\bar{z}} f \in C_0^\infty(\mathbb{C};\mathbb{C})$. Also,

$$\partial_{\bar{z}} T f = \partial_{\bar{z}}\partial_z P f = \partial_z \partial_{\bar{z}} P f = \partial_z f,$$

and

$$\partial_z T f = \partial_z P \partial_z f = P\partial_z \partial_z f = T\partial_z f.$$

This proves (12).

As for the second, note first that

$$\left| \int_{B_\epsilon(z)} \frac{\partial_z f}{w-z} dw \wedge d\bar{w} \right| \le 2\|\partial_z f\|_\infty \int_0^{2\pi} \int_0^{\epsilon} \frac{r dr d\theta}{re^{i\theta}} \to 0$$

as $\epsilon \to 0$. Consequently

$$Tf(z) = \frac{1}{2\pi i} \lim_{\epsilon\to 0} \int_{|w-z|\ge\epsilon} \frac{\partial_z f}{w-z} dw \wedge d\bar{w} = \frac{1}{2\pi i} \lim_{\epsilon\to 0} \int_{|w-z|\ge\epsilon} \frac{1}{w-z} d\left(f(w) d\bar{w}\right). \tag{14}$$

Now

$$\begin{aligned}\int_{|w-z|\ge\epsilon} \frac{1}{w-z} d\left(f(w)d\bar{w}\right) &= \int_{|w-z|\ge\epsilon} d\left(\frac{f(w)d\bar{w}}{w-z}\right) - \int_{|w-z|\ge\epsilon} d\left(\frac{1}{w-z}\right) \wedge f(w)d\bar{w} \\ &= -\int_{|w-z|=\epsilon} \frac{f(w)d\bar{w}}{w-z} + \int_{|w-z|\ge\epsilon} \frac{1}{(w-z)^2} f(w) dw \wedge d\bar{w}.\end{aligned}$$

But the first term above tends to zero as $\epsilon \to 0$: indeed

$$\int_{|w-z|=\epsilon} \frac{f(w)d\bar{w}}{w-z} = \int_0^{2\pi} \frac{f(z+\epsilon e^{i\theta})}{\epsilon e^{i\theta}} \epsilon e^{-i\theta} d\theta = \int_0^{2\pi} f(z+\epsilon e^{i\theta}) e^{-2i\theta} d\theta \to 0$$

by virtue of the continuity of f and the fact that $\int_0^{2\pi} e^{-2i\theta} d\theta = 0$. Hence (14) together with the above calculation imply the result. □

The operator T, with formula given by (13), is an example of a *singular integral operator*, *i.e.* an operator on functions on $\mathbb{R}^d$ given by convolution with an appropriately balanced[5] function which is singular at the origin with absolute value asymptotic to $|x|^{-d}$. Since $|x|^{-d}$ fails to be locally integrable on $\mathbb{R}^d$ the behavior of such operators appears subtle; however in [CaZy] Calderón and Zygmund proved a general theorem that in our case can be expressed as follows.

Theorem 5.6. *Let $1 < p < \infty$. Then there is a constant $C > 0$, depending only on p, such that the operator $T\colon C_0^\infty(\mathbb{C};\mathbb{C}) \to C(\mathbb{C};\mathbb{C})$ defined by (13) obeys*

$$\|Tf\|_p \le C\|f\|_p$$

for all $f \in C_0^\infty(\mathbb{C};\mathbb{C})$. Consequently T extends to a bounded linear operator $T\colon L^p(\mathbb{C};\mathbb{C}) \to L^p(\mathbb{C};\mathbb{C})$.

---

[4]Since $v \mapsto v^{-2}$ is not locally integrable on $\mathbb{C}$ the integrand on the right hand side in (13) is not absolutely convergent over all of $\mathbb{C}$, which is why the limiting process is necessary

[5]in the sense that, in particular, the restriction of the function to an arbitrarily small sphere around the origin has mean value zero

Proof of Theorem 5.6 for $p = 2$. When $p = 2$ the result follows from an application of Stokes' theorem (an alternate method for the $p = 2$ case, which more readily adapts to more general singular integral operators, is to use the Fourier transform and Plancherel's theorem). For $p \neq 2$ the proof is harder and is therefore consigned to the next subsection (and won't be covered in class unless there's demand for it).

First observe that, given $f \in C_0^\infty(\mathbb{C};\mathbb{C})$, if $M, R \in \mathbb{R}$ are chosen to have the property that $|f(w)| \leq M$ for all $w \in \mathbb{C}$ and $supp(f) \subset B_R(0)$, then if $|z| > 2R$ we have, for all $w \in supp(f)$, $|z-w| > \frac{|z|}{2}$ and hence

$$|Pf(z)| = \left|\frac{1}{2\pi}\int_{B_R(0)} \frac{f(w)}{w-z} dw \wedge d\bar{w}\right| \leq \frac{1}{\pi}\frac{M\pi R^2}{|z|/2} = \frac{2MR^2}{|z|} \text{ if } |z| > 2R.$$

Thus for $f \in C_0^\infty(\mathbb{C};\mathbb{C})$ there is $N_f \in \mathbb{R}$ such that $|Pf(z)| \leq N_f/|z|$ if $|z|$ is sufficiently large. Since $T = P \circ \partial_z$, by replacing N_f with $\max\{N_f, N_{\partial_z f}\}$ we in fact obtain

$$\max\{|Pf(z)|, |Tf(z)|\} \leq \frac{N_f}{|z|} \text{ for sufficiently large } |z|. \tag{15}$$

Now observe that

$$\begin{aligned}|Tf|^2 dz \wedge d\bar{z} &= (\partial_z Pf)\overline{Tf} dz \wedge d\bar{z} = d(Pf\overline{Tf}d\bar{z}) - Pf\partial_z\overline{Tf}dz \wedge d\bar{z} \\ &= d(Pf\overline{Tf}d\bar{z}) - Pf\overline{\partial_{\bar{z}}Tf}dz \wedge d\bar{z} = d(Pf\overline{Tf}d\bar{z}) - Pf\overline{\partial_z f}dz \wedge d\bar{z} = d(Pf\overline{Tf}d\bar{z}) - Pf\partial_{\bar{z}}\bar{f}dz \wedge d\bar{z} \\ &= d\left(Pf\overline{Tf}d\bar{z} + (Pf)\bar{f}dz\right) + (\partial_{\bar{z}}Pf)f dz \wedge d\bar{z} = d\left(Pf\overline{Tf}d\bar{z} + (Pf)\bar{f}dz\right) + |f|^2 dz \wedge d\bar{z}.\end{aligned}$$

Hence for any $R > 0$ we have

$$\int_{B_R(0)} |Tf|^2 dz \wedge \bar{d}z = \int_{\partial B_R(0)} (Pf\overline{Tf}d\bar{z} + (Pf)\bar{f}dz) + \int_{B_R(0)} |f|^2 dz \wedge d\bar{z}.$$

For R large enough that the support of f is contained in $B_R(0)$ and that (15) applies, the integral over $\partial B_R(0)$ above is bounded in absolute value by $2\pi N_f/R$. So sending $R \to \infty$ gives

$$\int_{\mathbb{C}} |Tf|^2 dz \wedge d\bar{z} = \int_{\mathbb{C}} |f|^2 dz \wedge d\bar{z},$$

so the result follows (for $p = 2$) with constant $C = 1$. □

Theorem 5.6 is the driving force behind many of the properties that we'll prove for pseudoholomorphic curves; before giving its proof in the case $p \neq 2$ let us indicate some consequences.

Corollary 5.7. *For $k \in \mathbb{N}$ and $1 < p < \infty$:*

(i) *The operator $T\colon C_0^\infty(\mathbb{C};\mathbb{C})$ extends to a bounded linear map $T\colon W^{k,p}(\mathbb{C};\mathbb{C}) \to W^{k,p}(\mathbb{C};\mathbb{C})$.*

(ii) *Fix a bounded open set Ω, and write $C_0^\infty(\Omega;\mathbb{C})$ for the space of smooth functions with support contained in Ω, $W_0^{k,p}(\Omega;\mathbb{C})$ for its completion with respect to the norm $\|\cdot\|_{k,p}$, and $W^{k,p}(\Omega;\mathbb{C})$ for the completion with respect to $\|\cdot\|_{k,p}$ of the space of restrictions to Ω of functions in $C_0^\infty(\mathbb{C};\mathbb{C})$. Then there is $C_\Omega > 1$ such that, if $f \in C_0^\infty(\Omega;\mathbb{C})$ we have*

$$\|Pf\|_{W^{k+1,p}(\Omega)} := \sum_{|\alpha|\leq k}\left(\int_\Omega |D^\alpha f|^p\right)^{1/p} \leq C_\Omega \|f\|_{k,p}.$$

Consequently P extends to a bounded operator

$$P\colon W_0^{k,p}(\Omega;\mathbb{C}) \to W^{k+1,p}(\Omega;\mathbb{C}).$$

Proof. Note first that since $T\partial_z = \partial_z T$ and $T\partial_{\bar{z}} = \partial_{\bar{z}} T$ we have $T \circ D^\alpha = D^\alpha \circ T$ for every multi-index α. So for $u \in C_0^\infty(\mathbb{C};\mathbb{C})$, for any $k \geq 0$ and $1 < p < \infty$ we have, where C is the constant from Theorem 5.6,

$$\|Tu\|_{k,p} = \sum_{|\alpha|\leq k} \|D^\alpha Tu\|_p = \sum_{|\alpha|\leq k} \|T(D^\alpha u)\|_p \leq C \sum_{|\alpha|\leq k} \|D^\alpha u\|_p = C\|u\|_{k,p}.$$

This proves (i).

As for (ii), since $\frac{\partial}{\partial x} = \partial_z + \partial_{\bar{z}}$ and $\frac{\partial}{\partial y} = i(\partial_z - \partial_{\bar{z}})$, we have, for $u \in C_0^\infty(\Omega; \mathbb{C})$,

$$\begin{aligned} \|Pu\|_{k+1,p} &\le \|Pu\|_p + 2\left(\|\partial_{\bar{z}} Pu\|_{k,p} + \|\partial_z Pu\|_{k,p}\right) \\ &\le M_\Omega \|u\|_p + 2\left(\|u\|_{k,p} + \|Tu\|_{k,p}\right) \le M_\Omega\|u\|_p + 2(C+1)\|u\|_{k,p} \le (M_\Omega + 2(C+1))\,\|u\|_{k,p}. \end{aligned}$$

□

Corollary 5.8. *If $k \ge 0$ and $1 < p < \infty$ there is a constant C with the following property. For any $u \in W^{k+1,p}(\mathbb{C}; \mathbb{C})$ we have*

$$\|u\|_{k+1,p} \le C\left(\|\partial_{\bar{z}} u\|_{k,p} + \|u\|_p\right). \tag{16}$$

Proof. We have, for $u \in C_0^\infty(\mathbb{C}; \mathbb{C})$

$$\|u\|_{k+1,p} \le \|u\|_p + 2\left(\|\partial_{\bar{z}} u\|_{k,p} + \|\partial_z u\|_{k,p}\right) = \|u\|_p + 2\left(\|\partial_{\bar{z}} u\|_{k,p} + \|T\partial_{\bar{z}} u\|_{k,p}\right) \le \|u\|_p + 2(C+1)\|\partial_{\bar{z}} u\|_{k,p},$$

which proves (16) when $u \in C_0^\infty(\mathbb{C}; \mathbb{C})$.

If we just assume $u \in W^{k+1,p}(\mathbb{C}; \mathbb{C})$, choose a sequence $\{u_n\}_{n=1}^\infty$ such that $u_n \to u$ in $W^{k+1,p}$. We then also have $\partial_{\bar{z}} u_n \to \partial_{\bar{z}} u$ in $W^{k,p}$ and $u_n \to u$ in L^p, so using (16) for u_n and sending $n \to \infty$ yields (16) for u.

□

Corollary 5.9. *Let $1 < p < \infty$, $k \ge 0$ and fix two bounded open subsets $\Omega', \Omega \subset \mathbb{C}$ such that $\overline{\Omega'} \subset \Omega$. Then there is a constant C, depending on k, p, Ω', Ω, such that for any $u \in C^\infty(\mathbb{C}; \mathbb{C})$ we have*

$$\|u\|_{W^{k+1,p}(\Omega';\mathbb{C})} \le C\left(\|\partial_{\bar{z}} u\|_{W^{k,p}(\Omega;\mathbb{C})} + \|u\|_{L^p(\Omega;\mathbb{C})}\right).$$

Proof. We proceed by induction on k; assume that the corollary has been proven for all integers l with $0 \le l < k$ (of course if $k = 0$ this isn't assuming anything, so we don't need to treat the base case separately). Choose a smooth function $\chi\colon \mathbb{C} \to [0,1]$ such that $\chi|_{\Omega'} = 1$ and $supp(\chi)$ is compact and contained in Ω. We can then choose an open set Ω'' so that $supp(\chi) \subset \Omega''$ and $\overline{\Omega''} \subset \Omega$.

If $u \in C^\infty(\mathbb{C}; \mathbb{C})$ then $\chi u \in C_0^\infty(\mathbb{C}; \mathbb{C})$, $\chi u|_{\Omega'} = u|_{\Omega'}$, and χu is supported in Ω''. Clearly

$$\|u\|_{W^{k+1,p}(\Omega';\mathbb{C})} = \|\chi u\|_{W^{k+1,p}(\Omega';\mathbb{C})} \le \|\chi u\|_{k+1,p}.$$

For some constant M depending on the sup norms of the derivatives of χ up to order $k+1$ (and not on u), the product rule $\partial_{\bar{z}}(\chi u) = (\partial_{\bar{z}}\chi)u + \chi(\partial_{\bar{z}} u)$ and the fact that $supp(\chi) \subset \Omega''$ give an estimate

$$\|\partial_{\bar{z}}(\chi u)\|_{k,p} \le M\left(\|u\|_{W^{k,p}(\Omega'';\mathbb{C})} + \|\partial_{\bar{z}} u\|_{W^{k,p}(\Omega'';\mathbb{C})}\right).$$

By the inductive hypothesis (applied to the sets $\Omega'' \subset \Omega$) we have a constant M' with

$$\|u\|_{W^{k,p}(\Omega'';\mathbb{C})} \le M'\left(\|\partial_{\bar{z}} u\|_{W^{k-1,p}(\Omega;\mathbb{C})} + \|u\|_{L^p(\Omega;\mathbb{C})}\right).$$

Combining the above observations with Corollary 5.8 applied to χu yields (for various constants C_i):

$$\begin{aligned} \|u\|_{W^{k+1,p}(\Omega';\mathbb{C})} &\le \|\chi u\|_{k+1,p} \le C_1\left(\|\partial_{\bar{z}}(\chi u)\|_{k,p} + \|\chi u\|_p\right) \\ &\le C_2\left(\|\partial_{\bar{z}} u\|_{W^{k,p}(\Omega'';\mathbb{C})} + \|\partial_{\bar{z}} u\|_{W^{k-1,p}(\Omega;\mathbb{C})} + \|u\|_{L^p(\Omega;\mathbb{C})}\right) \\ &\le C_3\left(\|\partial_{\bar{z}} u\|_{W^{k,p}(\Omega;\mathbb{C})} + \|u\|_{L^p(\Omega;\mathbb{C})}\right), \end{aligned}$$

as desired. □

Theorem 5.10 (Fundamental elliptic estimate for $\partial_{\bar{z}}$)**.** *Let $1 < p < \infty$ and $k \ge 0$. Fix two bounded open subsets $\Omega, \Omega' \subset \mathbb{C}$ with $\overline{\Omega'} \subset \Omega$. Then there is a constant C, depending on Ω and Ω', with the following property. If $u \in L^p(\Omega; \mathbb{C})$ and if the weak derivative $\partial_{\bar{z}} u$ exists and belongs to $W^{k,p}(\Omega; \mathbb{C})$, then $u \in W^{k+1,p}(\Omega'; \mathbb{C})$, and*

$$\|u\|_{W^{k+1,p}(\Omega')} \le C\left(\|\partial_{\bar{z}} u\|_{W^{k,p}(\Omega;\mathbb{C})} + \|u\|_{L^p(\Omega;\mathbb{C})}\right).$$

Proof. As in the previous corollary, we use induction on k. Assume the theorem proven for all integers l with $0 \le l < k$, fix a smooth function $\chi\colon \mathbb{C} \to [0,1]$ with compact support in Ω such that $\chi|_{\Omega'} = 1$, and choose an open subset Ω'' containing $supp(\chi)$ but whose closure is contained in Ω. We assume inductively that $u \in W^{k,p}(\Omega''';\mathbb{C})$ (possibly $k = 0$) whenever Ω''' has closure contained in Ω. Since by assumption $\partial_{\bar{z}}u \in W^{k,p}(\Omega;\mathbb{C})$ we then have $\chi u \in W^{k,p}(\mathbb{C};\mathbb{C})$, with $\partial_{\bar{z}}(\chi u) = \chi\partial_{\bar{z}}u + (\partial_{\bar{z}}\chi)u$ (to see this, set $u_n = \beta_n * u$ and use the fact that $\beta_n * D^\alpha u = D^\alpha(\beta_n * u)$ (resp. $\beta_n * \partial_{\bar{z}}u = \partial_{\bar{z}}(\beta_n * u)$) when $D^\alpha u$ (resp. $\partial_{\bar{z}}u$) exists to deduce the appropriate weak derivative relationships for u by taking limits of the corresponding relationships for u_n; details are left to the reader). Thus

$$\|\partial_{\bar{z}}(\chi u)\|_{k,p} \le A\left(\|\partial_{\bar{z}}u\|_{W^{k,p}(\Omega'';\mathbb{C})} + \|u\|_{W^{k,p}(\Omega'';\mathbb{C})}\right), \tag{17}$$

where the constant A depends on χ (and thus on Ω, Ω').

By the choice of Ω'', the support of $\beta_n * (\chi u)$ is supported in Ω'' for all sufficiently large n. Furthermore the same argument as that given in the proof of Lemma 4.7 shows that $\beta_n * (\partial_{\bar{z}}(\chi u)) = \partial_{\bar{z}}(\beta_n * (\chi u))$. Write $g_n = \beta_n * (\chi u)$. Since $\partial_{\bar{z}}(\chi u) \in W^{k,p}(\mathbb{C};\mathbb{C})$ and $\partial_{\bar{z}}g_n = \beta_n * (\partial_{\bar{z}}(\chi u))$ Corollary 4.8 shows that $\partial_{\bar{z}}g_n \to \partial_{\bar{z}}(\chi u)$ in $W^{k,p}$. Also, $g_n \to \chi u$ in L^p. So since the g_n are smooth functions with (for n sufficiently large) support in Ω'' and with $\partial_{\bar{z}}g_n$ Cauchy in $W^{k,p}$ and g_n Cauchy in L^p, Corollary 5.9 shows that $\{g_n\}_{n=1}^\infty$ is Cauchy in $W^{k+1,p}(\Omega'';\mathbb{C})$. Hence the L^p-limit of g_n, namely χu, belongs to $W^{k+1,p}(\Omega'';\mathbb{C})$. Hence since $\chi|_{\Omega'} = 1$ we have $u \in W^{k+1,p}(\Omega';\mathbb{C})$. Furthermore, the estimate in Corollary 5.9 for the g_n is easily seen to imply the desired estimate for u by sending $n \to \infty$ to deduce the result for χu and then appealing to (17) and to the inductive hypothesis.

□

Corollary 5.11. *Suppose that $u \in L^p(\mathbb{C};\mathbb{C})$ for some $1 < p < \infty$ and that the weak derivative $\partial_{\bar{z}}u$ exists and belongs to $C^\infty(\mathbb{C};\mathbb{C}) \cap L^p(\mathbb{C};\mathbb{C})$. Then $u \in C^\infty(\mathbb{C};\mathbb{C})$.*

Proof. The preceding theorem shows that, given any bounded open set Ω, $u \in W^{k+1,p}(\Omega;\mathbb{C})$ for every integer k. Since functions in $W^{k+1,p}(\Omega;\mathbb{C})$ are k-times differentiable on Ω if $p > 2$ and $(k-1)$-times differentiable if $1 < p \le 2$ the result is immediate.

□

5.1. Proof of the Calderón–Zygmund Theorem for $p \neq 2$.

Exercise 5.12. Assuming that Theorem 5.6 holds for $p > 2$, prove it for $1 < p < 2$ by using the duality between L^q and L^p where $\frac{1}{p} + \frac{1}{q} = 1$.

Since we've proven Theorem 5.6 for $p = 2$ (in fact we've proven the equality $\|Tg\|_2 = \|g\|_2$), the preceding exercise shows that its proof will be complete once we handle the case $p > 2$. We'll follow a recent paper of F. Yao [Y]; Calderón and Zygmund's original proof, by contrast, went by proving the case $1 < p < 2$ first and then using duality as in the exercise to deduce the result for $p > 2$. In all proofs that I know of, the case $p = 2$ (already proven) is itself directly used in the proof for $p \neq 2$.

If $B \subset \mathbb{C}$ is a measurable subset (with respect to the usual 2-dimensional Lebesgue measure on $\mathbb{C} = \mathbb{R}^2$) denote by $|B|$ the Lebesgue measure of B. Also, if $g\colon \mathbb{C} \to \mathbb{C}$ is a function and $t \in \mathbb{R}$ we use the shorthand $\{|g| > t\} = \{z \in \mathbb{C} | |g(z)| > t\}$. Notation for the Lebesgue measure on $\mathbb{C}$ will typically be omitted from the various integrals that appear below. Here is the main lemma:

Lemma 5.13 ([Y]). *There are constants $C, K > 0$ such that, for any $f \in C_0^\infty(\mathbb{C},\mathbb{C})$ and any numbers $\mu, \delta > 0$ with $\delta < 1$, we have*

$$|\{|Tf| > \mu\}| \le \frac{C}{\mu^2}\left(\delta^2 \int_{\{|Tf| > \mu/K\}} |Tf|^2 + \int_{\{|f| \ge \delta\mu/K\}} |f|^2\right)$$

Proof of Theorem 5.6 ($p > 2$), assuming Lemma 5.13. The proof uses the following general formulas, for a measurable function $g\colon \mathbb{C} \to \mathbb{C}$ and $p > 2$:

$$\int_{\mathbb{C}} |g|^p = p \int_0^\infty \nu^{p-1} |\{|g| > \nu\}| d\nu$$
$$= (p-2) \int_0^\infty \nu^{p-3} \left(\int_{\{|g|>\nu\}} |g|^2 \right) d\nu.$$

If these formulas are unfamiliar, it's easy to prove them: first prove them for g equal to a simple function $\sum_{i=1}^m a_i \chi_{A_i}$, and then use the fact that an arbitrary measurable function has a sequence of simple functions converging monotonically to it and appeal to the monotone convergence theorem.

Given this and Lemma 5.13, we have

$$\begin{aligned}
\int_{\mathbb{C}} |Tf|^p &= p \int_0^\infty \mu^{p-1} |\{|Tf| > \mu\}| d\mu \\
&\le Cp \int_0^\infty \mu^{p-3} \left(\delta^2 \int_{\{|Tf|>\mu/K\}} |Tf|^2 + \int_{\{|f| \ge \delta\mu/K\}} |f|^2 \right) d\mu \\
&= CpK^{p-2} \left(\delta^2 \int_0^\infty \nu^{p-3} \int_{\{|Tf|>\nu\}} |Tf|^2 d\nu + \delta^{2-p} \int_0^\infty \nu^{p-3} \int_{\{|f|>\nu\}} |f|^2 d\nu \right) \\
&= C'\delta^2 \int_{\mathbb{C}} |Tf|^p + C'\delta^{2-p} \int_{\mathbb{C}} |f|^p,
\end{aligned}$$

where the constant $C' > 1$ depends on p but not on δ. Now choose for δ the value that causes $C'\delta^2 = 1/2$. Rearranging the above equation then gives

$$\frac{1}{2} \int_{\mathbb{C}} |Tf|^p \le C'\delta^{2-p} \int_{\mathbb{C}} |f|^p;$$

thus for a suitable constant C'' depending only on p we have

$$\|Tf\|_p \le C''\|f\|_p.$$

□

The rest of this subsection will be devoted to the proof of Lemma 5.13. Choose positive numbers $\delta, \lambda > 0$ with $\delta < 1$. (λ will be a certain constant times the number μ in the hypothesis of the lemma.) Define

$$E_\lambda = \{|Tf| > \lambda\}.$$

If B is any bounded open set, define

$$J[B] = \fint_B \left(|Tf|^2 + \delta^{-2}|f|^2 \right).$$

Lemma 5.14. *Given λ, δ, there is a family $\{B_{\rho_i}(x_i)\}_{i\in S}$ (where $S \subset \mathbb{N}$) of pairwise disjoint balls such that*

$$E_\lambda \subset \cup_{i\in S} B_{5\rho_i}(x_i), \quad J[B_{\rho_i}(x_i)] = \lambda^2, \quad (\forall \rho > \rho_i)(J[B_\rho(x_i)] < \lambda^2),$$

and, for any $\theta > 1$,

$$\sum_{i\in S} \int_{B_{\theta\rho_i}(x_i)} |f|^2 \le 2\theta^2 \left(\delta^2 \int_{\{|Tf|>\lambda/2\}} |Tf|^2 + \int_{\{|f|>\delta\lambda/2\}} |f|^2 \right). \tag{18}$$

Proof. For any $x \in E_\lambda$, since $|Tf(x)| > \lambda$ (and since we assume $f \in C_0^\infty(\mathbb{C}; \mathbb{C})$ so that Tf is continuous[6]) we will have $J[B_r(x)] > \lambda^2$ when r is sufficiently small (depending on x). On the other hand since by the $p = 2$ case of the theorem $Tf \in L^2$ we will have $J[B] < \lambda^2$ whenever B is any ball of radius larger than some number

[6]For more general f, we could apply the Lebesgue differentiation theorem and arrive at the same conclusion up to the removal of a measure-zero set from E_λ

$R > 0$. Hence for every $x \in E_\lambda$, the number $r_x = \sup\{r | J[B_r(x)] \geq \lambda^2\}$ will be finite and positive and will obey $J[B_{r_x}(x)] = \lambda^2$. Further we have $r_x \leq R$ for all x. Define

$$\mathcal{B} = \{B_{r_x}(x) | x \in E_\lambda\}.$$

Now recall the following standard lemma (whose proof is a good exercise):

Lemma 5.15 (Vitali Covering Lemma)**.** *Let $\mathcal{B}$ be any collection of balls of positive radius in a metric space, with the property that for some $R \in \mathbb{R}$ each ball in $\mathcal{B}$ has radius at most R. Then $\mathcal{B}$ has a subcollection $\{B_{\rho_i}(x_i) | i \in S\}$ where $S \subset \mathbb{N}$ whose members are pairwise disjoint and such that*

$$\cup_{B \in \mathcal{B}} B \subset \cup_{i \in S} B_{5\rho_i}(x_i).$$

Accordingly let $\{B_{\rho_i}(x_i)\}_{i \in S}$ be the subcollection of $\mathcal{B}$ supplied by the Vitali Covering Lemma. Since E_λ is contained in the union of the balls in $\mathcal{B}$, these balls clearly satisfy all the conclusions of the lemma except possibly (18), and we now show that (18) holds as well.

In this direction, for any i we have

$$\begin{aligned} \int_{B_{\theta\rho_i}(x_i)} |f|^2 &\leq |B_{\theta\rho_i}(x_i)| \fint_{B_{\theta\rho_i}(x_i)} |f|^2 \leq \delta^2 |B_{\theta\rho_i}(x_i)| J[B_{\theta\rho_i}(x_i)] \\ &< \delta^2 |B_{\theta\rho_i}(x_i)| J[B_{\rho_i}(x_i)] = \delta^2\theta^2 |B_{\rho_i}(x_i)| J[B_{\rho_i}(x_i)]. \end{aligned} \tag{19}$$

Now

$$\begin{aligned} |B_{\rho_i}(x_i)| J[B_{\rho_i}(x_i)] &= \int_{B_{\rho_i}(x_i)} \left(|Tf|^2 + \delta^{-2}|f|^2\right) \\ &\leq \int_{B_{\rho_i}(x_i)} \left(\max\{\lambda/2, |Tf|\}^2 + \delta^{-2} \max\{\delta\lambda/2, |f|\}^2\right) \\ &\leq \frac{\lambda^2}{4}|B_{\rho_i}(x_i)| + \int_{B_{\rho_i}(x_i) \cap \{|Tf| > \lambda/2\}} |Tf|^2 + \frac{\lambda^2}{4}|B_{\rho_i}(x_i)| + \delta^{-2} \int_{B_{\rho_i}(x_i) \cap \{|f| > \delta\lambda/2\}} |f|^2 \\ &= \frac{1}{2}|B_{\rho_i}(x_i)| J[B_{\rho_i}(x_i)] + \int_{B_{\rho_i}(x_i) \cap \{|Tf| > \lambda/2\}} |Tf|^2 + \delta^{-2} \int_{B_{\rho_i}(x_i) \cap \{|f| > \delta\lambda/2\}} |f|^2 \end{aligned}$$

(where the last line follows from the fact that $J[B_{\rho_i}(x_i)] = \lambda^2$). Rearranging and substituting into (19) then gives

$$\int_{B_{\theta\rho_i}(x_i)} |f|^2 \leq 2\delta^2\theta^2 \left(\int_{B_{\rho_i}(x_i) \cap \{|Tf| > \lambda/2\}} |Tf|^2 + \delta^{-2} \int_{B_{\rho_i}(x_i) \cap \{|f| > \delta\lambda/2\}} |f|^2 \right).$$

Summing over i and using the fact that the various $B_{\rho_i}(x_i)$ are pairwise disjoint then proves (18). □

Now, for each $i \in S$ write

$$f = g_{\lambda,i} + h_{\lambda,i} \text{ with } g_{\lambda,i}(x) = \begin{cases} f(x) & x \in B_{25\rho_i}(x_i) \\ 0 & otherwise \end{cases}$$

In particular $g_{\lambda,i}, h_{\lambda,i} \in L^2(\mathbb{C}; \mathbb{C})$, so $Tg_{\lambda,i}$ and $Th_{\lambda,i}$ are well-defined as L^2 functions. Also, as one can see by approximating $h_{\lambda,i}$ by smooth functions and using the L^2-boundedness of T, $Th_{\lambda,i}$ can and will be chosen within its almost-everywhere-equivalence class so that the following holds: if $x \in B_{25\rho_i}(x_i)$ (so in particular $x \notin supp(h_{\lambda,i})$), then the integral $\int_{\mathbb{C}} \frac{h_{\lambda,i}(w)}{2\pi i(x-w)^2} dw \wedge d\bar{w}$ exists and is equal to $Th_{\lambda,i}(x)$.

Lemma 5.16. *There is a constant $M > 1$, independent of f, λ, and i, such that if $x \in B_{5\rho_i}(x_i)$ then $|Th_{\lambda,i}(x)| \leq M\lambda$.*

Proof. First we will bound $|Th_{\lambda,i}(z) - Th_{\lambda,i}(x)|$ for $x, z \in B_{5\rho_i}(x_i)$. Suppose $w \in \mathbb{C}$ with $h_{\lambda,i}(w) \neq 0$. By definition $w \notin B_{25\rho_i}(x_i)$, so

$$w \in B_{5^{m+1}\rho_i}(x_i) \setminus B_{5^m\rho_i}(x_i) \text{ for some } m \geq 2.$$

Then $|x-z| < 10\rho_i$ and $(5^m - 5)\rho_i \le |w-x|, |w-z| \le (5^{m+1}+5)\rho_i$, so

$$\left|\frac{1}{(x-w)^2} - \frac{1}{(z-w)^2}\right| = \left|\frac{(x-z)(2w-x-z)}{(x-w)^2(z-w)^2}\right| \le \frac{20(5^{m+1}+5)\rho_i^2}{(5^m-5)^4\rho_i^4} \le \frac{300}{5^{3m}\rho_i^2}.$$

So, for some constant C,

$$\begin{aligned}|Th_{\lambda,i}(z) - Th_{\lambda,i}(x)| &= \left|\frac{1}{2\pi i}\int_{\mathbb{C}} h_{\lambda,i}(w)\left(\frac{1}{(x-w)^2} - \frac{1}{(z-w)^2}\right) dw \wedge d\bar{w}\right| \le \sum_{m=2}^{\infty} \frac{C}{5^{3m}\rho_i^2}\int_{B_{5^{m+1}\rho_i}(x_i)\setminus B_{5^m\rho_i}(x_i)} |f| \\ &\le \sum_{m=2}^{\infty} \frac{C}{5^{3m}\rho_i^2}\int_{B_{5^{m+1}\rho_i}(x_i)} |f| \le \sum_{m=2}^{\infty} \frac{C}{5^{3m}\rho_i^2}|B_{5^{m+1}\rho_i}(x_i)|\left(⨍_{B_{5^{m+1}\rho_i}(x_i)} |f|^2\right)^{1/2} \\ &= \sum_{m=2}^{\infty} \frac{25\pi C}{5^m}\left(⨍_{B_{5^{m+1}\rho_i}(x_i)} |f|^2\right)^{1/2} \le \sum_{m=2}^{\infty} \frac{25\pi C}{5^m}\delta\lambda \\ &\le C'\delta\lambda\end{aligned}$$

where we've used the Schwarz inequality and (in the penultimate inequality) the fact that $⨍_{B_{5^{m+1}\rho_i}(x_i)} |f|^2 \le \delta^2 J[B_{5^{m+1}\rho_i}(x_i)] \le \delta^2\lambda^2$.

To use this to get the desired pointwise bound, note that

$$|Th_{\lambda,i}(x)| \le ⨍_{B_{5\rho_i}(x_i)} |Th_{\lambda,i}(z)| d^2z + ⨍_{B_{5\rho_i}(x_i)} |Th_{\lambda,i}(z) - Th_{\lambda,i}(x)| d^2z.$$

We have $|Th_{\lambda,i}(z) - Th_{\lambda,i}(x)| \le C'\delta\lambda < C'\lambda$ since $\delta < 1$, giving a bound of $C'\lambda$ for the second term. So it remains only to bound $⨍_{B_{5\rho_i}(x_i)} |Th_{\lambda,i}(z)| d^2z$.

For this, we note

$$⨍_{B_{5\rho_i}(x_i)} |Th_{\lambda,i}(z)| d^2z \le ⨍_{B_{5\rho_i}(x_i)} |Tf(z)| d^2z + ⨍_{B_{5\rho_i}(x_i)} |Tg_{\lambda,i}(z)| d^2z,$$

and we have

$$⨍_{B_{5\rho_i}(x_i)} |Tf| \le \left(⨍_{B_{5\rho_i}(x_i)} |Tf|^2\right)^{1/2} \le J[B_{5\rho_i}(x_i)]^{1/2} < \lambda,$$

while

$$\begin{aligned}⨍_{B_{5\rho_i}(x_i)} |Tg_{\lambda,i}| &\le |B_{5\rho_i}(x_i)|^{-1/2}\left(\int_{B_{5\rho_i}(x_i)} |Tg_{\lambda,i}|^2\right)^{1/2} \\ &= |B_{5\rho_i}(x_i)|^{-1/2}\left(\int_{\mathbb{C}} |g_{\lambda,i}|^2\right)^{1/2} = |B_{5\rho_i}(x_i)|^{-1/2}\left(\int_{B_{25\rho_i}(x_i)} |f|^2\right)^{1/2} = 5\left(⨍_{B_{25\rho_i}(x_i)} |f|^2\right)^{1/2} \le 5\lambda\delta.\end{aligned}$$

(Here we've used the fact that by the proof of the $p = 2$ case $\|Tg_{\lambda,i}\|_2 = \|g_{\lambda,i}\|_2$.)

We have now bounded all three terms on the right in the inequality

$$|Th_{\lambda,i}(x)| \le ⨍_{B_{5\rho_i}(x_i)} |Th_{\lambda,i}(z) - Th_{\lambda,i}(x)| d^2z + ⨍_{B_{5\rho_i}(x_i)} |Tg_{\lambda,i}(z)| d^2z + ⨍_{B_{5\rho_i}(x_i)} |Tf(z)| d^2z$$

by constants times λ, from which the lemma follows. □

We can now complete the proof of Lemma 5.13 and hence of Theorem 5.6. Given μ and δ as in Lemma 5.13, choose λ so that $2M\lambda = \mu$, where $M > 1$ is as in Lemma 5.16. We are attempting to bound the measure $|\{|Tf| > \mu\}|$. Since $\mu > \lambda$, we have

$$\{|Tf| > \mu\} \subset E_\lambda \subset \cup_{i\in S} B_{5\rho_i}(x_i).$$

Now if $x \in B_{5\rho_i}(x_i) \cap \{|Tf| > \mu\}$, then since $Th_{\lambda,i}|_{B_{5\rho_i}(x_i)} \leq M\lambda = \mu/2$ we have $|Tg_{\lambda,i}(x)| > \mu/2$. Thus,

$$\left|B_{5\rho_i}(x_i) \cap \{|Tf| > \mu\}\right| \leq \left|B_{5\rho_i}(x_i) \cap \{|Tg_{\lambda,i}| > \mu/2\}\right| \leq \frac{4}{\mu^2}\int_{\mathbb{C}} |Tg_{\lambda,i}|^2 = \frac{4}{\mu^2}\int_{\mathbb{C}} |g_{\lambda,i}|^2 = \frac{4}{\mu^2}\int_{B_{25\rho_i}(x_i)} |f|^2$$

where again we've used the $p = 2$ case of the theorem.

Hence, using (18), we get

$$\begin{aligned} |\{|Tf| > \mu\}| &\leq \sum_{i \in S} \left|B_{5\rho_i}(x_i) \cap \{|Tf| > \mu\}\right| \\ &\leq \frac{4}{\mu^2}\sum_{i\in S}\int_{B_{25\rho_i}(x_i)} |f|^2 \leq \frac{8\cdot 25^2}{\mu^2}\left(\delta^2 \int_{\{|Tf|>\lambda/2\}} |Tf|^2 + \int_{\{|f|>\delta\lambda/2\}} |f|^2\right). \end{aligned}$$

Writing $K = 4M$ (since $2M\lambda = \mu$), this precisely recovers the statement of Lemma 5.13.

6. Local properties of J-holomorphic curves

Our preparation concerning the linear Cauchy–Riemann operator puts us in position to begin to study local properties of pseudoholomorphic curves, which, we recall, are solutions to a nonlinear analogue of the Cauchy–Riemann equation. We assume given a $2n$-dimensional symplectic manifold (M, ω) with an ω-tame almost complex structure J, and we consider maps $u\colon \Sigma \to M$ obeying (at least in a weak sense) $u_* \circ j = J \circ u_*$ where Σ is a surface and j is an almost complex structure on it. For convenience (and for simplicity of the statements of some theorems) we will always assume that the almost complex structure J is C^∞, though methods similar to those below still give results when much less regularity is assumed on J (see [AL, Chapter V] for instance). At least assuming that u is continuous (as we shall), for any given point $z_0 \in int(\Sigma)$ we can study the map locally near z_0 by choosing a complex coordinate $z = s + it$ around z_0 and a coordinate chart around $u(z_0)$ in M; in terms of these coordinates the equation for u to be pseudoholomorphic becomes

$$\frac{\partial u}{\partial s} + J(u(z))\frac{\partial u}{\partial t} = 0 \tag{20}$$

where u is now viewed as a map from a disc around the origin in $\mathbb{C}$ to $\mathbb{C}^n = \mathbb{R}^{2n}$ with $u(0) = \vec{0}$, and J is an almost complex structure on $\mathbb{R}^{2n}$. If the matrix-valued function J were identically equal to multiplication by i, i.e. to $J_0 := \begin{pmatrix} 0 & -Id \\ Id & 0 \end{pmatrix}$, then this would just be the classical Cauchy–Riemann equation for a vector-valued function u. While J will not be identically equal to J_0 (indeed, as mentioned at the start of the course, the Nijenhuis tensor gives an obstruction to choosing coordinates in such a way that $J = J_0$), there is no loss of generality in assuming that J coincides with J_0 at the origin $\vec{0}$ of $\mathbb{R}^{2n}$ with J_0: indeed, an exercise from last semester shows that, because $J(\vec{0})$ is a linear map of $T_{\vec{0}}\mathbb{R}^{2n} = \mathbb{R}^{2n}$ with $J(\vec{0})^2 = -Id$, there is $A \in GL(2n, \mathbb{R})$ such that $AJ(\vec{0})A^{-1} = J_0$, and then composing our initial coordinate chart on M around $u(z_0)$ with the matrix A reduces us to the case where

$$J(\vec{0}) = J_0.$$

We now rewrite the equation (20) in terms of the operators $\partial_{\bar{z}}$ and ∂_z from the previous section. By definition we have

$$\partial_{\bar{z}} = \frac{1}{2}\left(\frac{\partial}{\partial s} + J_0\frac{\partial}{\partial t}\right), \quad \partial_z = \frac{1}{2}\left(\frac{\partial}{\partial s} - J_0\frac{\partial}{\partial t}\right),$$

so

$$\frac{\partial}{\partial s} = \partial_z + \partial_{\bar{z}}, \quad \frac{\partial}{\partial t} = J_0(\partial_z - \partial_{\bar{z}}).$$

So (20) becomes

$$(Id - J(u(z))J_0)\partial_{\bar{z}}u + (Id + J(u(z))J_0)\partial_z u = 0.$$

Recall that we're assuming that u is continuous and that $J(\vec{0}) = J(u(0)) = J_0$, so at $z = 0$ we have $Id + J(u(z))J_0 = 0$ and $Id - J(u(z))J_0 = 2Id$. So after possibly replacing the domain with a smaller disc D around the origin, we may assume that $Id - J(u(z))J_0$ is invertible for all $z \in D$. Then for a suitable open set $V \subset \mathbb{R}^{2n}$ containing $u(D)$, we may define a map $q\colon V \to \mathbb{R}^{2n\times 2n}$ by

$$q(\vec{v}) = (Id - J(\vec{v})J_0)^{-1}(Id + J(\vec{v})J_0).$$

We then have

$$q(u(0)) = 0 \quad \text{and } \partial_{\bar{z}}u + q(u(z))\partial_z u = 0. \tag{21}$$

We've thus shown that, for a J-holomorphic curve $u\colon \Sigma \to M$, near any point $z_0 \in int(\Sigma)$ there are coordinate neighborhoods in terms of which u appears as a map to $\mathbb{R}^{2n}$ obeying (21)[7]. Since $q(u(0)) = 0$ and $q \circ u$ is continuous it is reasonable to view (21), at least on a small enough neighborhood of 0 as to ensure that $q \circ u$ is small, as a small perturbation of the linear Cauchy–Riemann equation. We will see that a variety of the notable local properties of holomorphic functions (i.e. of solutions to the linear Cauchy–Riemann equation) extend to J-holomorphic curves (i.e. solutions to these perturbed Cauchy–Riemann equations).

6.1. **Smoothness.** One of the more striking results that one encounters in a basic complex analysis course is the fact that any function which is holomorphic (initially a condition just on certain of its first partial derivatives) is in fact C^∞. In this section we will prove that the same holds for J-holomorphic curves:

Theorem 6.1. *Let J be a C^∞ almost complex structure on a smooth manifold M, and let $u\colon \Sigma \to M$ be a map from a surface Σ with complex structure j to M which is continuous and belongs to $W^{1,p}_{loc}(\Sigma, M)$ for some $p > 1$. Assume that u obeys the Cauchy–Riemann equation $u_* \circ j = J \circ u_*$. Then u is C^∞ on the interior of Σ.*

We clarify what it means for a continuous function u to belong to $W^{1,p}_{loc}(\Sigma, M)$: namely, for any $z_0 \in \Sigma$ there should be coordinate neighborhoods $D \subset \mathbb{R}^2$ of z_0 and $V \subset \mathbb{R}^{2n}$ of $u(z_0)$ such that, in terms of these coordinate neighborhoods, $u|_D$ belongs to $W^{1,p}(D, \mathbb{R}^{2n})$, which in turn is defined as the completion of the space of restrictions to D of compactly supported smooth functions $\mathbb{R}^2 \to \mathbb{R}^{2n}$ with respect to the (1,p)-Sobolev norm. (It's a bit trickier, though possible, to make sense of this when u isn't assumed continuous, but we won't need to do this). Assuming that $u \in W^{1,p}_{loc}(\Sigma, M)$, in any sufficiently small coordinate chart it has weak derivatives of order 1, and the Cauchy–Riemann equation $u_* \circ j = J \circ u_*$ may be understood as a condition on these weak derivatives. Of course, once we prove the theorem, we will know that these weak derivatives are genuine derivatives and so the equation will hold in the traditional sense.

Proving the theorem is equivalent to showing that, for any z_0 in the interior of Σ and any k, there is a coordinate neighborhood of z_0 in which u is C^k; if it helps to do so, we may choose this neighborhood to be very small and dependent on k. Thus we can focus in on the small neighborhoods D discussed earlier, in which u solves the equation (21)

$$\partial_{\bar{z}}u + (q \circ u)(z)\partial_z u = 0,$$

where $u(0) = \vec{0}$ and $q(\vec{0}) = 0$. Here q is a $2n \times 2n$-matrix-valued smooth function on an open subset of $\mathbb{R}^{2n}$ that contains $u(D)$; we then can and do extend q to a compactly supported smooth matrix-valued function on all of $\mathbb{R}^{2n}$ which doesn't affect the validity of (21) for $z \in D$.

Our approach will be an example of what is called "bootstrapping;" the point, roughly, is that we can increase the regularity u by appealing to the results of the last section, and then, once we know that u is slightly more regular than it was before, we can repeat the process by appealing to those results again. First we note the following, which is relevant due to the appearance of the term $q \circ u$ in (21):

[7] To be clear about the necessary regularity assumptions here, we require that u be continuous, and in order to make sense of (20) and therefore of (21) we assume that the first-order *weak* derivatives of u exist (when it's written out in local coordinates—one can verify as in Proposition 6.2 below that this notion is independent of the choice of coordinates). Thus it would certainly be enough to assume $u \in C^0 \cap W^{1,p}$ for some $p > 1$, where again membership in $W^{1,p}$ is tested by local coordinate charts.

Proposition 6.2. *Let $q \in C_0^\infty(\mathbb{R}^D;\mathbb{R}^N)$ with $q(\vec{0}) = 0$, and let $k \geq 1$ and $1 < p < \infty$. Then there is a constant C (depending on q) such that if $u\colon \mathbb{R}^d \to \mathbb{R}^D$ is a compactly supported function and $u \in W^{k,p}(\mathbb{R}^d;\mathbb{R}^D) \cap C^{k-1}(\mathbb{R}^d;\mathbb{R}^D)$ then $q \circ u$ is continuous and belongs to $W^{k,p}(\mathbb{R}^d;\mathbb{R}^N)$, with*

$$\|q \circ u\|_{k,p} \leq C(1 + \|u\|_{C^0}) \sum_{l=0}^{k} (\|u\|_{k,p} + \|u\|_{C^{k-1}})^{l+1}.$$

Of course, Corollary 4.13 shows that if $p > d$ (where $d = 2$ in the case of interest) then any function in $W^{k,p}$ automatically belongs to C^{k-1}, with C^{k-1} norm bounded by a constant times the $W^{k,p}$ norm, so the estimate above simplifies to $\|q \circ u\|_{k,p} \leq C \sum_{l=1}^{k+1} \|u\|_{k,p}^l$. The important feature of the right hand side above is that it tends to zero as $\|u\|_{k,p}, \|u\|_{C^{k-1}}$ tend to zero.

The proof relies on the following easy result:

Lemma 6.3. *If $p \in C_0^\infty(\mathbb{R}^D, \mathbb{R}^N)$ and if $\{u_m\}_{m=1}^\infty$ is a sequence of C^k functions which converges in C^k to $u\colon \mathbb{R}^d \to \mathbb{R}^D$, then $q \circ u_m \to q \circ u$ in C^k.*

(Here "convergence in C^k" means that all partial derivatives of order up to and including k converge uniformly.)

Proof of Lemma 6.3. We proceed by induction on k. For $k = 0$, note that p has bounded first-order derivatives and hence is Lipschitz, so we have a constant M with $|p(u_m(x)) - p(u(x))| \leq M|u_m(x) - u(x)|$. Thus since $u_m - u \to 0$ uniformly it follows that $p \circ u_m - p \circ u \to 0$ uniformly.

Now let $k \geq 1$ and assume the lemma for all integers l with $0 \leq l < k$. Let $u_m \to u$ in C^k. Since $u_m \to u$ in C^{k-1}, by the assumed $l = k-1$ case of the lemma to prove that $p \circ u_m \to p \circ u$ in C^k it's enough to show that, whenever $|\alpha| = k-1$, we have $D^\alpha\left(\frac{\partial}{\partial x_i}(p \circ u_m)\right) \to D^\alpha\left(\frac{\partial}{\partial x_i}(p \circ u)\right)$ uniformly for each i. Now where u_m^l denotes the lth component of u_m

$$\begin{aligned} D^\alpha\left(\frac{\partial}{\partial x_i}(p \circ u_m)\right) &= D^\alpha\left(\sum_l \left(\frac{\partial p}{\partial y_l} \circ u\right)\frac{\partial u_m^l}{\partial x_i}\right) \\ &= \sum_l \sum_{\beta+\gamma=\alpha} D^\beta\left(\frac{\partial p}{\partial y_l} \circ u_m\right) D^\gamma \frac{\partial u_m^l}{\partial x_i}. \end{aligned}$$

Recalling that $|\alpha| = k-1$ the inductive hypothesis (applied with p replaced by $\frac{\partial p}{\partial y_l}$) shows that in each of the above terms we have $D^\beta\left(\frac{\partial p}{\partial y_l} \circ u_m\right) \to D^\beta\left(\frac{\partial p}{\partial y_l} \circ u\right)$ uniformly, while $D^\gamma \frac{\partial u_m^l}{\partial x_i} \to D^\gamma \frac{\partial u^l}{\partial x_i}$ uniformly by the assumption that $u_m \to u$ in C^k. So since sums and products of uniformly convergent sequences of bounded functions converge uniformly, it follows that $D^\alpha\left(\frac{\partial}{\partial x_i}(p \circ u_m)\right) \to D^\alpha\left(\frac{\partial}{\partial x_i}(p \circ u)\right)$ uniformly, completing the proof. □

Proof of Proposition 6.2. Of course $q \circ u$ is continuous since it is the composition of two continuous functions. For any multi-index α, since the order-$(|\alpha| + 1)$ derivatives of q are bounded, we have a constant M_α such that

$$|D^\alpha q(\vec{v}_1) - D^\alpha q(\vec{v}_2)| \leq M_\alpha \|\vec{v}_1 - \vec{v}_2\| \tag{22}$$

for $\vec{v}_i \in \mathbb{R}^D$. Applying (22) with α equal to the zero multi-index, $\vec{v}_1 = u(z)$, and $\vec{v}_2 = 0$ shows that (since we assume $q(\vec{0}) = 0$) $|q \circ u(z)| \leq M|u(z)|$ for all z, and therefore that

$$\|q \circ u\|_{C^0} \leq M\|u\|_{C^0} \text{ and } \|q \circ u\|_p \leq M\|u\|_p. \tag{23}$$

Now the mollifications $u_m = \beta_m * u$ converge in $W^{k,p}$ to u by Corollary 4.8, and in C^{k-1} by Lemma 4.3 (together with the fact that $\beta_m * D^\alpha u = D^\alpha(\beta_m * u)$). Consequently by Lemma 6.3 $q \circ u_m \to q \circ u$ in C^{k-1}, and therefore also in $W^{k-1,p}$ (in particular in L^p) since $q \circ u_m, q \circ u$ are compactly supported. We will show that $\{q \circ u_m\}$ is Cauchy in $W^{k,p}$ and that

$$\|q \circ u_m\|_{k,p} \leq C(1 + \|u_m\|_{C^0}) \sum_{l=0}^{k} (\|u_m\|_{k,p} + \|u_m\|_{C^{k-1}})^{l+1}; \tag{24}$$

from this it would immediately follow that the L^p limit $q \circ u$ belongs to $W^{k,p}$ and obeys the same bound. Note that the first paragraph of the proof shows that (24) holds when $k = 0$ (with $\|\cdot\|_{C^{k-1}}$ interpreted as 0), while we've already noted that u_m is Cauchy in $W^{0,p} = L^p$, so assume inductively that these assertions hold for all integers l with $0 \leq l < k$.

By this inductive hypothesis, it's enough to show that, if $|\alpha| = k-1$, the order k derivatives $D^\alpha(\frac{\partial u_m}{\partial x_i})$ are Cauchy in L^p and have their L^p norms bounded by the right hand side of (24). As in the proof of Lemma 6.3 we have

$$D^\alpha\left(\frac{\partial}{\partial x_i}(q \circ u_m)\right) = \sum_l \sum_{\beta+\gamma=\alpha} D^\beta\left(\frac{\partial q}{\partial y_l} \circ u_m\right) D^\gamma \frac{\partial u_m^l}{\partial x_i}. \tag{25}$$

Since $u_m \to u$ in $W^{k,p}$ the term $D^\gamma \frac{\partial u_m^l}{\partial x_i}$ is Cauchy in L^p for each γ appearing above, while since all β appearing above have order at most $k-1$ Lemma 6.3 shows that $D^\beta\left(\frac{\partial p}{\partial y_l} \circ u_m\right)$ is Cauchy in L^∞. Since the product of a sequence which is Cauchy in L^∞ with one which is Cauchy in L^p is always Cauchy in L^p, this proves that $D^\alpha\left(\frac{\partial}{\partial x_i}(q \circ u_m)\right)$ is Cauchy in L^p when $|\alpha| = k-1$ completing the proof that $q \circ u_m$ is Cauchy in $W^{k,p}$ and hence that $q \circ u \in W^{k,p}$.

It remains to prove the bound (24) for $\|q \circ u_m\|_{k,p}$. By the inductive hypothesis the bound already holds for $\|q \circ u_m\|_{k-1,p}$, so we need only bound the $\|D^\alpha(\frac{\partial(q \circ u)}{\partial x_i})\|_p$ for $|\alpha| = k-1$. Consider the various terms in (25). When $\beta = 0$, the term arising from β is bounded in C^0 norm by $M(1 + \|u_m\|_{C^0})$ (here M is a constant, depending among other things on $\frac{\partial q}{\partial y_l}(\vec{0})$) while the term arising from γ is bounded in L^p norm by $\|u_m\|_{k,p}$. When $\beta \neq 0$, we have

$$D^\beta\left(\frac{\partial q}{\partial y_l} \circ u_m\right) = D^\beta\left(\left(\frac{\partial q}{\partial y_l} - \frac{\partial q}{\partial y_l}(\vec{0})\right) \circ u_m\right),$$

so since $|\beta| \leq k-1$ we may apply the inductive hypothesis with q replaced by $\frac{\partial q}{\partial y_l} - \frac{\partial q}{\partial y_l}(\vec{0})$ (since this vanishes at $\vec{0}$) to get a bound

$$\left\|D^\beta\left(\frac{\partial q}{\partial y_l} \circ u_m\right)\right\|_p \leq C(1 + \|u_m\|_{C^0}) \sum_{l=0}^{k-1} (\|u_m\|_{k-1,p} + \|u_m\|_{C^{k-2}})^{l+1}.$$

Meanwhile when $\beta \neq 0$ the corresponding γ has $|\gamma| < k-1$, so we have a bound $\left\|D^\gamma \frac{\partial u_m^l}{\partial x_i}\right\|_{C^0} \leq \|u\|_{C^{k-1}}$. Combining all of these bounds does indeed give

$$\left\|D^\alpha\left(\frac{\partial(q \circ u)}{\partial x_i}\right)\right\|_p \leq C(1 + \|u\|_{C^0}) \sum_{l=0}^{k} (\|u\|_{W^{k,p}} + \|u\|_{C^{k-1}})^{l+1},$$

completing the induction. □

Proposition 6.2 puts us in position to prove the following, which by induction on k and Corollary 4.13 proves Theorem 6.1 in the case that $p > 2$.

Proposition 6.4. *Let $k \geq 1$, $p > 2$, and suppose that D is a disc around the origin in $\mathbb{C}$ such that $u\colon D \to \mathbb{R}^{2n}$ belongs to $W^{k,p}(D, \mathbb{R}^{2n})$ and obeys*

$$\partial_{\bar{z}} u + (q \circ u)(z) \partial_z u = 0, \quad u(0) = \vec{0}$$

where $q \in C^\infty(V, \mathbb{R}^{2n \times 2n})$ is a matrix-valued function on an open subset of $\mathbb{R}^{2n}$ containing $u(D)$ with $q(\vec{0}) = 0$. Then there is a disc $D' \subset D$ such that D' contains 0 and $u \in W^{k+1,p}(D', \mathbb{R}^{2n})$.

Proof. The proof is significantly simplified by the following observation (the "renormalization trick"):

Claim 6.5. Given any $\epsilon > 0$, it suffices to prove the proposition in the case that $\|u\|_{k,p} < \epsilon$.

To justify this, if $0 < \delta < 1$ denote by $u_\delta \colon D \to \mathbb{R}^{2n}$ the function $u_\delta(z) = u(z)$. (Note that since we are taking the domain of u_δ to be just D, u_δ only takes into account the behavior of u on a disc of radius δ times the radius of the original disc D). Obviously there is a disc D' around 0 on which u belongs to $W^{k+1,p}(D';\mathbb{R}^{2n})$ if and only if the same statement holds (perhaps with a different disc) for u_δ. Our claim will thus follow if we show that $\|u_\delta\|_{W^{k,p}(D;\mathbb{C})} \to 0$ as $\delta > 0$. Now since $k \geq 1$ and $p > 2$ and since $u(0) = 0$ we have by Corollary 4.13 an estimate $|u(z)| \leq C\|u\|_{k,p}|z|^{1-2/p}$, so for some u-dependent constant M the supremum of u_δ over D will be at most $M'\delta^{1-2/p}$. This implies that $\|u_\delta\|_{L^p(D;\mathbb{R}^{2n})} \to 0$ as $\delta > 0$. Meanwhile for $1 \leq |\alpha| \leq k$ we have $(D^\alpha u_\delta)(z) = \delta^{|\alpha|}(D^\alpha u)(\delta z)$, and so

$$\int_D |(D^\alpha u_\delta)(z)|^p d^2z = \delta^{p|\alpha|-2}\int_{\delta D}|D^\alpha u(w)|^p d^2w \leq \delta^{p|\alpha|-2}\|D^\alpha u\|^p_{L^p(D;\mathbb{R}^{2n})} \to 0$$

as $\delta \to 0$. This confirms that $\|u_\delta\|_{W^{k,p}(D;\mathbb{R}^{2n})} \to 0$ as $\delta \to 0$, proving Claim 6.5.

We now proceed with the proof of Proposition 6.4. Choose a pair of concentric discs D_0, D_1 around the origin, with

$$D_0 \subset \bar{D}_0 \subset D_1 \subset \bar{D}_1 \subset D.$$

Moreover let $\alpha, \chi \colon \mathbb{C} \to [0,1]$ be two smooth cutoff functions, with

$$\chi|_{D_0} = 1,\ supp(\chi) \subset D_1,\ \alpha|_{D_1} = 1,\ supp(\alpha) \subset D.$$

Also extend q to a compactly supported function (still denoted q) from $\mathbb{R}^{2n}$ to $\mathbb{R}^{2n\times 2n}$, coinciding with q on a neighborhood of $u(D_1)$ (for instance we could multiply q by a smooth cutoff function). The function χu is then supported in the disc D_1 (so extends by zero to a compactly supported continuous function on $\mathbb{C}$), and by the product rule and the assumed PDE for u obeys

$$\partial_{\bar{z}}(\chi u) + (q \circ u)(z)\partial_z(\chi u) = (\partial_{\bar{z}}\chi)u + (q \circ u)(z)(\partial_z\chi)u. \tag{26}$$

(While we've potentially changed q on $u(D \setminus D_1)$, the equation still holds for $z \in D \setminus D_1$ bacause all terms involved are zero there.) Writing $u' = \alpha u$ (and extending this by zero outside the support of α), since all terms are zero in (26) except where $\alpha u = u$, $u' \colon \mathbb{C} \to \mathbb{R}^{2n}$ is a compactly supported smooth function with

$$\partial_{\bar{z}}(\chi u) + (q \circ u')(z)\partial_z(\chi u) = (\partial_{\bar{z}}\chi)u + (q \circ u')(z)(\partial_z\chi)u. \tag{27}$$

The product rule gives a bound $\|u'\|_{k,p} \leq C\|u\|_{W^{k,p}(D;\mathbb{R}^{2n})}$. By Corollary 4.13 we also have an estimate $\|u'\|_{C^{k-1}} \leq C'\|u'\|_{k,p}$. Hence by Proposition 6.2 for any $\eta > 0$ there is $\epsilon > 0$ such that if $\|u\|_{W^{k,p}(D;\mathbb{R}^{2n})} < \epsilon$ then $q \circ u' \in W^{k,p}(\mathbb{C};\mathbb{R}^{2n})$ with $\|q \circ u'\|_{k,p} < \eta$. Hence by Claim 6.5, given any $\eta > 0$ we may (and do) assume that u obeys (27) where the function $q \circ u' \colon \mathbb{C} \to \mathbb{R}^{2n\times 2n}$ belongs to $W^{k,p}$ with $\|q \circ u'\|_{k,p} < \eta$.

Write

$$h(z) = (\partial_{\bar{z}}\chi)(z)u(z) + (q \circ u')(z)(\partial_z\chi)(z)u(z)$$

for the right hand side of (27). By Corollary 4.14 we have $h \in W^{k,p}(\mathbb{C};\mathbb{R}^{2n})$ (using that $k \geq 1$ and $p > 2$). Now where T is the Calderon–Zygmund operator of the previous section, so that in particular $T \circ \partial_{\bar{z}} = \partial_z$ on $W^{k,p}(\mathbb{C},\mathbb{R}^{2n})$, (27) can be written

$$(Id + (q \circ u') \cdot T)\partial_{\bar{z}}(\chi u) = h.$$

The Calderón–Zygmund Theorem 5.6, together with Corollary 4.14, shows that we have a bound

$$\|(q \circ u') \cdot Tg\|_{k,p} \leq C\|q \circ u'\|_{k,p}\|g\|_{k,p} \leq C\eta\|g\|_{k,p}$$

by our assumption on $q \circ u'$. For η choose the value $\frac{1}{2C}$, so that $(q \circ u') \cdot T$ has operator norm at most $\frac{1}{2}$ as a linear endomorphism of $W^{k,p}$. Since $W^{k,p}$ is a Banach space, $Id + (q \circ u') \cdot T$ is then invertible, with inverse $B := \sum_{m=0}^{\infty}(-(q \circ u') \cdot T)^m$. Thus we have

$$\partial_{\bar{z}}(\chi u) = Bh \in W^{k,p}(\mathbb{C},\mathbb{R}^{2n}).$$

Hence by Theorem 5.10 $\chi u \in W^{k+1,p}(D;\mathbb{R}^{2n})$, so since u coincides with χu on D_0 we obtain $u \in W^{k+1,p}(D_0;\mathbb{R}^{2n})$, completing the proof. □

Proposition 6.4 proves that a map $u\colon \Sigma \to M$ satisfying the nonlinear Cauchy-Riemann equation which belongs (locally) to $W^{1,p}$ for some $p > 2$ is smooth on the interior of the domain: if $z_0 \in int(\Sigma)$, repeated application of Proposition 6.4 shows that for every $k \geq 1$ there is a neighborhood of z_0 on which $u \in W^{k,p}$, so Corollary 4.13 (applied to the product of u and a cutoff function) shows that u is C^{k-1} on a smaller neighborhood of z_0. Thus for each $k \geq 1$ u is C^{k-1} on neighborhoods of every point in $int(\Sigma)$, which is to say that u is C^{k-1} on $int(\Sigma)$.

It remains to prove Theorem 6.1 when $1 < p \leq 2$. Evidently by what we have already done it is enough to prove that if $u\colon D \to \mathbb{R}^{2n}$ obeys $\partial_{\bar{z}}u + (q \circ u)\partial_z u = 0$ and $u(0) = \vec{0}$ and if u is continuous and $W^{1,p}$, then for some $p^* > 2$ and some disc D' around 0 we have $u \in W^{1,p^*}(D'; \mathbb{R}^{2n})$. The proof of this is structurally similar to that of Proposition 6.4. Since we assume u is continuous and since $u(0) = \vec{0}$, given $\epsilon > 0$ there is $\delta > 0$ such that whenever $z \in D$ we have $|u(\delta z)| < \epsilon$. So since the conclusion holds for u if and only if it holds for u_δ (possibly with a different disc D', where again $u_\delta(z) = u(\delta z)$), for any given $\epsilon > 0$ it suffices to consider the case that $\|u\|_{C^0} < \epsilon$. Letting χ and the C_0^∞ extension of q be as in the proof of Proposition 6.4 we have an equation

$$\partial_{\bar{z}}(\chi u) + (q \circ u)\partial_z(\chi u) = (\partial_{\bar{z}}\chi)u + (q \circ u)(\partial_z \chi u)u =: h.$$

Now where $p^* = \frac{2p}{2-p}$ if $p < 2$ and p^* is any number larger than 2 if $p = 2$, Theorem 4.17 gives an embedding $W^{1,p}(D, \mathbb{R}^{2n}) \to L^{p^*}(D, \mathbb{R}^{2n})$ (for $p = 2$ this uses that, since D has finite measure, we have $W^{1,2}(D, \mathbb{R}^{2n}) \subset W^{1,p'}(D, \mathbb{R}^{2n})$ for any $p' < 2$). Note that $p^* > 2$. So $u \in L^{p^*}$, and then since $q \circ u$ is bounded on the support of χ it follows that $h \in L^{p^*}$. Also, since q is compactly supported and smooth it is Lipschitz, so (since $u(0) = \vec{0}$) $\|q \circ u\|_{C^0} \leq M\|u\|_{C^0}$.

We thus have

$$(Id + (q \circ u) \cdot T)\partial_{\bar{z}}(\chi u) \in L^{p^*}.$$

By the Calderón–Zygmund theorem T is a bounded linear operator on L^{p^*}, and then multiplication by $q \circ u$ gives an operator on L^{p^*} of norm at most $\|q \circ u\|_{C^0} \leq M\|u\|_{C^0}$. So by assuming $\|u\|_{C^0} < \epsilon$ for sufficiently small ϵ we guarantee that $(q \circ u) \cdot T$ has norm at most $1/2$ on L^{p^*} and hence that $Id + (q \circ u)T$ is invertible on L^{p^*}. Hence we obtain $\partial_{\bar{z}}(\chi u) \in L^{p^*}$, which by Theorem 5.10 proves that $\chi u \in W^{1,p^*}$. So since u agrees with χu on a disc around 0 we have proven that there is a disc around zero on which u belongs to W^{1,p^*}. Since $p^* > 2$ this completes the proof of Theorem 6.1, as we may inductively apply Proposition 6.4.

6.2. **The Carleman Similarity Principle and its applications.** For an open subset D containing $0 \in \mathbb{C}$, call two functions $u, v\colon D \to \mathbb{R}^{2n}$ *similar near* 0 if the following holds: there is an open D' with $0 \in D' \subset D$ and a function $\Phi\colon D' \to GL(2n, \mathbb{R})$ such that $v(z) = \Phi(z)u(z)$ for each $z \in D'$. Consistently with our earlier assertion that pseudoholomorphic curves are locally very much like holomorphic functions to $\mathbb{C}^n$, we will see that any pseudoholomorphic curve with domain containing $0 \in \mathbb{C}$ is similar to a holomorphic function, with the caveat that we will only be able to get fairly limited regularity on the "similarity" Φ.

In fact we will consider a bit more generally solutions $u\colon D \to \mathbb{R}^{2n}$ to equations of form

$$\frac{\partial u}{\partial s} + \mathsf{J}(z)\frac{\partial u}{\partial t} + C(z)u(z) = 0, \tag{28}$$

where as usual we write $z = s + it$, and where for each $z \in D$ the terms $\mathsf{J}(z)$ and $C(z)$ are $2n \times 2n$-matrices, with $\mathsf{J}(z)$ an almost complex structure and J and C varying smoothly with z (actually significantly weaker regularity assumptions can be made on J and C if you prefer; see [MS2, Section 2.3] for precise statements). By what we have already done, any J-holomorphic curve u (for J a smooth almost complex structure) solves such an equation; just set $C = 0$ and $\mathsf{J} = J \circ u$—the latter is smooth since we now know that u is smooth. What follows will show, essentially, that the "similarity class" of solutions to (28) is independent of the choice of J and C, at least if we are willing to settle for non-smooth "similarities" Φ. First we remove the dependence on J.

Lemma 6.6. *If D is a disc around $0 \in \mathbb{C}$ and $u\colon D \to \mathbb{R}^{2n}$ is a solution of (28), then there is an open set D' with $0 \subset D' \subset D$ and smooth functions $\Phi\colon D' \to GL(2n, \mathbb{R})$, $B\colon D' \to \mathbb{R}^{2n \times 2n}$, and $v\colon D' \to \mathbb{R}^{2n}$ such that, for*

$z \in D'$,

$$v(z) = \Phi(z)u(z) \quad \frac{\partial v}{\partial s} + J_0 \frac{\partial v}{\partial t} + B(z)v(z) = 0.$$

Proof. For Φ, set $\Phi(z) = \Psi(z)^{-1}$ where $\Psi\colon D' \to GL(2n, \mathbb{R})$ such that $\Psi(z)^{-1}\mathsf{J}(z)\Psi(z) = J_0$. This function Ψ can be constructed using the implicit function theorem applied to the map $\mathcal{F}\colon GL(2n, \mathbb{R}) \to \{J \in GL(2n, \mathbb{R}) | J^2 = -1\}$ defined by $\Psi \mapsto \Psi J_0 \Psi^{-1}$, the point being that, choosing $\Psi(0)$ such that $\mathcal{F}(\Psi(0)) = \mathsf{J}(0)$, the linearization of $\mathcal{F}$ is surjective at $\Psi(0)$, so the the implicit function theorem solves the equation $\mathcal{F}(\Psi(z)) = \mathsf{J}(z)$ for z sufficiently close to zero. Another way of expressing the existence of Ψ is to say that the almost complex structure on the trivial bundle $\mathbb{R}^{2n} \times D \to D$ given by acting by $J(z)$ on the fiber over z has a trivialization as a *complex* vector bundle on a neighborhood D' of the origin. This is a special case of the fact that if a vector bundle carries an almost complex structure then the transition functions defining the vector bundle can be taken to live in $GL(n, \mathbb{C})$ (in fact they can even be taken in $U(n)$).

Since $u = \Psi v$ and $J\Psi = \Psi J_0$, we then have

$$\begin{aligned} 0 = \frac{\partial u}{\partial s} + \mathsf{J}\frac{\partial u}{\partial t} + Cu &= \frac{\partial}{\partial s}(\Psi v) + \mathsf{J}\frac{\partial}{\partial t}(\Psi v) + C\Psi v \\ &= \Psi\left(\frac{\partial v}{\partial s} + J_0\frac{\partial v}{\partial t}\right) + \left(\frac{\partial \Psi}{\partial s} + \mathsf{J}\frac{\partial \Psi}{\partial t} + C\Psi\right)v. \end{aligned}$$

Thus the lemma holds with

$$B(z) = \Psi(z)^{-1}\left(\frac{\partial \Psi}{\partial s}(z) + \mathsf{J}(z)\frac{\partial \Psi}{\partial t}(z) + C\Psi(z)\right).$$

□

The ring $\mathbb{R}^{2n\times 2n}$ of $2n\times 2n$ square matrices (i.e. of $\mathbb{R}$-linear transformations of $\mathbb{R}^{2n}$) contains as a subring a copy of the *complex*-linear transformations $\mathbb{C}^{n\times n}$ of $\mathbb{C}^n$, where we identify the standard almost complex structure J_0 on $\mathbb{R}^{2n}$ with multiplication by i. A real-linear transformation of $\mathbb{R}^{2n}$ is complex linear if and only if it commutes with multiplication by i; thus

$$\mathbb{C}^{n\times n} = \{B \in \mathbb{R}^{2n\times 2n} | BJ_0 = J_0 B\}.$$

Lemma 6.7. *Lemma 6.6 continues to hold with $B\colon D' \to \mathbb{R}^{2n\times 2n}$ replaced by a bounded (but not necessarily continuous) function $B'\colon D' \to \mathbb{C}^{n\times n}$.*

Proof. We just need to find, for each z, a matrix $B'(z) \in C^{n\times n}$ such that $B'(z)v(z) = B(z)v(z)$. If $v(z) = 0$ we'll just set $B'(z) = 0$. For $v(z) \neq 0$, $B'(z)$ will be the matrix representing the following complex-linear transformation: first project $\mathbb{C}^n$ onto the complex line through $v(z)$ (using the orthogonal projection induced by the standard inner product), and then compose this with the linear transformation $\mathbb{C}v(z) \to \mathbb{C}^n$ which sends $(a + ib)v(z)$ to $(a + ib)B(z)v(z)$. Thus in formulas:

$$B'(z)w = \begin{cases} 0 & v(z) = 0 \\ \frac{\overline{v(z)}^T w}{\overline{v(z)}^T v(z)} B(z)v(z) & v(z) \neq 0. \end{cases}$$

Clearly $B'(z) \in \mathbb{C}^{n\times n}$ and $B'(z)v(z) = B(z)v(z)$ for each z, and the entries of the matrix $B'(z)$ are uniformly bounded by a constant times $\|B\|_{C^0}$, so $B' \in L^\infty(D'; \mathbb{C}^{n\times n}$. □

Now suppose $u\colon D \to \mathbb{R}^{2n}$ and $B' \in L^\infty(D; \mathbb{C}^{n\times n})$ with

$$\frac{\partial u}{\partial s} + J_0\frac{\partial u}{\partial t} + B'(z)u(z) = 0 \tag{29}$$

We attempt to find a subdisk D' containing 0 and a map $\Phi\colon D' \to \mathbb{C}^{n\times n} \subset \mathbb{R}^{2n\times 2n}$ so that $v := \Phi u$ is holomorphic. Since $\Phi J_0 = J_0\Phi$, we see that

$$\begin{aligned}\frac{\partial v}{\partial s} + J_0\frac{\partial v}{\partial t} &= \left(\frac{\partial \Phi}{\partial s} + J_0\frac{\partial \Phi}{\partial t}\right)u + \Phi\left(\frac{\partial u}{\partial s} + J_0\frac{\partial u}{\partial t}\right) \\ &= \left(\frac{\partial \Phi}{\partial s} + J_0\frac{\partial \Phi}{\partial t}\right)u - \Phi(z)B'(z)u(z).\end{aligned}$$

From this it immediately follows that (recalling $\partial_{\bar{z}} = \frac{1}{2}(\partial_s + J_0\partial_t)$):

Proposition 6.8. *If u satisfies (29) and if $\Phi\colon D' \to \mathbb{C}^{n\times n}$ obeys*

$$\partial_{\bar{z}}\Phi = 2\Phi B'$$

then $v = \Phi u\colon D' \to \mathbb{R}^{2n}$ is holomorphic with respect to the standard complex structure J_0.

The above equation can indeed be solved for Φ:

Theorem 6.9. *Let D be a disc containing $0 \in \mathbb{C}$, and let $A \in L^\infty(D;\mathbb{C}^{n\times n})$. Then there is a disc D' with $0 \in D' \subset D$ and a map $\Phi\colon D' \to \mathbb{C}^{n\times n}$ such that*

$$\partial_{\bar{z}}\Phi = \Phi(z)A(z) \text{ for } z \in D'.$$

Moreover for every $p < \infty$ we have $\Phi \in W^{1,p}(D';\mathbb{C}^{n\times n})$, and $\Phi(z)$ is invertible for each $z \in D'$.

Corollary 6.10. *If $u\colon D \to \mathbb{R}^{2n}$ satisfies (28) where D is a disk around the origin in $\mathbb{C}$, there is a disc D' such that $0 \in D' \subset D$ and a map $\Phi\colon D' \to GL(2n,\mathbb{R})$ of class $W^{1,p}$ for all $p < \infty$ such that $z \mapsto \Phi(z)u(z)$ is a holomorphic map.*

Proof of Corollary 6.10. Indeed, this follows directly (after appropriate renamings) by combining Theorem 6.9 with Lemmas 6.6 and 6.7 and Proposition 6.8. □

Proof of Theorem 6.9. Similarly to the proof of Proposition 6.4 we can employ a renormalization trick:

Claim 6.11. For any $\epsilon > 0$ it suffices to prove the theorem when $\|A\|_\infty < \epsilon$

Indeed, if $0 < \delta < 1$ the function $\Phi_\delta\colon D \to \mathbb{C}^{n\times n}$ defined by $\Phi_\delta(z) = \Phi(\delta z)$ has the property that, for any subdisc $D_0 \subset D$, $\partial_{\bar{z}}\Phi = \Phi A$ on $\{\delta z | z \in D_0\}$ if and only if $\partial_{\bar{z}}\Phi_\delta = (\Phi)(\delta A)$ on D_0, and of course $\|\delta A\|_\infty \to 0$ as $\delta \to 0$. So by choosing δ with $\|\delta A\|_\infty < \epsilon$ and finding a solution $\Phi'\colon D_0 \to \mathbb{C}^{n\times n}$ on a disk D_0 containing zero to $\partial_{\bar{z}}\Phi' = A(\delta\Phi')$ we may recover a solution to the original equation (having the same regularity as Φ', and defined on a smaller disc δD_0) by setting $\Phi(z) = \Phi'(\delta^{-1}z)$.

Accordingly assume $\|A\|_\infty < \epsilon$ where ϵ is a small number to be specified later. Choose a subdisk $D' \subset D$ containing 0 and a smooth cutoff function $\chi\colon \mathbb{C} \to [0,1]$ supported in D and with $\chi|_{D'} = 1$; we will cut our functions off using χ in order to work with compactly supported functions. Specifically, we search for a compactly supported function $\Psi\colon \mathbb{C} \to \mathbb{C}^{n\times n}$ with

$$\Psi(z) = \chi(z)\left(P(\chi(1+\Psi)A)\right)(z).$$

If Ψ solves this, then setting $\Phi = Id + \Psi$ we would have

$$\partial_{\bar{z}}\Phi = \partial_{\bar{z}}\Psi = \partial_{\bar{z}}P(\chi\Phi A) = \chi\Phi A = \Phi A \text{ on } D',$$

at least assuming that $\chi\Phi A$ is within the class of functions on which P is well-defined with $\partial_{\bar{z}}P$ equal to the identity (recall that we've only established this for compactly supported smooth functions). Indeed, we have the following:

Proposition 6.12. *Fix functions $f_1, f_2 \in C_0^\infty(\mathbb{C};\mathbb{C})$ and define a linear map $R_{f_1,f_2}\colon C_0^\infty(\mathbb{C};\mathbb{C}) \to C_0^\infty(\mathbb{C};\mathbb{C})$ by*

$$(R_{f_1 f_2}u)(z) = f_1(z)(P(f_2u))(z).$$

Then for any $1 < p < \infty$ there is a constant C such that we have

$$\|R_{f_1 f_2}u\|_{1,p} \leq C\|u\|_p.$$

Hence $R_{f_1 f_2}$ extends to a bounded linear operator $R_{f_1 f_2}\colon L^p(\mathbb{C};\mathbb{C}) \to W^{1,p}(\mathbb{C};\mathbb{C})$

Proof. Recall that for $g \in C_0^\infty(\mathbb{C};\mathbb{C})$ we have

$$Pg(z) = \int_{\mathbb{C}} \frac{1}{\pi(z-w)} g(w)d^2w$$

(since $d^2w = -2idw \wedge d\bar{w}$), i.e., where $K(z) = \frac{1}{\pi z}$, $Pg = K * g$.

Let Ω be a compact set containing the supports of f_1 and f_2. Let $M = \max\{\|f_1\|_\infty, \|f_2\|_\infty$, and let $\tilde{\Omega} = \{z \in \mathbb{C} | (\exists x, y \in \Omega)(z = x - y)\}$. Note that $\tilde{\Omega}$ is compact and so has finite measure. For $u \in C_0^\infty(\mathbb{C};\mathbb{C})$ we have

$$\|R_{f_1 f_2}u\|_p^p = \int_{\mathbb{C}} |f_1(z)|^p \left| \int_{\mathbb{C}} K(z-w)(f_2(w)u(w))d^2w \right|^p d^2z \leq M^p \int_{\Omega} \left| \int_{\Omega} K(z-w)(f_2(w)u(w))d^2w \right|^p d^2z. \tag{30}$$

If K belonged to L^1, we could have used Young's inequality to estimate $\|K * (f_2u)\|_p \leq \|K\|_1 \|f_2u\|_p$. Now K does not belong to L^1; however letting

$$\tilde{K}(z) = \begin{cases} K(z) & z \in \tilde{\Omega} \\ 0 & z \notin \tilde{\Omega} \end{cases}$$

the last formula in (30) is unchanged if K is replaced by $\tilde{K}$ and is bounded above by $\|\tilde{K} * (f_2u)\|_p^p$. But since $\tilde{\Omega}$ is a bounded subset of $\mathbb{C}$ one readily checks that $\tilde{K}$ is integrable, and so Young's inequality gives

$$\|\tilde{K} * (f_2u)\|_p \leq \|\tilde{K}\|_1 \|f_2u\|_p \leq M\|\tilde{K}\|_1 \|u\|_p,$$

and hence

$$\|R_{f_1 f_2}u\|_p \leq M^2 \|\tilde{K}\|_1 \|u\|_p.$$

Meanwhile we have

$$\partial_{\bar{z}}(R_{f_1 f_2}u) = (\partial_{\bar{z}} f_1)P(f_2u) + f_1 \partial_{\bar{z}} P(f_2u) = (\partial_{\bar{z}} f_1)P(f_2u) + f_1 f_2 u$$

and

$$\partial_z(R_{f_1 f_2}u) = (\partial_z f_1)P(f_2u) + f_1(\partial_z P(f_2u)) = (\partial_z f_1)P(f_2u) + f_1 T(f_2u).$$

If $M' = \max\{\|\partial_{\bar{z}} f_1\|_\infty, \|\partial_z f_1\|_\infty, \|f_2\|_\infty\}$ the above shows that

$$\|(\partial_{\bar{z}} f_1)P(f_2u)\|_p, \|(\partial_z f_1)P(f_2u)\|_p \leq M'^2 \|\tilde{K}\|_1 \|u\|_p,$$

while $\|f_1 f_2 u\|_p \leq M^2 \|u\|_p$ and by the Calderón-Zygmund theorem there is a universal constant C such that $\|f_1 T(f_2u)\|_p \leq MC\|f_2u\|_p \leq M^2 C\|u\|_p$.

Thus $R_{f_1 f_2}u$, $\partial_{\bar{z}}(R_{f_1 f_2}u)$ and $\partial_z(R_{f_1 f_2}u)$ all have L^p norms bounded by a universal constant times $\|u\|_p$, so the proposition follows. □

Given this proposition, choose an arbitrary $p > 2$. Then the map $R_{\chi,\chi}\colon C_0^\infty(\mathbb{C};\mathbb{C}^{n\times n}) \to C_0^\infty(\mathbb{C};\mathbb{C}^{n\times n})$ defined by $R_{\chi,\chi}(X) = \chi P(\chi X)$ extends to a linear map $L^p \to W^{1,p}$, with a bound $\|R_{\chi,\chi}X\|_{1,p} \leq C_\chi \|X\|_p$ for some constant C_χ. By Claim 6.11 we may assume $\|A\|_\infty < \frac{1}{3C_\chi M}$ where $M > 1$ is a constant as in Morrey's inequality: $\|u\|_{C^0} \leq C\|u\|_{W^{1,p}}$. Extend A from its initial domain D to all of $\mathbb{C}$ by setting it equal to zero on $\mathbb{C} \setminus D$. Thus $A \in L^\infty(\mathbb{C};\mathbb{C}^{n\times n}) \cap L^p(\mathbb{C};\mathbb{C}^{n\times n})$ since D is bounded. For $\Psi \in L^p(\mathbb{C};\mathbb{C}^{n\times n})$ we hence have $(1+\Psi)A \in L^p(\mathbb{C};\mathbb{C}^n)$ (for $\Psi A \in L^p$ because $\Psi \in L^p$ and $A \in L^\infty$). So we may define a map $G\colon L^p(\mathbb{C};\mathbb{C}^{n\times n}) \to W^{1,p}(\mathbb{C};\mathbb{C}^{n\times n}$ by

$$G(\Psi) = R_{\chi,\chi}((1+\Psi)A) = \chi P(\chi(1+\Psi)A).$$

Observe that, for $\Psi, \Psi' \in L^p(\mathbb{C}; \mathbb{C}^{n\times n})$,

$$\|G(\Psi)-G(\Psi')\|_p = \|\chi P(\chi(\Psi-\Psi')A)\|_p = \|R_f((\Psi-\Psi')A)\|_p \le C_\chi \|(\Psi-\Psi')A\|_p \le C_\chi \|A\|_\infty \|\Psi-\Psi'\|_p < \frac{1}{3M}\|\Psi-\Psi'\|_p$$

by our assumption on $\|A\|_\infty$. Thus G is a contractive mapping, so it has a unique fixed point Ψ (which may be obtained as the limit of the sequence ψ_n defined by $\psi_0 = 0$ and $\psi_{n+1} = G(\psi_n)$). This map Ψ belongs to L^p and obeys

$$\Psi = R_f((1+\Psi)A).$$

But since R_f maps L^p to $W^{1,p}$ and $(1+\Psi)A \in L^p$ when $\Psi \in L^p$ it follows that in fact $\Psi \in W^{1,p}$. In particular $\partial_{\bar z}\Psi$ is a well-defined element of L^p and

$$\partial_{\bar z}\Psi = \partial_{\bar z}\left(R_f((1+\Psi)A)\right) = (\partial_{\bar z}\chi)P(\chi(1+\Psi)A) + \chi\partial_{\bar z}P(\chi(1+\Psi)A).$$

Now the equation $\partial_{\bar z}P = Identity$ on C_0^∞ extends by continuity to L^p, so the second term above is $\chi\cdot(\chi(1+\Psi)A)$. In particular, *on* D' (where χ is identically 1 and so $\partial_{\bar z}\chi = 0$) we obtain

$$\partial_{\bar z}\Psi = (1+\Psi)A \text{ on } D'.$$

Moreover, since we chose our $p > 2$ and since $\Psi \in W^{1,p}$, Ψ is in fact continuous. Hence $(1+\Psi)A \in L^\infty(D'; \mathbb{C}^{n\times n})$. Since D' is bounded it follows that $(1+\Psi)A \in L^q$ for every $q < \infty$. So since $\partial_{\bar z}\Psi = (1+\Psi)A$ it follows from Theorem 5.10 that $\Psi \in W^{1,q}(D''; \mathbb{C}^{n\times n})$ for any open subset D'' of D' with closure contained in D'.

Finally, note that if $X \in W^{1,p}$ with $\|X\|_{1,p} \le 1/2M$, then

$$\|G(X)\|_{W^{1,p}} \le C_\chi\|A\|_\infty \|1+X\|_p \le \frac{1}{3M}\|1+X\|_p \le \frac{1}{3M}\cdot\frac{3}{2} = \frac{1}{2M}.$$

So by the construction of Ψ (obtained as the limit of a sequence constructed iteratively applying G to zero) it follows that $\|\Psi\|_{W^{1,p}} < 1/2M$, and so (by the choice of M) $\|\Psi\|_{C^0} \le 1/2$. Hence the matrix $\Phi(z) = 1+\Psi(z)$ is invertible for every $z \in D$, defines a map $D'' \to \mathbb{C}^{n\times n}$ which belongs to $W^{1,p}$ for every $p < \infty$, and at all points in D'' we have $\partial_{\bar z}\Phi = \Phi A$. □

Recall that a holomorphic function $u\colon D \to \mathbb{C}^n$ is determined as soon as we specify the values of u and of all of its derivatives at a single point. Thus if $u, v\colon D \to \mathbb{C}^n$ are two different holomorphic functions such that the function $w = u - v$ has the property that, for every $k \in \mathbb{N}$, $\lim_{|z|\to 0} \frac{|w(z)|}{|z|^k} = 0$ (which evidently forces the derivatives of order k of w to vanish at 0), then $w = 0$ and so $u = v$. The Carleman similarity principle allows us to carry this over to J-holomorphic curves (and indeed even a bit more generally):

Proposition 6.13. *Let D be a disc in $\mathbb{C}$, J an almost complex structure on $\mathbb{C}^n$, and $C\colon D \to \mathbb{R}^{2n\times 2n}$ a smooth function, and suppose that $u_0, u_1\colon D \to \mathbb{C}^n$ each satisfy*

$$\frac{\partial u_i}{\partial s} + J(u_i(z))\frac{\partial u_i}{\partial t} + C(z)u_i(z) = 0.$$

If $z_0 \in D$ and for all $k \in \mathbb{N}$ we have

$$\lim_{z\to z_0} \frac{u_1(z) - u_0(z)}{|z-z_0|^k} = 0,$$

then there is a neighborhood D' of z_0 such that $u_0 = u_1$ throughout D'.

Proof. Let $w = u_1 - u_0$. Subtracting the respective equations satisfied by u_1 and u_0 gives

$$\begin{aligned} 0 &= \frac{\partial w}{\partial s} + \left(J(u_1(z))\frac{\partial u_1}{\partial t} - J(u_0(z))\frac{\partial u_0}{\partial t}\right) + C(z)w(z) \\ &= \frac{\partial w}{\partial s} + J(u_1(z))\frac{\partial w}{\partial t} + C(z)w(z) + (J(u_1(z)) - J(u_0(z)))\frac{\partial u_0}{\partial t}. \end{aligned}$$

Now

$$(J(u_1(z)) - J(u_0(z)))\frac{\partial u_0}{\partial t} = \int_0^1 \frac{d}{d\tau}\left(J(u_0(z) + tw(z))\right)\frac{\partial u_0}{\partial t}d\tau = B(z)w(z)$$

by the chain rule, where B is a certain matrix valued function which depends on the derivative of J and on $\frac{\partial u_0}{\partial t}$. So w solves an equation of the form

$$\frac{\partial w}{\partial s} + J(u_1(z))\frac{\partial w}{\partial t} + (B(z) + C(z))w(z) = 0$$

for $z \in D$, whence Theorem 6.9 gives a disc D' containing z_0 and map $\Phi\colon D' \to GL(2n, \mathbb{R})$ such that $w' := z \mapsto \Phi(z)w(z)$ is a holomorphic function on D', and $\Phi \in W^{1,p}$ for every $p < \infty$. So by Corollary 4.13, there is a constant C such that $|\Phi(z)| < C$ for all $z \in D'$. Then for every $k \in \mathbb{N}$

$$\frac{|w'(z)|}{|z - z_0|^k} \leq \frac{|\Phi(z_0)w(z)|}{|z - z_0|^k} \leq C\frac{|w(z)|}{|z - z_0|^k} \to 0$$

as $z \mapsto z_0$. Since w' is holomorphic this forces w' to be identically zero on D', and therefore $u_1 - u_0 = w = \Phi^{-1}w'$ is identically zero on D'. □

Corollary 6.14. *(Unique continuation) Let Σ be a connected Riemann surface, M a smooth manifold with almost complex structure J, and $u_0, u_1\colon \Sigma \to M$ to J-holomorphic curves which agree to infinite order at some point $z_0 \in \Sigma$. Then $u_0 = u_1$ throughout Σ.*

Of course, by definition, two functions agree to infinite order at a point if their derivatives of all orders coincide there.

Proof. The set $S = \{z \in \Sigma | u_0$ and u_1 agree to infinite order at $z\}$ is the intersection of closed sets and therefore is closed, and S is nonempty since it contains z_0. In local coordinates around any point $z \in S$, u_0 and u_1 satisfy the hypotheses of Proposition 6.13, and so Proposition 6.13 shows that S contains a neighborhood of z, proving that S is open. Since S is a nonempty, open, and closed subset of the connected surface Σ it follows that $S = \Sigma$. □

Corollary 6.15. *Let $u\colon \Sigma \to M$ be a J-holomorphic map from a connected Riemann surface Σ. If u is not constant, then the set of critical points of u is discrete.*

Proof. We must show that any critical point z_0 of u has a neighborhood in which z_0 is the only critical point; thus it suffices to work in a coordinate neighborhood of z_0, in which u appears as a map $D \to \mathbb{C}^n$ obeying $\frac{\partial u}{\partial s} + \mathsf{J}(z)\frac{\partial u}{\partial t} = 0$. Differentiating this equation with respect to s gives, where $v = \frac{\partial u}{\partial s}$,

$$\begin{aligned} 0 &= \frac{\partial^2 u}{\partial s} + \frac{\partial \mathsf{J}}{\partial s}\frac{\partial u}{\partial t} + \mathsf{J}\frac{\partial^2 u}{\partial s \partial t} \\ &= \frac{\partial v}{\partial s} + \mathsf{J}(z)\frac{\partial v}{\partial t} + \frac{\partial \mathsf{J}}{\partial s}\mathsf{J}(z)v(z). \end{aligned}$$

There is then a disc D' around z_0 and a map $\Phi\colon D' \to GL(2n, \mathbb{R})$ such that Φv is holomorphic, and hence either has only isolated, finite order zeros or vanishes identically on D'. So since each $\Phi(z)$ is invertible v either has isolated, finite order zeroes or vanishes identically on D'. The equation $\frac{\partial u}{\partial s} + J\frac{\partial u}{\partial t}$ shows that $v = \frac{\partial u}{\partial s}$ vanishes at a point if and only if the derivative du vanishes there (this latter being a coordinate indepenent notion). It hence follows that the set on which du vanishes to infinite order is open and closed in Σ. So since Σ is connected either du vanishes identically (so u is constant) or else there is no point at which du vanishes to infinite order, which in particular implies that there is no open subset on which du vanishes identically. So since the above discussion shows that if du had a non-isolated zero it would vanish on a neighborhood of that zero, this proves that if u is nonconstant then all of its zeros are isolated. □

6.3. **Isolation of intersection points and somewhere injectivity.** We now consider intersection points between (local) pseudoholomorphic curves. Let D be a small disc in $\mathbb{C}$ around 0, let J be an almost complex structure on a smooth manifold M, and let $u, v\colon D \to M$ be two J-holomorphic curves such that $u(0) = v(0)$. The results of the previous subsection show that the set of points $z \in D$ at which $u(z) = v(z)$ is either discrete or equal to all of D; however a this does not address the question of how the images of u and v intersect each other, since of course we might have $u(z) = v(w)$ for some $z \neq w$. Nonetheless, the Carleman similarity principle can again be used to prove a similar dichotomy for intersection points, at least away from critical points:

Theorem 6.16. *Let $u, v\colon D \to M$ be two J-holomorphic maps where $D \subset \mathbb{C}$ is a disc containing 0, and suppose that $u(0) = v(0)$ and that $du(0) \neq 0$. Then one of the following holds:*

- *There are neighborhoods $U, V \subset D$ of 0 such that $v(V) \subset u(U)$. In this case, there is a holomorphic map $\phi\colon V \to U$ such that*
$$v|_V = u \circ \phi.$$
- *There are neighborhoods $U, V \subset D$ of 0 such that*
$$u(U) \cap v(V) = \{u(0)\}.$$

Proof. Since $du(0) \neq 0$ and u is J-holomorphic, the map $du(0)\colon T_0D \to T_{u(z_0)}M$ is injective; by continuity the same is true for $du(z)$ for all z sufficiently close to 0. Thus (at least on a smaller disc around 0) u is an immersion, and so it restricts to a (possibly still smaller) neighborhood D' of 0 as an embedding. Consequently coordinates may be chosen on a neighborhood of $u(0) \in M$ in such a way that, with respect to these coordinates, $u|_{D'}$ appears as an embedding with image contained in $\mathbb{R}^2 \times \{0\}$; write the image of $u|_{D'}$ as $W \times \{0\}$ where W is an open neighborhood of the origin in $\mathbb{R}^2$. Note that the fact that u is J-holomorphic implies that the tangent space to $u(D')$ is preserved by J, so in these coordinates the restriction of J to $W \times \{0\}$ preserves the factor TW. Furthermore, by possibly changing our initial choice of coordinates and shrinking D', we may arrange that, additionally, the restriction of J to $W \times \{0\}$ preserves the factor $T_0\mathbb{R}^{2n-2}$. (To carefully construct these coordinates, proceed as follows: choose a Riemannian metric g on M which is invariant under J, and in a neighborhood D' of zero choose a frame $\{X_3, X_4, \ldots, X_{2n}\}$ for the g-orthogonal complement of $T(u(D'))$ such that, for each $w \in u(D')$, $J(X_{2j-1}(w)) = X_{2j}(w)$. Then the coordinates $(x_1, x_2, \ldots, x_{2n})$ will correspond to the point $exp_{u(x_1+ix_2)}\left(\sum_{j=3}^{2n} x_jX_j\right)$.)

In terms of the coordinates of the previous paragraph, we may thus view $u|_{D'}$ and $v|_{D'}$ as maps $D' \to \mathbb{C} \times \mathbb{C}^{n-1}$. Express them accordingly as $(u_1, \tilde{u})$ and $(v_1, \tilde{v})$ where $u_1, v_1\colon D' \to \mathbb{C}$ and $\tilde{u}, \tilde{v}\colon D' \to \mathbb{C}^{n-1}$; we have arranged that $\tilde{u} = \vec{0}$ and that u_1 is a diffeomorphism onto an open subset of $\mathbb{C}$. With respect to the splitting, J has the block form $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ where we have arranged that $B(w, \vec{0}) = C(w, \vec{0})$ for each w, and therefore that $A(w, \vec{0})$ and $D(w, \vec{0})$ are almost complex structures on their respective domains $\mathbb{C}$ and $\mathbb{C}^{n-1}$. Let $\tilde{\pi}\colon \mathbb{C} \times \mathbb{C}^{n-1} \to \mathbb{C}^{n-1}$ be the projection onto the second factor. Note in particular that we then have, for a general vector $(x_1, \tilde{x}) \in \mathbb{C}^n$,
$$\tilde{\pi}\left(J(w, \vec{0})(x_1, \tilde{x})\right) = D(w, \vec{0})\tilde{x}.$$
We have
$$\begin{aligned} 0 &= \tilde{\pi}_*\left(\frac{\partial v}{\partial s} + J(v_1(z), \tilde{v}(z))\frac{\partial v}{\partial t}\right) \\ &= \frac{\partial \tilde{v}}{\partial s} + \tilde{\pi}_*\left(J(v_1(z), \vec{0})\frac{\partial v}{\partial t} + (J(v_1(z), \tilde{v}(z)) - J(v_1(z), \vec{0}))\frac{\partial v}{\partial t}\right) \\ &= \frac{\partial \tilde{v}}{\partial s} + D(v_1(z), \vec{0})\frac{\partial \tilde{v}}{\partial t} + \int_0^1 \frac{d}{d\sigma}\left(J(v_1(z), \sigma\tilde{v}(z))\frac{\partial v}{\partial t}\right) d\sigma. \end{aligned}$$
But recall that $D(v_1(z), \vec{0})$ is an almost complex structure on $\mathbb{C}^{n-1}$, and note that the integral above can, by the chain rule, be expressed as $B(z)\tilde{v}(z)$ for some (complicated) function $B\colon D' \to \mathbb{R}^{(2n-2)\times(2n-2)}$. Hence the function $\tilde{v}\colon D' \to \mathbb{C}^{n-1}$ obeys the hypothesis of Corollary 6.10. Consequently, on a smaller disc D'' around the origin, $\tilde{v}\colon D'' \to \mathbb{C}^{n-1}$ either vanishes identically or vanishes only at 0. Recalling that $v = (v_1, \tilde{v})$ and $u = (u_1, \vec{0})$, in the case that $\tilde{v}|_{D''}$ vanishes only at 0 we clearly have $v(D'') \cap u(D') = \{u(0)\}$. On the other hand if $\tilde{v}$ is identically zero on D'', then since $u_1\colon D'' \to \mathbb{C}$ is an embedding (in particular its image contains a neighborhood of $u_1(0)$) and since $v_1(0) = u_1(0)$, for a suitably small neighborhood V of 0 we will have $v_1(V) \subset u_1(D')$ and so (since $\tilde{v} = \tilde{u} = 0$ in this case) $v(V) \subset u(D')$. Also, since u_1 is an embedding and $v_1(V) \subset u_1(D')$, the restriction of u_1

to $u_1^{-1}(v_1(V))$ has an inverse u_1^{-1}, and setting $\phi = u_1^{-1} \circ v_1$ we obviously have $v_1 = u_1 \circ \phi$, from which it follows that $v = u \circ \phi$ since the $\mathbb{C}^{n-1}$-components of u and v are both 0. To see that ϕ is holomorphic, note that, where as above A is the almost complex structure on $u_1(D') \subset \mathbb{C} \times \{0\}$ obtained by restricting J, the maps u_1 and v_1 are both A-holomorphic, from which it immediately follows that $\phi = u_1^{-1} \circ v$ is holomorphic.

□

Corollary 6.17. *Let (M, J) be an almost complex manifold, let $u\colon \Sigma \to M$ be a J-holomorphic curve, where the surface Σ may not be connected, and let $\Sigma' \subset \Sigma$ be any compact subset of the interior of Σ. Assume that the restriction of u to each connected component of Σ is nonconstant. Then $C := \{z \in \Sigma' | du(z) = 0\}$ is finite. Furthermore, the set*

$$A = \{z \in M | (\exists w \in \Sigma')(w \neq z,\ u(w) = u(z) \text{ and } u(D) \cap u(D') = \{u(z)\} \text{ for some open sets } D, D' \text{ with } w \in D, z \in D')\}$$

is at most countable, and can have accumulation points only at points of C.

In fact, it turns out that points of A also can't accumulate at C, but this is a harder result and we omit the proof (it follows from a result of Micallef and White [MW] which is also proved in [MS2, Appendix E], and Sikorav [Sik] gave a proof which may be a bit more easily digestible based on what we've already done). From this deeper theorem it follows that A is actually finite.

Proof. Since any discrete subset of a compact topological space is finite, the first statement follows directly from Corollary 6.15. For the second statement, suppose that a sequence $\{z_m\}_{m=1}^{\infty}$ in A had a limit $z \in \Sigma' \setminus (A \cup C)$. These z_m would have corresponding points w_m, which (since they belong to the compact set Σ') would have a subsequence converging to some $w \in \Sigma'$. We would have $u(z_m) = u(w_m)$, so by continuity $u(z) = u(w)$. Since we assume that $z \notin C$, u restricts to a neighborhood of z as an embedding. Consequently it must be that $w \neq z$, since if $w = z$ the fact that $u(z_m) = u(w_m)$ would force $w_m = z_m$ once m is large enough for the sequences to enter the neighborhood of z on which u is an embedding. Since $w \neq z$, Theorem 6.16 applies and we find that either $z \in A$ (a contradiction) or that the first alternative in Theorem 6.16 holds for w and z (with what is called u in Theorem 6.16 being the restriction of our given u to a coordinate neighborhood of z, and what is called v in Theorem 6.16 being the restriction of u to a coordinate neighborhood of w). But since $w_m \to w$ and $z_m \to z$, the first alternative of 6.16 for w and z implies the first alternative for w_m and z_m once m is large enough. But the assumption on the z_m shows that instead the second alternative holds for them. Thus any convergent sequence in A has limit either in A or in C; since C is finite this proves that $A \cup C$ is closed. By continuity, the last clause of the corollary follows immediately.

To see that A is countable, note that if Σ'' is any compact subset of $\Sigma' \setminus C$ then $A \cap \Sigma''$ is closed by what we have shown above, and so $A \cap \Sigma''$ is compact. But it follows immediately from the definition of A that A is discrete, so since $A \cap \Sigma''$ is compact it is finite. So since $\Sigma' \setminus C$ can be written as a countable union of compact sets it follows that $A \cap (\Sigma' \setminus C)$ is countable, and hence that A is countable since $A \subset \Sigma'$ and C is finite. □

For a J-holomorphic map $u\colon \Sigma \to M$ where Σ is a closed surface, introduce the following terminology:

- A point $z \in \Sigma$ is called an *injective point* if $u^{-1}\{u(z)\} = \{z\}$ and $du(z) \neq 0$.
- u is called *somewhere injective* if there exists an injective point $z \in \Sigma$.
- u is called a multiple cover if there exists a holomorphic map $\phi\colon \Sigma \to \bar{\Sigma}$ of degree greater than one and a J-holomorphic map $v\colon \bar{\Sigma} \to M$ such that $u = v \circ \phi$.
- u is called simple if it is not a multiple cover.

Clearly if u is somewhere injective it must be simple, since if it were a multiple cover then any point $z \in \Sigma$ would either be a critical point of ϕ and hence of u or else would have $\#u^{-1}\{u(z)\}$ equal to at least the degree of the (branched) covering map ϕ. More interestingly, the converse holds.

Theorem 6.18. *Let $u\colon \Sigma \to M$ be a nonconstant J-holomorphic curve from a closed connected surface Σ to an almost complex manifold (M, J). Then there is a closed surface $\bar{\Sigma}$, a J-holomorphic map $v\colon \bar{\Sigma} \to M$ such that all but countably many points of $\bar{\Sigma}$ are injective points, and a holomorphic map $\phi\colon \Sigma \to \bar{\Sigma}$ such that $u = v \circ \phi$.*

Corollary 6.19. *Any simple J-holomorphic map $u\colon \Sigma \to M$ is somewhere injective, and indeed all but countably many points of Σ are injective points.*

Indeed, if we assume the theorem, ϕ would need to have degree 1 and hence would be invertible, so the asserted property for u would follow from that for v.

Sketch of the proof of Theorem 6.18. Where $A, C \subset \Sigma$ are the sets of Corollary 6.17 (with $\Sigma' = \Sigma$), let $\Sigma_0 = u(\Sigma) \setminus u(A \cup C))$. If $z, w \in u^{-1}(\Sigma_0)$ with $u(z) = u(w)$ then Theorem 6.16 and the fact that neither z nor w lies in $A \cup C$ show that there are neighborhoods U of z and V of w and a holomorphic map $\phi\colon V \to U$ with holomorphic inverse such that $v|_V = u \circ \phi$, and such that $u|_U$ and $v|_V$ are diffeomorphisms onto their images. Consequently $u|_{u^{-1}(\Sigma_0)}\colon u^{-1}(\Sigma_0) \to \Sigma_0$ is a (unbranched) covering map.

Define an equivalence relation $\sim$ on Σ by declaring that $z \sim w$ if and only if, given neighborhoods U_0 of z and V_0 of w, there are open sets U, V with $z \in U \subset U_0$, $w \subset V \subset V_0$ such that $u(V) = u(W)$. The surface $\bar{\Sigma}$ is by definition the quotient $\bar{\Sigma} = \Sigma/\sim$ by this equivalence relation. Where $\phi\colon \Sigma \to \bar{\Sigma}$ is the quotient map, it's clear from the definition that there is a function $v\colon \bar{\Sigma} \to M$ uniquely defined by the property that $u = v \circ \phi$ (set $v(z)$ equal to $u(z')$ for any z' with $\phi(z') = z$, which is independent of the choice of z' by the definition of $\sim$. Note that for $z, w \in u^{-1}(\Sigma_0)$, $z \sim w$ if and only if $u(z) = u(w)$, so $\bar{\Sigma}$ contains Σ_0 as a subset with countable (in fact finite, using the Micallef–White theorem) complement; further by Theorem 6.16 it holds that any point of $\Sigma_0 \subset \bar{\Sigma}$ is an injective point for v.

It remains to show that $\bar{\Sigma}$ carries the structure of an almost complex 2-manifold and that v is J-holomorphic with respect to the (almost) complex structure on $\bar{\Sigma}$; we will only outline this. The coordinate charts and almost complex structure on $\bar{\Sigma} \setminus \phi(C)$ are directly induced from those on Σ: if $x \in \bar{\Sigma} \setminus \phi(C)$ then for any $z, w \in \Sigma$ with $\phi(z) = \phi(w) = x$ there are neighborhoods U, V of z and w to which u restricts to an embedding and such that, for some holomorphic diffeomorphism $\phi\colon V \to U$, $u|_V = u|_U \circ \phi$, and we may use a coordinate chart defined on U (or on V) as a coordinate chart on $\phi(z) = \phi(w) = x$, with the almost complex structure on this coordinate chart in $\bar{\Sigma}$ the same as that from Σ.

The situation is more subtle near $\phi(C)$ and here (at least it seems to me[8]) one needs to appeal to the Micallef–White theorem or something similar ([MW],[Sik]). For $z \in \phi^{-1}(\phi(C))$ we will sketch the construction of a holomorphic coordinate chart near $\phi(z) \in \Sigma$ (of course it will need to be true that the transition functions between the charts for different z mapped to the same point by ϕ are holomorphic). Using the Micallef–White theorem, for a sufficiently small neighborhood of z there will be no points of A; thus on a punctured disc neighborhood D^* of z the restriction $u\colon D^* \to M$ will be a finite covering map onto its image $u(D^*)$. Let m_z be the degree of this covering map. The image of this covering map will be another punctured disc, and so the covering transformation group is $\Gamma_z := \mathbb{Z}/m_z\mathbb{Z}$. Since no point of D^* belongs to $A \cup C$, all of the covering transformations in Γ_z are *holomorphic* diffeomorphisms of D^*, and they extend continuously to maps of $D = D^* \cup \{z\}$ sending z to itself. Hence by the removable singularities theorem in complex analysis it follows that each $\psi \in \Gamma_z$ extends to a holomorphic map $\psi\colon D \to D$ Identifying D with a subset of $\mathbb{C}$ by a coordinate chart around z, the map $w \mapsto \prod_{\psi \in \Gamma_z} \psi(w)$ descends to a map $\phi(D) \to \mathbb{C}$. One can show that this map is a homeomorphism to its image, and (using removable singularities again) that transition maps between two such maps are holomorphic.

Note that ϕ is an immersion on $\Sigma \setminus \phi^{-1}(\phi(C))$, so at points of $\bar{\Sigma} \setminus \phi(C)$ it's easy to see that the fact that u is J-holomorphic implies that v is J-holomorphic. To argue that v is J-holomorphic on all of $\bar{\Sigma}$, one can appeal to a removable singularities theorem for J-holomorphic curves (which we haven't yet proven) stating that a continuous map which is J-holomorphic except at finitely many points is in fact J-holomorphic everywhere. Or one can (perhaps with help from the Micallef–White theorem which gives a normal form for u near the critical points z as above) show that v at least belongs to $W^{1,p}$ for some $p > 1$ and solves the Cauchy–Riemann equation

[8] McDuff and Salamon seem to imply that their version of this result, [MS2, Proposition 2.5.1], doesn't require one to know that A cannot accumulate at C, but it appears to me that without this fact one wouldn't be able to draw some conclusions that they draw in the course of their proof. In particular, if any neighborhood of a critical point z potentially contained infinitely many points of A, then in place of a m_z-fold cover of a once-punctured disc near z we'd have a m_z-fold cover of an infinitely-punctured disc, and the covering transformation group for this cover might have more than m_z elements, which would pose problems for defining the coordinate chart near $\phi(z)$.

weakly, so (since v is also continuous) our smoothness Theorem 6.1 shows that v is a genuine J-holomorphic curve.

□

7. Manifold structures on spaces of J-holomorphic curves

The goal for this section is to prove (at least part of) the following theorem.

Theorem 7.1. *Let (Σ, j) be a closed almost complex 2-manifold of genus g, let (M, ω) be a symplectic manifold, let $A \in H_2(M;\mathbb{Z})$, and let k be a sufficiently large integer. In the space $\mathcal{J}^k_\tau(M, \omega)$ of ω-tame almost complex structures of class C^k, there is a subset $\mathcal{J}^{reg,k}_\tau(M, \omega)$ which contains a countable intersection of open and dense sets with the property that, if $J \in \mathcal{J}^{reg,k}_\tau(M, \omega)$, then*

$$\mathcal{M}^*(\Sigma, J, A) = \{u\colon\ \Sigma \to M | u_*[\Sigma] = A,\ u \text{ is } J\text{-holomorphic and simple}\}$$

is a canonically oriented C^{k-1} manifold of dimension $2(1 - g) + 2\langle c_1(TM), A\rangle$.

Specifically, I will prove here (modulo some background from infinite dimensional differential topology) that $\mathcal{M}(\Sigma, J, A)$ is a manifold of some finite dimension; I won't say much if anything about how to orient it or to compute its dimension.

Recall that $u\colon\ \Sigma \to M$ is J-holomorphic if and only if

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

We will take the view, roughly, that $\bar{\partial}_J$ is a map between (infinite-dimensional) manifolds,[9] will show that, for suitable J, the zero locus $\mathcal{M}^*(\Sigma, J, A)$ is a manifold by virtue of 0 being a regular value of $\bar{\partial}_J$.

This requires first some background about infinite-dimensional manifolds.

7.1. Banach manifolds.

We record here some facts about Banach manifolds; I don't intend to give proofs, but encourage you to fill in the details of the proofs yourself if you are curious, or to consult some reference on the subject ([MS2, Appendix A] contains some of the relevant details; there are various books (for instance [Lan]) on differential topology that give a comprehensive introduction to smooth manifolds without assuming finite-dimensionality, so that their treatment applies to Banach manifolds as well).

Given two normed vector spaces V and W let $\mathcal{B}(V, W)$ be the vector space of bounded linear maps from V to W. Recall that $\mathcal{B}(V, W)$ carries the "operator norm" defined by, for $A \in \mathcal{B}(V, W)$,

$$\|A\| = \sup_{v \in V : \|v\|=1} \|Av\|_W.$$

If W is a Banach space (i.e. is complete with respect to its norm), then $\mathcal{B}(V, W)$ is a Banach space with respect to the operator norm.

Suppose that $\mathcal{F}\colon\ U \to Y$ is some function, where U is an open set in a Banach space X and where Y is also a Banach space. Of course the norms on X and Y induce metrics, and so it makes sense to ask whether the function $\mathcal{F}$ is continuous. Likewise, we shall now observe that it makes sense to ask if $\mathcal{F}$ is differentiable, or indeed C^∞. Indeed, for $x \in X$, we say that $\mathcal{F}$ is differentiable at x if there exists a bounded linear map $(D\mathcal{F})_x\colon\ X \to Y$ such that

$$\lim_{h \to 0} \frac{\|\mathcal{F}(x + h) - \mathcal{F}(x) - (D\mathcal{F})_x h\|_Y}{\|h\|_X} = 0.$$

Assuming that $\mathcal{F}$ is differentiable at every $x \in U$, we then have a map $D\mathcal{F}\colon\ U \to \mathcal{B}(X, Y)$ from an open set in one Banach space to another Banach space. Accordingly we say that $\mathcal{F}\colon\ U \to Y$ is C^1 if $D\mathcal{F}\colon\ U \to \mathcal{B}(X, Y)$ is C^0 (*i.e.*, if it is continuous). Now that we know what it means for a map from one Banach space to another to be C^1, we can say that $\mathcal{F}\colon\ U \to Y$ is C^2 if the map $D\mathcal{F}\colon\ U \to \mathcal{B}(X, Y)$ is C^1, and then, inductively, we say that

[9] or rather, this is true locally; to formulate this globally we will have a vector bundle $\mathcal{E} \to \mathcal{B}^*$ and a section $u \mapsto (u, \bar{\partial}_J u)$ of this vector bundle

$\mathcal{F}: U \to Y$ is C^{r+1} if $D\mathcal{F}: U \to \mathcal{B}(X, Y)$ is C^r. Unsurprisingly, $\mathcal{F}: U \to Y$ is called C^∞ (or *smooth*) if $\mathcal{F}$ is C^r for every natural number r.

Most of the standard results of multivariable differential calculus (for instance the chain rule) continue to hold in this possibly-infinite-dimensional setting.

A C^r *Banach manifold* $\mathcal{B}$ (for some $1 \leq r \leq \infty$) is then defined to be a second countable Hausdorff space equipped with an atlas $\{(U_\alpha, E_\alpha, \phi_\alpha)\}_{\alpha\in A}$, where the U_α form an open cover of $\mathcal{B}$; the E_α are Banach spaces, and the $\phi_\alpha: U_\alpha \to E_\alpha$ are embeddings of open subsets such that the transition functions $\phi_\beta \circ \phi_\alpha^{-1}: \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$ are C^r-diffeomorphisms of open subsets of the Banach spaces E_α, E_β (i.e. they are bijective C^r maps with C^r inverses). Note that the convention here is that the E_α are not (necessarily) all the same Banach space; it's not difficult to show (consider the derivatives of the transition functions) that if $\mathcal{B}$ is connected then the E_α are isomorphic as Banach spaces.[10]

For an example of a Banach manifold, let Σ, M be two finite-dimensional smooth manifolds with Σ compact, choose $p > \dim\Sigma$, and let $\mathcal{B} = W^{1,p}(\Sigma, M)$ (i.e. $\mathcal{B}$ consists of continuous maps which in local coordinates in source and target belong to the Sobolev class $W^{1,p}$; as we've noted before the chain rule implies that this notion doesn't depend on the choice of local coordinate charts). A topology is induced on this space by prescribing a neighborhood base around any given $u \in \mathcal{B}$ by the requirement that $u_n \to u$ if and only if, with respect to all sufficiently small coordinate charts $V \subset \Sigma$ such that $u(V)$ is contained in a coordinate chart W in M, the restriction $u_n|_V$ also has image contained in W for sufficiently large n and (in terms of these coordinates) $u_n|_V \to u|_V$ in $W^{1,p}$.

To actually construct an atlas on $\mathcal{B}$, one can proceed as follows. We will construct data (U_u, E_u, ϕ_u) as in the definition of a Banach manifold for any *smooth* $u: \Sigma \to M$. Choose and fix a Riemannian metric on M. Given $u \in C^\infty(\Sigma, M)$, let $E_u = W^{1,p}(\Sigma, u^*TM)$ be the space of $W^{1,p}$ sections of the bundle $u^*TM \to \Sigma$. (Recall that u^*TM is the bundle whose fiber over $z \in \Sigma$ is the tangent space $T_{u(z)}M$; the implied norm on sections of this bundle may be taken with respect to our chosen Riemannian metric.) Where V_u is a sufficiently small $W^{1,p}$-neighborhood of the zero section of u^*TM, Corollary 4.13 implies that any $\xi \in V_u$ will, for every z, have $|\xi(z)|$ bounded above by the injectivity radius of our fixed Riemannian metric on M. We can then define a map $\psi: V_u \to \mathcal{B}$ by, for $\xi \in V_u$, letting $\psi(\xi)$ be the map in $W^{1,p}(\Sigma, M)$ defined by

$$(\psi(\xi))(z) = exp_{u(z)}(\xi(z))$$

where exp is the exponential map of the Riemannian metric (i.e. $(\psi(\xi))(z)$ is obtained by starting at $u(z)$ and then going out along a geodesic in the direction $\xi(z) \in T_{u(z)}M$). The neighborhood U_u of u is then $\psi(V_u)$, and the chart $\phi_u: U_u \to E_u$ is given by ψ_u^{-1}. It's left to the reader to convince him/herself that the resulting transition maps $\phi_v \circ \phi_u^{-1}$ are smooth maps between open subsets of the Banach spaces $W^{1,p}(\Sigma, u^*TM)$ and $W^{1,p}(\Sigma, v^*TM)$.

Just as in the finite-dimensional context, the implicit function theorem holds in Banach manifolds. Note that the statement below contains references to the tangent space at a point to a Banach manifold (which is always a Banach space), to C^r maps between Banach manifolds, and to the linearization[11] of a C^r map as being a linear map on tangent spaces. I have not formally given definitions of these notions, but leave it to the reader to fill them in—all of this is just a straightforward generalization of the finite-dimensional case.

Theorem 7.2 (Implicit function theorem). *Let $\mathcal{F}: \mathcal{B} \to \mathcal{B}'$ be a C^r map from one Banach manifold to another and suppose $p \in \mathcal{B}'$ has the property that, for every $x \in \mathcal{B}$ with $\mathcal{F}(x) = b$, the linearization $(\mathcal{F}_*)_x: T_x\mathcal{B} \to T_p\mathcal{B}'$ has a bounded right inverse. Then $\mathcal{F}^{-1}(\{p\})$ is a C^r Banach manifold, and $T_x\mathcal{F}^{-1}(\{p\}) = \ker((\mathcal{F}_*)_x)$.*

You can find a proof of the implicit function theorem for Banach manifolds in [MS2, Section A.3]; the approach of the proof is sufficiently similar to that used in finite dimensions that it might be more instructive

[10]As is customary, the convention is that an isomorphism of Banach spaces need not be an isometry—just a bounded linear bijection (which will necessarily have a bounded inverse)

[11]I'll tend to use the word "linearization" where customary usage might call for "derivative" when discussing maps between Banach manifolds, simply because finite-dimensional derivatives will often simultaneously be appearing in their own right and I want to keep the two separate.

to review the proof of the implicit function theorem in your favorite differential topology or analysis book and persuade yourself that that proof generalizes to Banach manifolds.

By definition, $p \in \mathcal{B}'$ is called a *regular value* of $\mathcal{F}$ if it satisfies the hypothesis of the implicit function theorem above.

Remark 7.3. Of course, if $(\mathcal{F}_*)_x$ has a bounded linear right inverse then it is surjective. Conversely, a surjective linear map between vector spaces always has a linear right inverse (this is an exercise in the use of the axiom of choice), but when there are norms present it's not always possible to arrange for this right inverse to be bounded. I don't know any examples where this is obvious; however this sort of question has been studied for approximately as long as people have been thinking about Banach spaces and so there are many examples that can be found in the literature. In this direction, note that the question, "Is there a bounded surjection $D\colon A \to B$ from one Banach space to another without a bounded right inverse" is essentially equivalent to the question, "Is there a closed subspace $V \le A$ of a Banach space which does not have a closed complement, i.e. such that there is no closed $W \le A$ such that $A = V \oplus W$." (For one direction, given $V \le A$, let $B = A/V$ (this carries a complete norm given by assigning to a coset in A/V the minimal distance between any of its representatives and the subspace V), and let $D\colon A \to A/V$ be the projection; if D had a bounded right inverse Q (so $DQ = I$) then $Im(Q)$ would be a complement to V, and would be closed since $Im(Q) = \ker(QD - I)$.) One relatively simple example (found by Phillips in 1940) of a closed subspace V of a Banach space A without a closed complement is given by letting $A = l^\infty$ be the space of bounded sequences of real numbers with the sup norm and letting $V = c_0$ be the subspace consisting of sequences which converge to 0. There is a proof that these have the claimed property in [Meg, Theorem 3.2.20].

All that said, in practice, for the particular maps that we will encounter the hard part will be showing surjectivity and then the existence of a bounded right inverse will quickly follow; this is basically because a bounded linear surjection with a finite dimensional kernel automatically has a bounded right inverse (exercise; use the Hahn-Banach theorem).

7.2. **The linearization of the Cauchy–Riemann equation.** We now fix a closed almost complex 2-manifold (Σ, j), a $2n$-dimensional symplectic manifold (M, ω), and a real number $p > 2$. As above let $\mathcal{B} = W^{1,p}(\Sigma, M)$, and let $\mathcal{B}^*$ be the open subset of $\mathcal{B}$ consisting of maps $u\colon \Sigma \to M$ which are simple (i.e., such that there do not exist maps $\phi\colon \Sigma \to \Sigma'$ and $v\colon \Sigma' \to M$ where ϕ is a holomorphic map of degree larger than 1 and $u = v \circ \phi$).

Given $J \in \mathcal{J}^k_\tau(M, \omega)$, we intend to introduce a vector bundle $\mathcal{E} \to \mathcal{B}^*$ such that the Cauchy–Riemann operator $\bar{\partial}_J$ (or, more correctly, the map $u \mapsto (u, \bar{\partial}_J u)$) can be interpreted as a section of the bundle $\mathcal{E}$, which moreover has a chance of having a surjective linearization so that the implicit function theorem will apply. To do this it is of course necessary to decide what the fibers of the bundle will be—the fiber $\mathcal{E}_u$ over $u \in \mathcal{B}^*$ should be a space in which $\bar{\partial}_J u$ lives.

Now

$$\bar{\partial}_J u = \frac{1}{2}(du + J \circ du \circ j)$$

is, at any point $z \in \Sigma$, a $\mathbb{R}$-linear map $(\bar{\partial}_J u)(z)\colon T_z\Sigma \to T_{u(z)}M$. Recalling the pullback bundle $u^*TM \to \Sigma$ with fibers $(u^*TM)_z = T_{u(z)}M$, $\bar{\partial}_J u$ can thus be viewed as a bundle map $T\Sigma \to u^*TM$ of bundles over Σ (i.e. as a map between the total spaces which restricts to each fiber $T_z\Sigma$ as a homomorphism to $(u^*TM)_z$). However $\bar{\partial}_J u$ satisfies a constraint beyond this, namely for $v \in T_z\Sigma$

$$\bar{\partial}_J u(jv) = \frac{1}{2}(du \circ j - J \circ du)(v) = -J\left(\frac{1}{2}(du + J \circ du \circ j)\right)(v) = -J(\bar{\partial}_J u)(v).$$

We thus have

$$\bar{\partial}_J u \in \overline{Hom}_J(T\Sigma, u^*TM)$$

for any u, where the right hand side denotes bundle maps from $T\Sigma \to u^*TM$ which are complex antilinear with respect to the almost complex structures j on $T\Sigma$ and J on u^*TM. If u is assumed to be of class $W^{1,p}$, then $\bar{\partial}_J u$ will be of class L^p.

Accordingly we form a bundle $\mathcal{E} \to \mathcal{B}^*$ whose fiber over $u \in \mathcal{B}^*$ is the space $\mathcal{E}_u = L^p(\overline{Hom}_J(T\Sigma, u^*TM))$ of L^p, complex antilinear bundle maps $T\Sigma \to u^*TM$. I won't go into the details of showing that this is genuinely a (appropriately defined) Banach vector bundle beyond noting that to form the appropriate local trivializations one needs to, for $v \in \mathcal{B}^*$ sufficiently close to $u \in \mathcal{B}^*$, identify the space $L^p(\overline{Hom}_J(T\Sigma, v^*TM))$ with $L^p(\overline{Hom}_J(T\Sigma, u^*TM))$, and that this can be done by composing a bundle map $T\Sigma \to v^*TM$ (i.e. a collection of maps $T_z\Sigma \to T_{v(z)}M$) with parallel translation along shortest geodesics from $v(z)$ to $u(z)$ for all z in order to get a bundle map $T\Sigma \to u^*TM$. Here we use the Riemannian metric g_J associated in the usual way to $J \in \mathcal{J}^k_\tau(M, \omega)$.

With the bundle $\mathcal{E} \to \mathcal{B}^*$ constructed (set-theoretically, $\mathcal{E} = \{(u, \eta) | \eta \in L^p(\overline{Hom}_J(T\Sigma, u^*TM))\}$), we have a section $s\colon \mathcal{B}^* \to \mathcal{E}$ sending u to $(u, \bar{\partial}_J u)$. The moduli space $\mathcal{M}^*(\Sigma, J, A)$ which we wish to be a manifold consists of those $u \in \mathcal{B}^*$ such that $u_*[\Sigma] = A$ and $\bar{\partial}_J u = 0$. $\mathcal{M}^*(\Sigma, J, A)$ will be a submanifold of $\mathcal{B}^*$ provided that, for every $u \in \mathcal{M}^*(\Sigma, J, A)$ there is a neighborhood $\mathcal{U}$ of u such that $\mathcal{M}^*(\Sigma, J, A) \cap \mathcal{U}$ is a submanifold of $\mathcal{B}^*$. Taking $\mathcal{U}$ equal to the one of the trivializing neighborhoods for $\mathcal{E}$ as in the previous paragraph (so for $v \in \mathcal{U}$ we have a linear identification $\mathcal{E}_v \cong \mathcal{E}_u$ given by parallel translation), in terms of the trivialization $\bar{\partial}_J$ restricts to $\mathcal{U}$ as a map

$$\bar{\partial}_J\colon \mathcal{U} \to \mathcal{E}_u = L^p(\overline{Hom}_J(T\Sigma, u^*TM)).$$

Of course, parallel translation can also be used to show that all of the maps in $\mathcal{U}$ are homotopic (so represent the same homology class A). Together with the implicit function theorem, this establishes that:

Proposition 7.4. *$\mathcal{M}^*(\Sigma, J, A)$ will be a C^r manifold provided that the following holds: for every $u \in \mathcal{M}^*(\Sigma, J, A)$ and for sufficiently small trivializing neighborhoods $\mathcal{U}$ of u for the bundle $\mathcal{E} \to \mathcal{B}^*$, the map $\bar{\partial}_J\colon \mathcal{U} \to \mathcal{E}_u$ is of class C^r and its linearization at u has a bounded right inverse.*

Accordingly we should fix a J-holomorphic map $u\colon \Sigma \to M$ with a trivializing neighborhood $\mathcal{U} \subset \mathcal{B}^*$ and consider the linearization D^J_u of the $\mathcal{E}_u$-valued map $v \mapsto \bar{\partial}_J v$. Now the tangent space to $\mathcal{U}$ at u is equal to the space $W^{1,p}(\Sigma, u^*TM)$ of sections of the pullback bundle u^*TM; namely, to obtain a path $t \mapsto u_t$ in $\mathcal{U}$ with $u_0 = u$ and $\frac{d}{dt}u_t\big|_{t=0} = \xi \in W^{1,p}(\Sigma, u^*TM)$ we can set

$$u_t(z) = exp_{u(z)}(t\xi(z)).$$

The linearization $D^J_u\colon W^{1,p}(\Sigma, u^*TM) \to \mathcal{E}_u$ of $\bar{\partial}_J$ at u is then given by

$$\begin{aligned} D^J_u\xi &= \frac{d}{dt}\bar{\partial}_J u_t\Big|_{t=0} = \frac{d}{dt}\left(d(exp_u(t\xi)) + J(exp_u(t\xi))d(exp_u(t\xi)) \circ j\right)\Big|_{t=0} \\ &= \nabla\xi + J(u)\nabla\xi \circ j + (\nabla_\xi J) \circ du \circ j. \end{aligned}$$

(Here, for instance, $\nabla\xi \circ j$ should be interpreted as the map $T\Sigma \to u^*TM$ defined by $(\nabla\xi \circ j)(v) = \nabla_{u_* jv}\xi$. Also, the path $\bar{\partial}_J u_t$ is regarded as living in the fixed vector space $\mathcal{E}_u$ via the parallel translation that identifies each $\mathcal{E}_{u_t}$ with $\mathcal{E}_u$.)

Now the bundle $u^*TM \to \Sigma$ carries an almost complex structure $J(u)$, and so it admits local trivializations $u^*TM|_U \cong U \times \mathbb{R}^{2n}$ in terms of which the almost complex structure $J(u)$ appears as the standard almost complex structure J_0 on $\mathbb{R}^{2n}$ In terms of such a local trivialization (along with a local complex trivialization for $T\Sigma$) the above formula for D^J_u reduces to something with the form

$$\begin{aligned} D^J_u|_U\colon W^{1,p}(U; \mathbb{C}) &\to L^p(U; \overline{Hom}_{J_0}(\mathbb{C}; \mathbb{C}^n)) \\ \xi &\mapsto (\partial_{\bar{z}}\xi + A(z)\xi(z))\, d\bar{z} \end{aligned}$$

for some matrix-valued function A which incorporates the partial derivatives of J (contributed by the term $\nabla_\xi J$ in the formula for D^J_u) as well as some terms involving the Christoffel symbols of the metric arising from switching from covariant derivatives of ξ to derivatives with respect to coordinates in local trivializations. Note that A is a smooth function if J is; a bit more generally if J is C^r for some $r \geq 1$ (which implies that u is as well, as you can check by reexamining the proofs in Section 6) then A is C^{r-1}, which in turn implies that D^J_u is a C^{r-1} map of Banach spaces.

Choose a finite cover $U_1, \dots, U_k$ of Σ by trivializing charts in which the above local form holds, i.e. so that upon restriction to the various U_i we have $D_u^J\xi = (\partial_{\bar{z}}\xi + A_i(z)\xi)\,d\bar{z}$ for some matrix-valued function A_i on U_i (which may be assumed to be bounded and C^0 if J is C^1). Choose $\epsilon > 0$ such that (with respect to some fixed metric on Σ) such that for any $z \in \Sigma$ we have $B_\epsilon(z) \subset U_i$ for some i, and for each i let $V_i = \{z \in \Sigma | B_\epsilon(z) \subset U_i\}$. The V_i continue to cover Σ, with $\bar{V}_i \subset U_i$. Hence Theorem 5.10 gives constants C_i with

$$\|\xi\|_{W^{1,p}(V_i)} \leq C_i \left(\|\partial_{\bar{z}}\xi\|_{L^p(U_i)} + \|\xi\|_{L^p(V_i)}\right) \leq C_i \left(\|D_u^J\xi\|_{L^p(U_i)} + (1 + \|A_i\|_{C^0})\|\xi\|_{L^p(U_i)}\right).$$

The $\|A_i\|_{C^0}$ are constants, and we have $\|\xi\|_{W^{1,p}(\Sigma)} \leq \sum_{i=1}^k \|\xi\|_{W^{1,p}(V_i)}$, so this proves:

Proposition 7.5. *Assuming that J is C^1, for $u \in \mathcal{M}^*(\Sigma, J, A)$ the linearization $D_u^J\colon W^{1,p}(\Sigma, u^*TM) \to \mathcal{E}_u$ of $\bar{\partial}_J$ at u obeys, for some constant C depending on u but not on ξ,*

$$\|\xi\|_{W^{1,p}} \leq C\left(\|D_u^J\xi\|_{L^p} + \|\xi\|_{L^p}\right).$$

We wish for D_u^J to be surjective; ultimately we will show this (when it in fact holds) by proving that its range is closed on the one hand and dense on the other. As is probably not obvious, Proposition 7.5 implies that, regardless of J, D_u^J alway has closed range, as will follow from the functional analytic lemma below. To prepare for it, recall that a linear operator $K\colon X \to \mathbb{Z}$ from one Banach space to another is called *compact* if whenever $\{x_i\}_{i=1}^\infty$ is a bounded sequence in X it holds that the sequence $\{Kx_i\}_{i=1}^\infty$ in Z has a ocnvergent subsequence. Recall also that Theorem 4.16 (or rather a straightforward generalization of it which has essentially the same proof) shows that the inclusion $W^{1,p}(\Sigma; u^*TM) \hookrightarrow L^p(\Sigma; u^*TM)$ is a compact operator; in view of this Proposition 7.5 shows that the following lemma applies to the operator $D = D_u^J$.

Lemma 7.6. *Suppose that X, Y, Z are Banach spaces, and that $D\colon X \to Y$ is a bounded linear operator such that there exists a constant C and a compact operator $K\colon X \to Z$ such that, for all $x \in X$,*

$$\|x\|_X \leq C\left(\|Dx\|_Y + \|Kx\|_Z\right). \tag{31}$$

Then $\ker D$ *is finite dimensional and $Im(D) \leq Y$ is closed.*

Proof. Let $\{x_m\}_{m=1}^\infty$ be a sequence in $\ker D$ such that $\|x_m\|_X \leq 1$ for all m. Then the fact that K is compact shows that for some subsequence $\{x_{m_i}\}_{i=1}^\infty$ the sequence $\{Kx_{m_i}\}_{i=1}^\infty$ is Cauchy in Z. But since $Dx_{m_i} = 0$ for all i, (31) shows that

$$\|x_{m_i} - x_{m_k}\|_X \leq C\|Kx_{m_i} - Kx_{m_k}\|_Z,$$

and so $\{x_{m_i}\}_{i=1}^\infty$ is a Cauchy sequence in X, and therefore has a convergent subsequence. $\ker D$ is thus a closed subspace of X (hence a Banach space with respect to the subspace norm) with the property that its closed unit ball is compact; this implies that $\ker D$ is finite-dimensional.[12]

Now that we know that $\ker D$ is finite-dimensional, as noted at the end of Remark 7.3 it follows from the Hahn-Banach theorem that $\ker D$ has a closed complement, i.e. that there is a closed subspace $X' \leq X$ such that $X = X' \oplus \ker D$. Then X' is a Banach space with respect to the subspace norm, and where $D' = D|_{X'}$ the operator $D'\colon X' \to Y$ is an *injective* bounded linear operator with $Im(D') = Im(D)$. Obviously (31) continues to hold with X, D replaced by X', D'.

So we now show that $Im(D')$ (which as noted earlier is equal to $Im(D)$) is closed. We thus need to show that if we have a sequence $\{x_m\}_{m=1}^\infty$ of elements of X' such that $D'x_m \to y \in Y$, then $y \in Im(D')$. Suppose first that some subsequence $\{x_{m_i}\}_{i=1}^\infty$ of $\{x_m\}_{m=1}^\infty$ has $\|x_{m_i}\|_X$ bounded. The compactness of K shows that $\{Kx_{m_i}\}_{m=1}^\infty$ has a convergent subsequence $\{Kx_{m_{i_j}}\}$, and by assumption we have $D'x_{m_{i_j}} \to y$. Hence (31) implies that $\{x_{m_{i_j}}\}$ is a Cauchy sequence in X', whose limit x obeys $D'x = y$ by the boundedness of D'.

There remains the case where no subsequence of the $\{x_m\}_{m=1}^\infty$ is bounded; however we shall derive a contradiction in this case. Thus we are assuming that $\|x_m\|_X \to \infty$, with $x_m \in X'$. Let $a_m = \frac{x_m}{\|x_m\|_X}$, so $\|a_m\|_X = 1$ and since $D'x_m \to y$ we have $D'a_m \to 0$. The compactness of K shows that the Ka_m have a convergent subsequence, and

[12] See this Wikipedia page on Riesz's Lemma for two proofs of the fact that a normed vector space whose closed unit ball is compact must be finite-dimensional.

then since $D'a_m$ also converges it follows from (31) that $\{a_m\}_{m=1}^{\infty}$ has a convergent subsequence, say converging to a. Since $a_m \in X'$ and X' is closed we have $a \in X'$, while since $\|a_m\|_X = 1$ we have $\|a\|_X = 1$. But since $D'a_m \to 0$ we would have $D'a = 0$, and this contradicts the injectivity of D'.

So in fact if $x_m \in X'$ with $D'x_m \to y$ the x_m have a bounded subsequence, and so the paragraph before last shows that $y \in Im(D') = Im(D)$. This completes the proof that $Im(D)$ is closed.

□

Our goal is to show that, for generic $J \in \mathcal{J}^k_\tau(M,\omega)$, the operator $D^J_u\colon\ W^{1,p}(\Sigma; u^*TM) \to \mathcal{E}_u$ is surjective for every $u \in \mathcal{M}^*(\Sigma, J, A)$. Proposition 7.5 and Lemma 7.6 establish that $Im(D^J_u) \leq \mathcal{E}_u$ is always closed. It then follows from the Hahn-Banach theorem that D^J_u is surjective provided that

$$Ann(Im(D^J_u)) := \{\eta \in \mathcal{E}^*_u | \eta|_{Im(D^J_u)} = 0\} = \{0\}.$$

(Here $\mathcal{E}^*_u$ denotes as usual the space of *bounded* linear functionals on $\mathcal{E}_u$. To deduce this from the Hahn-Banach theorem, note that if $Im(D^J_u)$ were a proper closed subspace of $\mathcal{E}_u$ we could choose $x \in \mathcal{E}_u \setminus Im(D^J_u)$; define a bounded linear functional ℓ on $V := Im(D^J_u) \oplus \langle x\rangle$ by $\ell(tx + z) = t$ for $t \in \mathbb{R}, z \in Im(D^J_u)$; and then use the Hahn-Banach theorem to extend the domain of ℓ from V to all of $\mathcal{E}_u$.)

Recall that $\mathcal{E}_u = L^p(\overline{Hom}_J(T\Sigma, u^*TM))$; thus in appropriate trivializations an element $x \in \mathcal{E}_u$ has the form $x|_U = f(z)d\bar{z}$ where $f \in L^p(U;\mathbb{C}^n)$ for some coordinate chart $U \subset \Sigma$. Thus an element $\eta \in \mathcal{E}^*_u$ can be locally represented as $\eta|_U = g(z)d\bar{z}$ where $g \in L^q(U;\mathbb{C}^n)$ and $\frac{1}{p} + \frac{1}{q} = 1$; where $\langle\cdot,\cdot\rangle$ is the standard ($\mathbb{C}$-valued) inner product on $\mathbb{C}^n$, the pairing between $\mathcal{E}_u$ and $\mathcal{E}^*_u$ is given by (for at least for $x \in \mathcal{E}_u$ supported in one of these open sets U),

$$\eta(x) = Re\left(\frac{i}{2}\int_U \overline{g(z)}f(z)dz \wedge d\bar{z}\right).$$

Suppose that $\eta|_{Im(D^J_u)} = 0$ and fix a trivializing neighborhood U as in the previous paragraph. For any $\xi \in W^{1,p}(\Sigma, u^*TM)$ such that $supp(\xi) \subset U$ (so ξ can be viewed as a $W^{1,p}$ function $U \to \mathbb{C}^n$, recall that $D^J_u\xi = (\partial_{\bar{z}}\xi + A(z)\xi(z))d\bar{z}$ for a bounded continuous matrix-valued function A. So for any such ξ we have (again assuming that η is represented within U as $g(z)d\bar{z}$)

$$0 = \eta(D^J_u\xi) = Re\left(\frac{i}{2}\int_U \bar{g}(z)(\partial_{\bar{z}}\xi + A\xi)dz \wedge d\bar{z}\right).$$

It follows that $\bar{g} \in L^q(U;\mathbb{C}^n)$ is a weak solution to the equation $-\partial_{\bar{z}}\bar{g} + A(z)^T\bar{g}(z) = 0$. It then follows that $\partial_{\bar{z}}\bar{g} \in L^q(U;\mathbb{C}^n)$, so by Theorem 5.10 $\bar{g} \in W^{1,q}(U';\mathbb{C}^n)$ for any open set U' compactly contained in U. Assuming that J is C^2, so that A is C^1, we then have $A\bar{g} \in W^{1,q}(U';\mathbb{C}^n)$, so $\partial_{\bar{z}}\bar{g} \in W^{1,q}(U';\mathbb{C}^n)$, and then Theorem 5.10 shows that $\bar{g} \in W^{2,q}(U'';\mathbb{C}^n)$ for any open U'' compactly contained in U'. Hence Theorem 4.17 shows that, where $p^* = \frac{2q}{2-q} > 2$, we have $\bar{g} \in W^{1,p^*}(U'';\mathbb{C}^n)$.

Since there is a finite cover of Σ by sets U as above, this implies the following:

Proposition 7.7. *Assuming that J is C^2, any $\eta \in Ann(Im(D^J_u)) \subset \mathcal{E}^*_u$ belongs to $W^{1,p^*}\left(\Sigma; \left(\overline{Hom}_J(T\Sigma; u^*TM)\right)^*\right)$, and obeys $(D^J_u)^*\eta = 0$ where $(D^J_u)^*\colon\ W^{1,p^*}\left(\Sigma; \left(\overline{Hom}_J(T\Sigma; u^*TM)\right)^*\right) \to L^{p^*}\left(\Sigma; \left(\overline{Hom}_J(T\Sigma; u^*TM)\right)^*\right)$ is an operator whose which in suitable local trivializations (in which $\eta(z) = g(z)d\bar{z}$) has the coordinate expression $(D^J_u)^*(g(z)d\bar{z}) = \partial_z g + B(z)g(z)$ for a continuous matrix-valued function B.*

(Indeed, this follows from the discussion above the proposition since $\partial_{\bar{z}}\bar{g} = \overline{\partial_z g}$).

Now the same discussion that led to the conclusion (via Theorem 5.10 and Lemma 7.6) that D^J_u has closed range and finite dimensional kernel also shows that $(D^J_u)^*$ has closed range and finite-dimensional kernel, since its local coordinate expressions have the same basic form (modulo complex conjugation). Thus since $Ann(Im(D^J_u)) \leq \ker(D^J_u)^*$, this proves that $Ann(Im(D^J_u))$ is finite-dimensional. If $\eta_1, \ldots, \eta_k$ is a basis for $Ann(Im(D^J_u))$, then we may find $x_1, \ldots, x_k \in \mathcal{E}_u$ such that $\eta_i(x_j) = \delta_{ij}$. Then $Im(D^J_u) \oplus span\{x_1, \ldots, x_k\}$ is a closed subspace of $\mathcal{E}_u$ with trivial annihilator, and so is all of $\mathcal{E}_u$ by the Hahn-Banach theorem. This proves that the cokernel of D^J_u, i.e. $\frac{\mathcal{E}_u}{Im(D^J_u)}$, has finite dimension k. We have now proven:

Theorem 7.8. *If J is C^2 and if $u \in \mathcal{M}^*(\Sigma, J, A)$, then the linearization $D_u^J\colon W^{1,p}(\Sigma; u^*TM) \to \mathcal{E}_u$ is a* ***Fredholm*** *operator, i.e. it is a bounded linear operator with closed range, finite-dimensional kernel, and finite-dimensional cokernel.*

By definition the *index* of a Fredholm operator $D\colon X \to Y$ is the difference $ind(D) = \dim(\ker D) - \dim(\operatorname{coker} D)$. Notice that if X and Y are finite-dimensional, any linear operator D from X to Y is Fredholm, and that the rank-plus-nullity theorem shows that $ind(D) = \dim X - \dim Y$, independently of D. In the infinite-dimensional situation, of course not every operator is Fredholm, and it's not true that all Fredholm operators have the same index; however it *is* true that the Fredholm index is constant on any given connected component of the space of Fredholm operators. In particular, if we consider a path of Fredholm operators $D_t\colon X \to Y$, although the individual terms $\dim(\ker(D_t))$ and $\dim(\operatorname{coker}(D_t))$ will likely change as t varies, the difference $ind(D_t)$ will remain constant. See [Mr] for a concise treatment of these and other facts about Fredholm operators.

The Atiyah-Singer index theorem gives a formula for the indices of many naturally-occurring differential Fredholm operators in geometry in terms of topological data; in our case it can be read as asserting that:

Theorem 7.9. *If $u_*[\Sigma] = A$, $\dim M = 2n$, and if Σ has genus g, then the Fredholm operator $D_u^J\colon W^{1,p}(\Sigma; u^*TM) \to \mathcal{E}_u$ has index*

$$ind(D_u^J) = 2n(1-g) + 2\langle c_1(TM), A\rangle.$$

One essentially self-contained proof of this is given in [MS2, Appendix C]; in light of the stability properties of the indices of Fredholm operators mentioned earlier this can also be seen as a special case of extensions of the Riemann-Roch theorem in algebraic geometry due to Weil or to Hirzebruch (in particular we're using significantly less than the full strength of the Atiyah-Singer theorem here).

7.3. **Generic surjectivity.** Having shown that D_u^J is always Fredholm when J is C^k with $k \geq 2$ and $u \in \mathcal{M}^*(\Sigma, J, A)$, we will now show that (if k is large enough) for generic J it holds that D_u^J is surjective for all $u \in \mathcal{M}^*(\Sigma, J, A)$. We will first show that (for $k \geq 2$) the "universal moduli space"

$$\tilde{\mathcal{M}}^{*,k}(\Sigma, A) = \{(u, J) \in \mathcal{B}^* \times \mathcal{J}_\tau^k(M, \omega) | \bar{\partial}_J u = 0\}$$

is a manifold. In this direction, we form a bundle $\tilde{\mathcal{E}} \to \mathcal{B}^* \times \mathcal{J}_\tau^k(M, \omega)$ whose fiber over (u, J) is, as before, the space $\mathcal{E}_{u,J} = L^p(\overline{Hom}_J(T\Sigma, u^*TM))$. We have a section $(u, J) \mapsto \bar{\partial}_J u$ of this bundle, which works out (see [MS2, p. 48]) to be of class C^{k-1}.

Proposition 7.10. *If $(u, J) \in \tilde{\mathcal{M}}^{*,k}(\Sigma, A)$, then the linearization*

$$\mathcal{D}_{u,J}\colon T_{(u,J)}(\mathcal{B}^* \times \mathcal{J}_\tau^k(M, \omega)) \to \mathcal{E}_{u,J}$$

of the map $(u' J') \mapsto \bar{\partial}_{J'} u'$ at (u, J) is surjective, and in fact has a bounded right inverse.

Proof. Note that

$$\begin{aligned} T_{(u,J)}(\mathcal{B}^* \times \mathcal{J}_\tau^k(M, \omega)) &= T_u\mathcal{B}^* \oplus T_J\mathcal{J}_\tau^k(M, \omega) \\ &= W^{1,p}(\Sigma; u^*TM) \oplus \{Y \in C^k(TM, TM) | JY + YJ = 0\} \end{aligned}$$

(the equation $JY + YJ = 0$ arises from the fact that $J + \epsilon Y$ will continue to be an ω-tame almost complex structure to leading order in ϵ provided that $JY + YJ = 0$). Since $\bar{\partial}_{J'} u' = \frac{1}{2}(du + J(u) \circ du \circ j)$, the linearization $\mathcal{D}_{u,J}$ is given by the formulas

$$\mathcal{D}_{u,J}(\xi, 0) = D_u^J \xi$$

$$\mathcal{D}_{u,J}(0, Y) = \frac{1}{2} Y(u) \circ du \circ j$$

In particular the image of $\mathcal{D}_{u,J}$ contains the finite codimension, closed subspace $Im(D_u^J)$; hence $Im(\mathcal{D}_{u,J})$ is also closed in $\mathcal{E}_{u,J}$ with finite codimension (for instance because, where $\pi\colon \mathcal{E}_{u,J} \to \mathcal{E}_{u,J}/Im(D_u^J)$ is the projection, we have $Im(\mathcal{D}_{u,J}) = \pi^{-1}(\pi(Im(\mathcal{D}_{u,J})))$ and $\pi(Im(\mathcal{D}_{u,J}))$ is a subspace of a finite-dimensional space and so is closed with finite codimension). So by the Hahn-Banach theorem, in order to show that $\mathcal{D}_{u,J}$ is surjective it is enough to show that $Ann(Im(\mathcal{D}_{u,J})) = \{0\}$.

Since $Im(D^u_J) \leq Im(\mathcal{D}_{u,J})$, any element η of $Ann(Im(\mathcal{D}_{u,J}))$ belongs to $Ann(Im(D^J_u))$. In particular Proposition 7.7 implies that any such η is continuous. Suppose then that η is a not-identically-zero element of $Ann(Im(\mathcal{D}_{u,J}))$. Now $u\colon \Sigma \to M$ is simple and J-holomorphic, so Corollary 6.17 shows that there is an open, dense set $V \subset \Sigma$ such that for all $z \in V$ the map $du(z)\colon T_z\Sigma \to T_{u(z)}M$ is injective and $u^{-1}(\{u(z)\}) = \{z\}$. Choose a small open subset $U \subset V$ to which η restricts (with respect to appropriate trivializations) as $\eta|_U = g(z)d\bar{z}$ with $g\colon U \to \mathbb{C}^n$ nonvanishing on U. Let $W \subset M$ be an open subset such that $W \cap u(\Sigma) = u(U)$. Then since $u|_U$ is injective, since $du(z)$ is injective at each $z \in U$, and since our only constraint on Y is that it should anti-commute with J, we can choose Y to be supported in W and to have the property that $Y(u) \circ du \circ j = \beta(z)\eta(z_0)$ where β a nonnegative function supported in a small neighborhood of some $z_0 \in W$. Since (because all points of U are injective points for u) $u^{-1}(W) = U$, we will have

$$\eta(\mathcal{D}_{u,J}(0, Y)) = \frac{1}{2}Re\left(\int_U \beta(z)\overline{g(z)}g(z_0)\right) > 0$$

as long as the support of β is chosen to be a small enough neighborhood of z_0 that $Re(\overline{g(z)}g(z_0))$ does not change sign on this neighborhood; this is possible by the continuity of η. But this contradicts the assumption that $\eta \in Ann(Im(\mathcal{D}_{u,J}))$. Since the only assumption on η was that it was nonzero, this completes the proof that $\mathcal{D}_{u,J}$ is surjective at every $(u, J) \in \tilde{\mathcal{M}}^*(\Sigma, A)$.

To show that $\mathcal{D}_{u,J}$ has a bounded right inverse, note that since $\ker D^J_u$ is finite-dimensional kernel it has a closed complement $A \leq W^{1,p}(\Sigma; u^*TM)$. Choose $y_1, \dots, y_m \in \mathcal{E}_{u,J}$ to be a linearly independent set such that $Im(D^J_u) \oplus span\{y_i\} = \mathcal{E}_{u,J}$, and choose $Y_1, \dots, Y_m \in T_J\mathcal{J}^k_\tau(M, \omega)$ such that $\mathcal{D}_{u,J}(0, Y_i) = y_i$. Then

$$\tilde{A} = \left\{\left(a, \sum_{i=1}^m t_iY_i\right)|a \in A, t_i \in \mathbb{R}\right\}$$

is a closed subspace of $T_{(u,J)}\mathcal{B}^* \times \mathcal{J}^k_\tau(M, \omega)$ to which $\mathcal{D}_{u,J}$ restricts as a bounded linear isomorphism onto $\mathcal{E}_{u,J}$. The map $\mathcal{E}_{u,J} \to T_{(u,J)}\mathcal{B}^* \times \mathcal{J}^k_\tau(M, \omega)$ which sends η to the unique element x of $\tilde{A}$ having the property that $\mathcal{D}_{u,J}x = \eta$ will then be the desired right inverse; that this map is bounded follows immediately upon applying the open mapping theorem to $\mathcal{D}_{u,J}|_{\tilde{A}}$.

□

In particular it follows immediately from the implicit function theorem that $\tilde{\mathcal{M}}^{*,k}(\Sigma, A)$ is a C^{k-1} Banach manifold (for any choice of the integer $k \geq 2$).

Definition 7.11. *A C^1 map $\mathcal{F}\colon \mathcal{B} \to \mathcal{B}'$ from one Banach manifold to another is called* Fredholm *if, for every $x \in \mathcal{B}$, the linearization $(\mathcal{F}_*)_x\colon T_x\mathcal{B} \to T_{f(x)}\mathcal{B}'$ is a Fredholm operator.*

We have shown that $\tilde{\mathcal{M}}^{*,k}(\Sigma, A)$ is a C^{k-1}-Banach manifold (for $k \geq 2$ at least). Also, the space $\mathcal{J}^k_\tau(M, \omega)$ of ω-tame almost complex structures of class C^k is a Banach manifold, essentially because the C^k norm on endomorphisms of TM is complete[13] Consider the map

$$\pi\colon \tilde{\mathcal{M}}^{*,k}(\Sigma, A) \to \mathcal{J}^k_\tau(M, \omega)$$
$$(u, J) \mapsto J,$$

which is of class C^{k-1}.

[13]Note by contrast that the space of C^∞ endomorphisms of TM, or more generally the space of C^∞ sections of any vector bundle, is not a Banach space with respect to any standard norm; this technical point is why we've switched from considering smooth almost complex structures as we did earlier in the course to considering C^k almost complex structures. Note that although the results from Section 6.1 were formulated under the assumption that J was C^∞ in order to make their statements concise, their proofs generally extend to the case where J is just C^k for some finite positive k, and show for instance that a J-holomorphic curve for a C^k almost complex structure J will be of class $W^{k+1,p}$ for all $p < \infty$ and so in particular will be C^k.

Proposition 7.12. *The projection map $\pi\colon \tilde{\mathcal{M}}^{*,k}(\Sigma, A) \to \mathcal{J}^k_\tau(M,\omega)$ is Fredholm, and at any $(u, J) \in \tilde{\mathcal{M}}^{*,k}(\Sigma, A)$ we have $\dim\ker(\pi_*)_{(u,J)} = \dim\ker D^J_u$ and $\dim coker(\pi_*)_{(u,J)} = \dim coker D^J_u$. Consequently whenever $J \in \mathcal{J}^k_\tau(M,\omega)$ is a regular value for π the linearization $D^J_u\colon W^{1,p}(\Sigma; u^*TM) \to \mathcal{E}_u$ is surjective and has a bounded right inverse for every $u \in \pi^{-1}(J)$.*

Proof. We are to consider the linearizations

$$(\pi_*)_{(u,J)}\colon T_{(u,J)}\tilde{\mathcal{M}}^{*,k}(\Sigma, A) \to T_J\mathcal{J}^k_\tau(M,\omega).$$

Now

$$T_{(u,J)}\tilde{\mathcal{M}}^{*,k}(\Sigma, A) = \ker \mathcal{D}_{u,J} = \{(\xi, Y)|D^J_u\xi + LY = 0\}$$

where the map $L\colon T_J\mathcal{J}^k_\tau(M,\omega) \to \mathcal{E}_{u,J}$ is defined by $LY = \frac{1}{2}Y(u) \circ du \circ j$. Of course $(\pi_*)_{(u,J)}$ acts by $(\pi_*)_{(u,J)}(\xi, Y) = Y$. Thus

$$\ker(\pi_*)_{(u,J)} = \{(\xi, 0)|D^J_u\xi = 0\}$$

is finite-dimensional by Theorem 7.8. So to show that π is a Fredholm map it remains to show that $Im(\pi_*)_{(u,J)}$ is closed and of finite codimension. Now

$$\begin{aligned} Im(\pi_*)_{(u,J)} &= \{Y \in T_J\mathcal{J}^k_\tau(M,\omega)|(\exists\xi \in W^{1,p}(\Sigma; u^*TM))(D^J_u + LY = 0)\} \\ &= L^{-1}(Im(D^J_u)). \end{aligned}$$

By Theorem 7.8 $Im(D^J_u)$ is closed and of finite codimension, so since L is a bounded operator it immediately follows that $Im(\pi_*)_{(u,J)} = L^{-1}(Im(D^J_u))$ is closed and of finite codimension at most equal to $\dim\text{coker}\, D^J_u$. This completes the proof that π is a Fredholm map. To complete the proof of the first sentence of the proposition we need to show that $\dim\text{coker}\, Im(\pi_*)_{(u,J)} \geq \dim\text{coker}\, D^J_u$. Now recalling that $LY = \mathcal{D}_{u,J}(0, Y)$, the proof of Proposition 7.10 shows that a complement to $Im(D^J_u)$ is spanned by elements of the form $LY_1, \ldots, LY_m$ for suitable $Y_i \in T_J\mathcal{J}^k_\tau(M,\omega)$. Assuming the LY_i to be linearly independent (by decreasing m if necessary), the Y_i will be linearly independent elements of $T_J\mathcal{J}^k_\tau(M,\omega)$ whose span has trivial intersection with $Im(\pi_*)_{(u,J)} = L^{-1}(Im(D^J_u))$, proving that $\dim\text{coker}\,(\pi_*)_{(u,J)} \geq \dim\text{coker}\, D^J_u$. Since, as noted earlier, the reverse inequality follows directly from $Im(\pi_*)_{(u,J)} = L^{-1}(Im(D^J_u))$, this completes the proof of the first sentence of the proposition.

In particular, it follows that if J is a regular value of π, so that for every $u \in \pi^{-1}(J)$ the linearization $(\pi_*)_{(u,J)}$ is surjective, then for every $u \in \pi^{-1}(J)$ the linearization D^J_u is surjective. As has essentially been discussed earlier, this fact, combined with the fact that its kernel is finite-dimensional by Theorem 7.8, implies that it has a bounded right inverse. Indeed, the Hahn-Banach theorem can be used to construct a closed subspace $V \leq W^{1,p}(\Sigma; u^*TM)$ such that $\ker D^J_u \oplus V = W^{1,p}(\Sigma; u^*TM)$. The restriction to V of D^J_u is a bounded linear bijection whose inverse $Q\colon \mathcal{E}_u \to V$ is bounded by the open mapping theorem, and then if we think of Q as a map to all of $W^{1,p}(\Sigma; u^*TM)$ it will be a bounded right inverse to D^J_u.

□

Corollary 7.13. *If J is a regular value of the map $\pi\colon \tilde{\mathcal{M}}^{*,k}(\Sigma, A) \to \mathcal{J}^k_\tau(M,\omega)$ then $\mathcal{M}^*(\Sigma, J, A)$ is a C^{k-1} manifold of dimension $2n(1-g) + 2\langle c_1(TM), A\rangle$.*

Proof. Indeed, $\mathcal{M}^*(\Sigma, J, A) = \pi^{-1}(J)$, so if J is a regular value of π then Proposition 7.12 shows that D^J_u has a bounded right inverse (and in particular is surjective) for each $u \in \mathcal{M}^*(\Sigma, J, A)$. So by Proposition 7.4 $\mathcal{M}^*(\Sigma, J, A)$ is a manifold, whose dimension on a neighborhood of any given element u will be the dimension of the kernel of the linearization D^u_J. But since the cokernel of this operator is trivial, this dimension will be equal to the index of D^J_u, which is equal to the value given in the statement of the corollary (independently of u) by Theorem 7.9. □

Regular values of π do indeed exist in abundance, as follows from the following infinite-dimensional version of Sard's theorem:

Theorem 7.14 (Sard-Smale Theorem)**.** *Let $\mathcal{F}\colon \mathcal{B} \to \mathcal{B}'$ be a C^l Fredholm map between two (second-countable) Banach manifolds where $l \geq 1$, and assume that for all $x \in \mathcal{B}$ $ind((d\mathcal{F})_x) \leq l - 1$. Then the set of regular values of $\mathcal{F}$ is a countable intersection of open and dense subsets of $\mathcal{B}'$.*

See [MS2, Section A.5] for a proof. Of course, the Baire category theorem then implies that the set of regular values is dense in $\mathcal{B}'$.

Since Proposition 7.12 shows that the index of each $(d\pi)_{(u,J)}$ is equal to $ind(D_u^J) = 2n(1-g) + 2\langle c_1(TM), A\rangle$, if we take $k \geq \max\{2, 2+2n(1-g)+2\langle c_1(TM), A\rangle\}$ and set $\mathcal{J}_\tau^{reg,k}(M,\omega)$ equal to the set of regular values of π, this completes the proof of Theorem 7.1 (with the exception of the orientation issue, which we won't discuss (see [MS2, pp. 50–51 and Section A.2])).

Recall that the reason for working with C^k almost complex structures rather than C^∞ ones was mainly that doing so allowed us to stay within the category of Banach manifolds, the issue being that while the space of C^∞ sections of a vector bundle admits a complete metric (so that $\mathcal{J}_\tau(M,\omega)$ is what would be called a Fréchet manifold) it does not admit a complete norm. It is however true that Theorem 7.1 can be leveraged to show that for J in a countable intersection of open dense sets within the space $\mathcal{J}_\tau(M,\omega)$ of C^∞ almost complex structures it holds that $\mathcal{M}^*(\Sigma, J, A)$ is a manifold of the expected dimension (with D_u^J surjective at each point of $\mathcal{M}^*(\Sigma, J, A)$); see [MS2, pp. 52-53] for the argument.

7.4. **Variations on Theorem 7.1.** Let us briefly outline the argument that we used in proving Theorem 7.1, which asserts that for suitable generic J the space $\mathcal{M}^*(\Sigma, J, A)$ of J-holomorphic curves $\Sigma \to M$ representing the homology class A is a manifold:

(1) Theorem 7.8 showed that, for any J and any J-holomorphic map u, the linearization D_u^J of the Cauchy-Riemann operator at u is a Fredholm operator.
(2) Proposition 7.10 showed that the "universal moduli space"
$$\tilde{\mathcal{M}}^{*,k}(\Sigma, A) = \{(u,J) \in \mathcal{B}^* \times \mathcal{J}_\tau^k(M,\omega) | \bar{\partial}_J u = 0, u_*[\Sigma] = A\}$$
is a Banach manifold.
(3) Given this, Proposition 7.12 and the Sard-Smale theorem showed that the projection $\pi\colon \tilde{\mathcal{M}}^{*,k}(\Sigma, A) \to \mathcal{J}_\tau^k(M,\omega)$ has many regular values J, for any of which the moduli space $\mathcal{M}^*(\Sigma, J, A)$ will be a manifold of the expected dimension.

This basic outline can be used to prove many similar statements, as we now briefly sketch.

7.4.1. *Compatible almost complex structures.* If instead of ω-tame almost complex structures of class C^k we were to restrict to the space $\mathcal{J}^k(M,\omega)$ of ω-compatible almost complex structures of class C^k, the argument would go through verbatim, except that because the tangent space to $\mathcal{J}^k(M,\omega)$ is smaller than that to $\mathcal{J}_\tau^k(M,\omega)$ (due to the additional condition $\omega(Jv, Jw) = \omega(v,w)$) the allowed perturbations Y that are used in the proof of Proposition 7.10 would be subject to an additional constraint. However, one can still show that for any $\eta \in Ann(Im(D_u^J))$ there is $Y \in T_J\mathcal{J}^k(M,\omega)$ such that $\eta(\mathcal{D}_{u,J}(0,Y)) > 0$; the main relevant lemma here is [MS2, Lemma 3.2.2]. Once this is established the proofs go through without change to establish a version of Theorem 7.1 with $\mathcal{J}_\tau^k(M,\omega)$ replaced by $\mathcal{J}^k(M,\omega)$.

7.4.2. *Parametrized moduli spaces.* Assume that J_0, J_1 belong to the space $\mathcal{J}_\tau^{k,reg}(M,\omega)$ produced by Theorem 7.1. Now the space $\mathcal{J}_\tau^k(M,\omega)$ is contractible and in particular connected, so there are (many) paths $\{J_t\}_{t\in[0,1]}$ connecting J_0 to J_1. We *can't* expect to arrange that $J_t \in \mathcal{J}_\tau^{k,reg}(M,\omega)$ for every t. However, it is true that, for generic paths $\{J_t\}_{t\in[0,1]}$ from J_0 to J_1, the space
$$\mathcal{M}^*(\Sigma, \{J_t\}, A) = \{(t,u) \in [0,1] \times \mathcal{B}^* | \bar{\partial}_{J_t} u = 0\}$$
is a manifold of dimension $1 + 2n(1-g) + 2\langle c_1(TM), A\rangle$.

Indeed, this will hold provided that the path $\{J_t\}_{t\in[0,1]}$ has the property that that map $(t', u') \mapsto \bar{\partial}_{J_{t'}} u'$ has surjective linearization at every (t,u) such that $\bar{\partial}_{J_t} u = 0$. That the linearization is Fredholm follows quickly from Theorem 7.8. One then shows that, where $\mathcal{P}(J_0, J_1)$ denotes the space of C^k paths from J_0 to J_1, the map with domain $[0,1] \times \mathcal{B}^* \times \mathcal{P}(J_0, J_1)$ given by $(t, u, \{J_t\}) \mapsto \bar{\partial}_{J_t} u$ has surjective linearization at any of its zeros, and hence that the universal moduli space $\{t, u, \{J_t\} | \bar{\partial}_{J_t} u = 0\}$ is a Banach manifold. Then apply the Sard-Smale theorem to the projection of this Banach manifold onto $\mathcal{P}(J_0, J_1)$ to infer that generic paths from J_0 to J_1 are regular values for the projection, and that for any such path $\{J_t\}$ the parametrized moduli space $\mathcal{M}^*(\Sigma, \{J_t\}, A)$ is a manifold of the expected dimension.

7.4.3. *Multiply covered curves.* Throughout this section we have restricted attention to simple curves, insisting that the map u belong to the open subset $\mathcal{B}^*$ of $\mathcal{B} = W^{1,p}(\Sigma, M)$ of maps which are not multiple covers. It's not difficult to see that this is required to get a statement like Theorem 7.1: after all, if $\phi\colon \tilde{\Sigma} \to \Sigma$ is a holomorphic map of degree 2, then for any J-holomorphic map $v\colon \Sigma \to M$ representing the class A the map $v \circ \phi\colon \tilde{\Sigma} \to M$ is also J-holomorphic and represents the class $2A$, so the space of J-holomorphic maps $\tilde{\Sigma} \to M$ representing $2A$ (without assuming that the map is simple) would need to have dimension at least that of $\mathcal{M}^*(\Sigma, J, A)$. But the "expected dimension" $2n(1 - g(\Sigma')) + 2\langle c_1(TM), 2A\rangle$ might well be smaller than the dimension of $\mathcal{M}^*(\Sigma, J, A)$, which would pose problems for incorporating multiple covers into Theorem 7.1. This suggests that if one does want a version of Theorem 7.1 which doesn't assume that the maps involved are simple, one needs to somehow evade the issue that a multiple cover of a J-holomorphic map is also J-holomorphic.

The only place where the assumption that the maps u that we consider are simple is in the proof of Proposition 7.10, where we identify an small open subset $U \subset \Sigma$ for which the J-holomorphic map u obeys $u^{-1}(u(U)) = U$. The point here was that a perturbation Y of J that is supported in an open set $W \subset M$ with $W \cap u(\Sigma) = u(U)$ will have the property that $\eta(\mathcal{D}_{u,J}(0, Y))$ is equal to a certain integral over $U \subset \Sigma$. If instead u were a multiple cover there would be disjoint open subsets $U_1, \dots, U_m$ with $u(U_i) = u(U)$, and to find $\eta(\mathcal{D}_{u,J}(0, Y)$ we would need to sum the integrals over the various U_i. Cancellation between these integrals might prevent Y from being constructed in such a way as to ensure that $\eta(\mathcal{D}_{u,J}(0, Y)) \neq 0$.

If the perturbations Y were allowed to vary with $z \in \Sigma$, then the issue described above would not arise, since we could just have Y be nonzero on the single small open set U. Perturbing J by such a Y would result not in an almost complex structure on M, but rather in an almost complex structure on M which depends on a point in Σ. Thus we could instead consider, for a map $J \in C^k(\Sigma, \mathcal{J}^k_\tau(M, \omega))$, solutions $u\colon \Sigma \to M$ to the equation

$$du + \frac{1}{2} J(z, u(z)) \circ du \circ j = 0. \tag{32}$$

An exact analogue of the proof of Theorem 7.1 shows that, for generic $J \in C^k(\Sigma, \mathcal{J}^k_\tau(M, \omega))$, the space $\mathcal{M}(\Sigma, J, A)$ of solutions to (32) which represent the homology class A (without any assumption that the map is simple) is a manifold of the expected dimension. Solutions to (32) are not pseudoholomorphic curves (in particular it's not true that a multiple cover of a solution to (32) solves (32)); however it's straightforward to construct an almost complex structure $\tilde{J}$ on the *product* $\Sigma \times M$ such that u solves (32) if and only if the map $z \mapsto (z, u(z))$ is $\tilde{J}$-holomorphic. In particular this allows the regularity results from Section 6 to be brought to bear on solutions of (32).

7.4.4. *Curves with point constraints.* Choose submanifolds $N_1, \dots, N_l$ of M and, with notation as in Theorem 7.1, consider the set

$$\mathcal{M}^{*,k}(\Sigma, J, A, N_1, \dots, N_l) = \{(u, \vec{z}) \in \mathcal{B}^* \times \Sigma^l | u_*[\Sigma] = A, \bar{\partial}_J u = 0, u(z_i) \in N_i \text{ for each } i = 1, \dots, l\}.$$

The arguments in the proof of Theorem 7.1 can be used to show that for generic $J \in \mathcal{J}^k_\tau(M, \omega)$ (with k large enough) this set is a manifold of dimension $2n(1 - g) + 2\langle c_1(TM), A\rangle + \sum_{i=1}^l (2 - codim(N_i))$. (The $2 - codim(N_i)$ comes from the fact that the choice of z_i gives us two degrees of freedom, while once z_i is chosen requiring that $u(z_i) \in N_i$ cuts down the dimension by the codimension of N_i.) To show this, one first shows that the map

$$ev\colon \tilde{\mathcal{M}}^{*,k}(\Sigma, A) \times \Sigma^l \to M^l$$
$$(u, J, \vec{z}) \mapsto (u(z_1), \dots u(z_l))$$

is a submersion. This implies that the "universal moduli space"

$$\tilde{\mathcal{M}}^{*,k}(\Sigma, A, N_1, \dots, N_l) = ev^{-1}(N_1 \times \cdots \times N_l) = \{(u, J, \vec{z}) | (u, J) \in \tilde{\mathcal{M}}^{*,k}(\Sigma, A), u(z_i) \in N_i \text{ for each } i = 1, \dots, l\}$$

is a Banach manifold. Just as in the proof of Theorem 7.1 the Sard-Smale theorem can be applied to the projection $\tilde{\mathcal{M}}^{*,k}(\Sigma, A, N_1, \dots, N_l) \to \mathcal{J}^k_\tau(M, \omega)$ to obtain regular values J, for any of which $\mathcal{M}^{*,k}(\Sigma, J, A, N_1, \dots, N_l)$ will indeed be a manifold of the expected dimension.

By the same token, if we *fix* a tuple $\vec{z} \in \Sigma_l$, the space

$$\mathcal{M}^{*,k}(\Sigma, J, A, N_1, \dots, N_l, \vec{z}) = \{u \in \mathcal{B}^* | u_*[\Sigma] = A, \bar{\partial}_J u = 0, u(z_i) \in N_i\}$$

will, for generic J, be a manifold of dimension $2n(1-g)+2\langle c_1(TM), A\rangle - \sum_{i=1}^{l} codim(N_i)$. The argument for this is the same as in the previous paragraph, except that now the key point is that the evaluation map $\tilde{\mathcal{M}}^{*,k}(\Sigma, A) \to M^l$ given by $(u, J) \mapsto (u(z_1), \dots, u(z_l))$ (with $\vec{z}$ fixed, unlike in the last paragraph) is a submersion. The proof that the evaluation maps of this paragraph and the last are submersions may be an interesting exercise using the methods of this section and the last, or you can consult [MS2, Proposition 3.4.2].

As should be apparent, the arguments of the various parts of this subsection can be combined with each other to yield additional extensions of Theorem 7.1; details are left to the reader.

8. Compactness

The final ingredient needed before we can seriously begin applying the theory of pseudoholomorphic curves is a statement that spaces of such curves (now known to be manifolds in favorable cases) are compact, at least after we include some additional limit points. The key point will be (roughly speaking) that if a sequence of curves $u_n\colon \Sigma \to M$ representing the same homology class (or more generally satisfying an energy bound) has $\|du_n(p)\| \to \infty$ at some point p, then arbitrarily small neighborhoods U_n of p will, for n large, have $E(u_n|_{U_n}) > \hbar$ for some universal positive constant $\hbar$. Consequently (as the total energy is bounded) there can only be finitely many such small neighborhoods, and we will then focus in on these neighborhoods to understand the possible limiting behavior.

Throughout this section we will consider a connected symplectic manifold (M, ω) with an almost complex structure $J \in \mathcal{J}_\tau(M, \omega)$, which induces a Riemannian metric $g_J(v, w) = \frac{1}{2}(\omega(v, Jw) + \omega(w, Jv))$.[14] We won't insist that M be closed, but will assume that the behavior "at infinity" of the Riemannian manifold (M, g_J) is somewhat controlled. Specifically, we require (where for $v \in TM$ we write $|v| = g_J(v, v)^{1/2}$):

(1) (M, g_J) is complete (as a metric space with distance given by minimal lengths of paths, which is equivalent to various other kinds of completeness by the Hopf-Rinow theorem).
(2) There are constants $r_0, C_0 > 0$ such that for every $x \in M$ the exponential map $\exp_x\colon \{v \in T_xM \| v| \le r_0\} \to \overline{B(x, r_0)}$ is a diffeomorphism, with $\|(d\exp_x)_v\| \le C_0$ and $\|(d\exp_x^{-1})_y\| \le C_0$ for each $v \in \{v \in T_xM \| v| \le r_0\}$ and $y \in \overline{B(x, r_0)}$.
(3) There is a constant C_1 such that for all $x \in M$ and $v, w \in T_xM$ we have $|\omega(v, w)| \le C_1|v||w|$.

The second and third conditions above always hold on any Riemannian manifold with some constants r_0, C_0, C_1 *depending on* x, and we are thus requiring that it be possible to choose these constants independently of x. Easy covering arguments show that this can be done if M is closed (in which case the first condition of course also holds). We will be proving a number of estimates about J-holomorphic curves which will involve some constants; these constants will generally depend only on r_0 and C_0 above. If we are considering a sequence of almost complex structures $\{J_n\}_{n=1}^\infty$ which is Cauchy in (say) the C^2 norm, then the constants r_0 and C_0 associated to the J_n can be taken independent of n, in light of which the constants in our lemmas below can too.

If J is ω-compatible rather than just ω-tame then we have $|\omega(v, w)| = |g_J(Jv, w)| \le |Jv||w| = |v||w|$ and so we can just take $C_1 = 1$.

---

[14] J and hence g_J should probably be at least C^2-smooth in order to allow customary Riemannian geometry constructions such as the exponential map to be made and to enjoy the usual properties; there are however ways of getting around this requirement on J if one really needs to—for instance in [AL, Chapter V] Sikorav only requires J to be Hölder continuous—essentially by using in place of g_J a smooth Riemannian metric g on M which is C^0-close to g_J.

Recall that Proposition 2.1 and Remark 2.5 show that, if $u\colon \Sigma \to M$ is J-holomorphic, then

$$E(u) = Area(u) = \int_\Sigma u^*\omega \quad \text{(if } u \text{ is } J\text{-holomorphic)}$$

If J is ω-compatible, then Proposition 2.4 shows that *any* C^1 map $u\colon \Sigma \to M$ obeys $Area(u) \geq \int_\Sigma u^*\omega$. In our case, where J may just be tame but obeys (1)-(3) above, integrating (3) (applied to orthogonal vectors u_*e_1, u_*e_2) over Σ shows that

$$\int_\Sigma u^*\omega \leq C_1 Area(u) \quad \text{(for any } u\text{)}. \tag{33}$$

8.1. **The isoperimetric inequality and the monotonicity formula.** A number of important relationships exist between the area of a J-holomorphic curve; the length of its boundary (if there is one); and the diameter of its image. First, though, we prove a simple fact (a very weak analogue of the isoperimetric inequality) whose proof has nothing to do with J-holomorphicity. Throughout we continue assume that (M, ω, J) obey (1)-(3) above, so in particular we have the constants r_0, C_0.

Proposition 8.1. *There is a constant C such that if $\gamma\colon S^1 \to M$ is a C^1 curve having image contained in a ball $B(x_0, r_0/3)$ of radius $r_0/3$, then there is a C^1 map $v\colon D^2 \to B(x_0, r_0)$ such that $v|_{\partial D^2} = \gamma$ and*

$$Area(u) \leq CL(\gamma)^2$$

where $L(\gamma) = \int_0^{2\pi} \left|\frac{d\gamma}{d\theta}\right| d\theta$ is the length of γ.

Proof. We have $Im(\gamma) \subset B(x_0, r_0/3) \subset B(\gamma(1), 2r_0/3)$, so there is a map $\xi\colon S^1 \to \{v \in T_{\gamma(1)}M \| v| < 2r_0/3\}$ defined by the property that

$$\gamma(e^{i\theta}) = \exp_{\gamma(1)} \xi(\theta).$$

Define $v\colon D^2 \to M$ by

$$v(se^{i\theta}) = \exp_{\gamma(1)}(s\xi(e^{i\theta})).$$

Evidently for all $v \in D^2$ we have $v(z) \in B(\gamma(1), 2r_0/3) \subset B(x_0, r_0)$.

Now

$$\left|\frac{\partial v}{\partial s}\right| = \left|(d\exp_{\gamma(1)})_{s\xi(e^{i\theta})}\xi(e^{i\theta})\right| = |\xi(e^{i\theta})|$$

(the last equality follows from Gauss' lemma) while

$$\left|\frac{\partial v}{\partial \theta}\right| = \left|(d\exp_{\gamma(1)})_{s\xi(e^{i\theta})}\left(s\frac{d\xi}{d\theta}\right)\right| \leq C_0\left|\frac{d\xi}{d\theta}\right| = C_0\left|(d\exp^{-1}_{\gamma(1)})_{\gamma(e^{i\theta})}\frac{d\gamma}{d\theta}\right| \leq C_0^2\left|\frac{d\gamma}{d\theta}\right|.$$

Hence

$$Area(v) \leq \int_0^{2\pi}\int_0^1 \left|\frac{\partial v}{\partial s}\right|\left|\frac{\partial v}{\partial \theta}\right| ds d\theta \leq C_0^2 \int_0^{2\pi} |\xi(e^{i\theta})|\left|\frac{d\gamma}{d\theta}\right| d\theta \leq C_0^2 L(\gamma)\int_0^{2\pi}\left|\frac{d\gamma}{d\theta}\right| d\theta = C_0^2 L(\gamma)^2,$$

where we've used the observation that $|\xi(e^{i\theta})| = dist(\gamma(1), \gamma(e^{i\theta}))$ is no larger than the length of γ. □

Remark 8.2. In the special case that $(M, \omega, J) = (\mathbb{R}^{2n}, \omega_0, J_0)$ (where the constant r_0 can be taken to be ∞) the classical isoperimetric inequality asserts that the constant C can be taken equal to $\frac{1}{4\pi}$; note that this constant is achieved by any circle that is contained in a 2-plane in $\mathbb{R}^{2n}$. (Usually the isoperimetric inequality is expressed for curves in $\mathbb{R}^2$ rather than $\mathbb{R}^{2n}$, but the standard proof using Fourier series extends to higher dimensions without difficulty.)

Corollary 8.3. *If S is a possibly disconnected closed 1-manifold (i.e. a union of finitely many circles) and $\gamma\colon S \to M$ is a C^1 map with image contained in a ball $B(x_0, r_0/3)$, then there is a C^1 map $v\colon \Sigma' \to B(x_0, r_0)$ from a (possibly disconnected) 2-manifold with boundary such that $v|_{\partial\Sigma'} = \gamma$ and $Area(v) \leq CL(\gamma)^2$, where C is the constant of Proposition 8.1.*

Proof. γ is the union of finitely many curves $\gamma_1, \dots, \gamma_k\colon S^1 \to B(x_0, r_0/3)$, so Proposition 8.1 gives maps $v_1, \dots, v_k\colon D^2 \to B(x_0, r_0)$ with $v_i|_{\partial D^2} = \gamma_i$ and $Area(v_i) \le CL(\gamma_i)^2$. Take for Σ' a disjoint union of k copies of D^2 and for v the map whose restriction to the ith copy of D^2 is v_i. This satisfies our requirements, as

$$Area(v) = \sum_{i=1}^{k} Area(v_i) \le C \sum_{i=1}^{k} L(\gamma_i)^2 \le C\left(\sum_{i=1}^{k} L(\gamma_i)\right)^2 = CL(\gamma)^2.$$

□

Corollary 8.4. *Let $u\colon \Sigma \to M$ be a C^1 J-holomorphic map from a compact surface Σ with boundary whose image is contained in a ball $B(x_0, r_0/3)$. Then $Area(u) \le CC_1 L(u|_{\partial\Sigma})^2$, where C is the constant of Proposition 8.1.*

Proof. $u|_{\partial\Sigma}\colon \partial\Sigma \to B(x_0, r_0/3)$ is a C^1 map from a closed 1-manifold, so find $v\colon \Sigma' \to B(x_0, r_0)$ so that $v|_{\partial\Sigma'} = u|_{\partial\Sigma}$ and $Area(v) \le CL(u|_{\partial\Sigma})^2$. Now since u and v both have image in the $B(x_0, r_0)$, which is a diffeomorphic copy of a ball and therefore to which ω restricts as an exact 2-form, the fact that u and v agree on their boundaries implies via Stokes' theorem that $\int_\Sigma u^*\omega = \int_{\Sigma'} v^*\omega$. Hence

$$Area(u) = \int_\Sigma u^*\omega = \int_{\Sigma'} v^*\omega \le C_1 Area(v) \le CC_1 L(u|_{\partial\Sigma})^2,$$

where we've used (33). □

Now let $u\colon \Sigma \to M$ be a J-holomorphic map where Σ is compact, and let $p_0 \in \Sigma$ be a point with the property that

$$r_1 := dist(u(p_0), u(\partial\Sigma)) > 0.$$

Define

$$\mu\colon \Sigma \to \mathbb{R}$$
$$p \mapsto dist(u(p), u(p_0))$$

Then if

$$r < \min\{r_0/3, r_1\}$$

the maps

$$\mu|_{\mu^{-1}(0,r]} \text{ and } \mu^2|_{\mu^{-1}[0,r]}$$

are smooth maps whose domains are contained in the interior of Σ. Of course, if $s \in (0, r]$ is a critical value of μ, then s^2 is a critical value of μ^2. Since $\mu^{-1}[0, r]$ is compact, the critical points of $\mu^2|_{\mu^{-1}[0,r]}$ form a compact set, and so their image (i.e. the set of critical values of $\mu^2|_{\mu^{-1}[0,r]}$) is a compact subset of $[0, r^2]$; moreover this subset has measure zero by Sard's theorem. Consequently the regular values of $\mu|_{\mu^{-1}(0,r]}$ form an open, full-measure subset of $[0, r]$.

Proposition 8.5. *With notation as above, suppose that $0 < a < b \le r$ and that each point of $[a, b]$ is a regular value for μ. Then there is a diffeomorphism*

$$\Phi\colon \mu^{-1}(\{a\}) \times [a, b] \to \mu^{-1}[a, b]$$

such that, for $s \in [a, b]$, $\Phi\left(\mu^{-1}(\{a\}) \times \{s\}\right) = \mu^{-1}(\{s\})$.

Proof. (Sketch) Choose an auxiliary Riemannian metric on Σ, which (at least on the region where μ is differentiable, which includes $\mu^{-1}[a, b]$) produces a gradient vector field $\nabla\mu$. Since μ has no critical points in $\mu^{-1}[a, b]$ and since $\mu^{-1}[a, b]$ is compact, there is $\delta > 0$ such that $\|\nabla\mu\|_h \ge \delta$ everywhere on $\mu^{-1}[a, b]$. Suppose that $\gamma\colon [t_1, t_2] \to \mu^{-1}[a, b]$ is an integral curve of the vector field $\nabla\mu$. Then

$$\frac{d}{dt}(\mu(\gamma(t))) = d\mu(\dot\gamma(t)) = d\mu(\nabla\mu) \ge \delta^2 > 0.$$

Consequently an integral curve for $\nabla\mu$ which begins at $p \in \mu^{-1}(\{a\})$ will, for $a \leq s \leq b$, pass through $\mu^{-1}(\{s\})$ in time at most $\frac{s-a}{\delta^2}$, and moreover will pass through $\mu^{-1}(\{s\})$ only once since μ strictly increases along the curve as long as it remains in $\mu^{-1}[a,b]$. With this said, the diffeomorphism $\Phi\colon \mu^{-1}(\{a\}) \times [a,b] \to \mu^{-1}[a,b]$ can be defined by setting $\Phi(p,s)$ equal to the unique point of $\mu^{-1}(\{s\})$ that lies on the integral curve of $\nabla\mu$ that passes through p. Verification that this is indeed a diffeomorphism is left to the interested reader. □

Proposition 8.6. *Suppose that each point of $[a,b] \subset (0,r)$ is a regular value for μ, and for any $t \in [a,b]$ define*

$$A(t) = Area(u(\Sigma) \cap B(p_0,t)) \quad L(t) = Length(u(\Sigma) \cap \partial B(p_0,t)).$$

Then

$$A(b) - A(a) \geq (b-a) \min_{a \leq t \leq b} L(t).$$

Proof. Note that

$$A(b) - A(a) = Area(v) \text{ where } v = u \circ \Phi$$

and $\Phi\colon \mu^{-1}(\{a\}) \times [a,b] \to \mu^{-1}[a,b]$ is the diffeomorphism of the previous paragraph. (Since $\mu^{-1}(\{a\})$ is a closed one-dimensional submanifold of Σ it is a union of circles). Let θ be a local (angular) coordinate on $\mu^{-1}(\{a\})$ and let s be the coordinate on $[a,b]$. Thus

$$(\mu \circ \Phi)_* \partial_\theta = 0 \quad (\mu \circ \Phi)_* \partial_s = 1.$$

Write $\delta_{p_0}\colon B(u(p_0), r) \to \mathbb{R}$ for the function $x \mapsto dist(x, u(p_0))$, so that $\mu = \delta_{p_0} \circ u$. So for any tangent vector w to the domain $\mu^{-1}(\{a\}) \times [a,b]$ of v we have

$$g_J(\nabla\delta_{p_0}, v_* w) = (\delta_{p_0} \circ u \circ \Phi)_* w = \mu_* \Phi_* w.$$

In particular

$$g_J(\nabla\delta_{p_0}, v_*\partial_\theta) = 0 \quad g_J(\nabla\delta_{p_0}, v_*\partial_s) = 1.$$

Now $\nabla\delta_{p_0}$ has norm 1, so this implies that the orthogonal projection of $v_*\partial_s$ along $\nabla\delta_{p_0}$ has length 1; since $\nabla\delta_{p_0}$ is orthogonal to $v_*\partial_\theta$ this proves that $|v_*\partial_\theta \wedge v_*\partial_s|_{g_J} \geq |v_*\partial_\theta|_{g_J}$. Thus

$$\begin{aligned} A(b) - A(a) = Area(v) &= \int_a^b \int_{\mu^{-1}(\{a\})} |v_*\partial_\theta \wedge v_*\partial s|_{g_J} d\theta ds \geq \int_a^b \left(\int_{\mu^{-1}(\{a\})} |v_*\partial_\theta|_{g_J} d\theta \right) ds \\ &\geq (b-a) \min_{a \leq s \leq b} L(t). \end{aligned}$$

□

On any interval $[a,b]$ as in Proposition 8.6, by using the diffeomorphism Φ as in the proof of the proposition it's not difficult to see that A is differentiable and that L is continuous on $[a,b]$.[15] Now if s is any regular value of μ in $(0,r)$, since the regular values form an open set we can find an interval $[a,b]$ as in Proposition 8.6 such that $s \in [a,b]$. Then taking limits as $a \to s^-$ and $b \to s^+$, we see from Proposition 8.6 that

$$A'(s) \geq L(s) \text{ if } s \in (0,r) \text{ is a regular value of } \mu\ .$$

Since we assume $r < r_0/3$, Corollary 8.4 applies to show that, for $0 < s < r$, $L(s) \geq \frac{1}{\sqrt{CC_1}} \sqrt{A(s)}$. Therefore

$$\frac{d}{ds}\sqrt{A(s)} = \frac{A'(s)}{2\sqrt{A(s)}} \geq \frac{1}{2\sqrt{CC_1}} \text{ if } s \in (0,r) \text{ is a regular value of } \mu\ .$$

So if $[a,b] \subset (0,r)$ is an interval of regular values of μ then $\sqrt{A(b)} - \sqrt{A(a)} \geq \frac{b-a}{2\sqrt{CC_1}}$.

Now since the regular values of μ form an open, full measure subset of $[0,r]$, for any $\epsilon > 0$ we can find a finite disjoint collection of intervals in $(0,r)$ of regular values for μ having total length at least $(1-\epsilon)r$. We also know that $\sqrt{A(0)} = 0$, and that $t \mapsto \sqrt{A(t)}$ is a monotone increasing function. So on each of the intervals of regular

[15] For this statement and various others below one might worry about what happens if L vanishes for some s; however unless u is constant on each connected component this doesn't happen if s is a regular value, for instance because the Carleman similarity principle shows that a circle in Σ can't be mapped entirely to the same point.

values in our collection $\sqrt{A}$ increases by at least $\frac{1}{2\sqrt{CC_1}}$ times the length of the interval, while it also increases (by an unknown amount) on the remainder of $[0, r]$. Hence $\sqrt{A(r)} \geq \frac{(1-\epsilon)r}{2\sqrt{CC_1}}$. Since ϵ was arbitrary, this proves:

Theorem 8.7 (Monotonicity theorem)**.** *Let $u\colon \Sigma \to M$ be a nonconstant J-holomorphic map from a connected compact surface Σ and suppose $p_0 \in \Sigma$, $0 < r < r_0/3$, and*

$$u(\partial\Sigma) \cap B(u(p_0), r) = \emptyset.$$

Then

$$Area(u(\Sigma) \cap B(u(p_0), r)) \geq \frac{1}{4CC_1} r^2.$$

Remark 8.8. In the case that $(M, \omega, J) = (\mathbb{R}^{2n}, \omega_0, J_0)$, we have $r_0 = \infty$, $C_1 = 1$ since J_0 is ω_0-compatible, and as noted earlier $C = \frac{1}{4\pi}$ by the classical isoperimetric inequality. Thus $\frac{1}{4CC_1} = \pi$. So we recover the fact that a J_0-holomorphic curve passing through the origin of a ball of radius r in $\mathbb{R}^{2n}$ and with boundary contained in the boundary of the ball must have area at least πr^2. We appealed to this fact (which more generally holds for minimal surfaces) earlier in the sketch of the proof of the Gromov non-squeezing theorem.

Some elementary arguments now let us derive from Theorem 8.7 and Corollary 8.3 some facts whose hypotheses (unlike those of the results so far) don't make reference to balls in M:

Proposition 8.9. *There is a constant $\hbar > 0$ such that, if Σ is a compact surface and $u\colon \Sigma \to M$ is C^1 and J-holomorphic with $Area(u) < \hbar$, then $Area(u) \leq CC_1 L(u|_{\partial\Sigma})^2$. In particular if $\partial\Sigma = \emptyset$, then any nonconstant J-holomorphic map $u\colon \Sigma \to M$ has area at least $\hbar$.*

Proof. Where $C' = \frac{1}{4CC_1}$ is the constant appearing in Theorem 8.7, set $\hbar = \min\{CC_1, C'\}\left(\frac{r_0}{6}\right)^2$. Assume that $Area(u) < \hbar$ (and also that u is nonconstant, since the result is trivial if u is constant). Write $L = L(u|_{\partial\Sigma})$. Obviously if $L \geq \frac{r_0}{6}$ then $Area(u) \leq CC_1 L^2$, so it suffices to consider the case that $L < \frac{r_0}{6}$.

I claim that there must be a ball of radius $\frac{r_0}{3}$ which contains the image of u. More to the point, I claim that if this were not the case, there would be a point $p \in \Sigma$ such that $B(u(p), r_0/6) \cap u(\partial\Sigma) = \emptyset$; once this is established Theorem 8.7 shows that we'd have $Area(u) \geq C'(r_0/6)^2$ which contradicts our assumption on u. To prove this claim, note first that it trivially holds if $\partial\Sigma = \emptyset$. On the other hand if $\partial\Sigma \neq \emptyset$, choose any $p_0 \in \partial\Sigma$, so by assumption $u(\Sigma) \not\subset B(u(p_0), r_0/3)$. So if p_1 has $dist(u(p_1), u(p_0)) \geq r_0/3$, then since $L < r_0/6$ we have $dist(u(p_1), u(\partial\Sigma)) \geq r_0/6$. This confirms that if no ball of radius $r_0/3$ contains the image of u then there would be a point p_1 in Σ with $B(u(p), r_0/6) \cap u(\partial\Sigma) = \emptyset$, which as discussed above contradicts Theorem 8.7 and the assumption on $Area(u)$.

But now that we know there is a ball of radius $r_0/3$ containing the image of u, Proposition 8.4 immediately gives $Area(u) \leq CC_1 L^2$, completing the proof. □

Proposition 8.10. *Let Σ be a compact connected surface with exactly two boundary components $\partial_-\Sigma, \partial_+\Sigma$. If $\epsilon < r_0/3$ and if $u\colon \Sigma \to M$ is a C^1 J-holomorphic map with $Area(u) < C'\epsilon^2$ and $L(u|_{\partial_-\Sigma}), L(u|_{\partial_+\Sigma}) < \epsilon$, then the diameter of $u(\Sigma)$ is less than 5ϵ.*

The main case of interest here is where Σ is an annulus. Here and below the diameter of a subset $S \subset M$ refers to the supremal distance (as measured by the metric on M, which takes into account paths that leave S) between any two points of S.

Proof. It's enough to show that if we had $diam(u(\Sigma)) \geq 5\epsilon$ while $L(u|_{\partial_\pm\Sigma}) < \epsilon$ then there would be $p \in \Sigma$ with $dist(u(p), u(\partial\Sigma)) > \epsilon$, since in this case Theorem 8.7 would contradict our assumption on $Area(u)$. The proof of this claim splits naturally into two cases.

Case 1: $\delta := dist(u(\partial_-\Sigma), u(\partial_+\Sigma)) \leq 2\epsilon$. Now the diameters of the individual loops $u(\partial_-\Sigma), u(\partial_+\Sigma)$ are each less than $\epsilon/2$ (since if p and q are two points on one $\partial_-\Sigma$ or $\partial_+\Sigma$ such that $u(p)$ and $u(q)$ are maximally far apart, then $u(\partial_\pm\Sigma)$ contains two distinct paths from p to q (one "clockwise" and the other "counterclockwise")). So the assumption $\delta \leq 2\epsilon$ implies that any two points $u(p), u(q) \in u(\partial\Sigma)$ have $dist(u(p), u(q)) < 3\epsilon$. So if $diam(u(\Sigma)) \geq 5\epsilon$, then choosing $x, y \in \Sigma$ so that $dist(u(x), u(y)) \geq 5\epsilon$, we have

$$\begin{aligned} 5\epsilon &\leq dist(u(x), u(\partial\Sigma)) + dist(u(y), u(\partial\Sigma)) + \max\{d(u(p), u(q)) | p, q \in \partial\Sigma\} \\ &< dist(u(x), u(\partial\Sigma)) + dist(u(y), u(\partial\Sigma)) + 3\epsilon, \end{aligned}$$

and so one or the other of $dist(u(x), u(\partial\Sigma))$ and $dist(u(y), u(\partial\Sigma))$ must be at least ϵ.

Case 2: $\delta = dist(u(\partial_-\Sigma), u(\partial_+\Sigma)) > 2\epsilon$. Choose $p \in \partial_-\Sigma$ and $q \in \partial_+\Sigma$ such that $dist(u(p), u(q)) = \delta$. Let $\gamma\colon [0, 1] \to \Sigma$ be a continuous path from p to q. The continuous function $t \mapsto dist(u(\gamma(t)), \partial_-\Sigma)$ then takes values 0 at $t = 0$ and δ at $t = 1$, so by the Intermediate Value Theorem there is t_0 such that $dist(u(\gamma(t_0), \partial_-\Sigma)) = \delta/2$. But by the definition of δ as the infimal distance from $u(\partial_-\Sigma)$ to $u(\partial_+\Sigma)$ (and by the triangle inequality) we then necessarily have $dist(u(\gamma(t_0)), \partial_+\Sigma) \geq \delta - \delta/2 = \delta/2$. So since $\delta/2 > \epsilon$, $\gamma(t_0) \in \Sigma$ has the desired property.

□

8.2. **Removal of singularities and Gromov's Schwarz Lemma.** An important consequence of the monotonicity theorem (and its corollary, Proposition 8.10) is another analogue of a property of standard holomorphic functions, namely that a holomorphic map of a punctured disk extends to a holomorphic map of the whole disc provided that it doesn't diverge too severely as one approaches the puncture. To set this up, for $0 < r < R$, introduce the notations:

$$\begin{aligned} D(R) &= \{z \in \mathbb{C} | |z| \leq R\}, \\ D^*(R) &= \{z \in \mathbb{C} | 0 < |z| \leq R\}, \\ A(r, R) &= \{z \in \mathbb{C} | r \leq |z| \leq R\}. \end{aligned}$$

Theorem 8.11 (Removal of Singularities)**.** *Suppose that $u\colon D^*(R) \to M$ is a J holomorphic map such that*

$$Area(u) = \int_{D^*(R)} \left|\frac{\partial u}{\partial x} \wedge \frac{\partial u}{\partial y}\right|_{g_J} dxdy < \infty.$$

Then u extends to a continuous function on all of $D(R)$.

Note that once we know that u extends to a continuous function, since the hypothesis that the area is finite (so also the energy $\int |du|^2$ is finite, as these are equal for J-holomorphic curves) implies that u is of class $W^{1,2}$, Theorem 6.1 implies that u is C^∞ on the interior of $D(R)$ (or, if J is just C^k, u is C^k), and so in particular $u\colon D(R) \to M$ is a genuine J-holomorphic curve.

Proof. Consider arbitrary numbers η, ρ with $0 < \eta < \rho < R$, and for $t \in [\eta, \rho]$ define

$$\alpha(t) = Area(u|_{A(\eta,t)}) \quad \lambda(t) = Length(u|_{\partial D(t)}).$$

Since the standard polar coordinate basis vectors $\partial_r, \partial_\theta$ have $|\partial_r| = 1$ and $|\partial_\theta| = r$, and so $j\partial_\theta = -r\partial_r$, we see that, since u is J-holomorphic,

$$\begin{aligned} \alpha(t) &= \int_\eta^t \int_0^{2\pi} \left|\frac{\partial u}{\partial \theta} \wedge \frac{\partial u}{\partial r}\right|_{g_J} d\theta dr = \int_\eta^t \int_0^{2\pi} \left|\frac{\partial u}{\partial \theta} \wedge J\frac{\partial u}{\partial \theta}\right|_{g_J} r^{-1} d\theta dr \\ &= \int_\eta^t \left(\int_0^{2\pi} \left|\frac{\partial u}{\partial \theta}\right|^2 r^{-1} d\theta\right) dr. \end{aligned}$$

So

$$\alpha'(t) = t^{-1} \int_0^{2\pi} \left|\frac{\partial u}{\partial \theta}\right|^2 d\theta.$$

Meanwhile

$$\lambda(t)^2 = \left(\int_0^{2\pi} \left|\frac{\partial u}{\partial \theta}\right| d\theta\right)^2 \leq 2\pi \int_0^{2\pi} \left|\frac{\partial u}{\partial \theta}\right|^2 d\theta$$

by the Schwarz inequality. This shows that

$$\alpha'(t) \geq \frac{1}{2\pi t}\lambda(t)^2.$$

Consequently we have

$$\alpha(\rho) \geq \alpha(\rho) - \alpha(\eta) \geq \int_\eta^\rho \frac{1}{2\pi t}\lambda(t)^2 dt$$

$$\geq \frac{1}{2\pi}\log\left(\frac{\rho}{\eta}\right)\min_{\eta\leq t\leq\rho}\lambda(t)^2. \tag{34}$$

Now since $Area(u) < \infty$, we have $\alpha(t) \to 0$ as $t \to 0$. Let C' be as in Proposition 8.10, and let $\epsilon > 0$ be given. Choose a number $\rho_1 > 0$ such that $\alpha(\rho_1) < C'\epsilon^2$. Since $\log(\rho_1/\eta) \to \infty$ as $\eta \to 0^+$, we can choose a value $\eta_1 \in (0, \rho_1)$ such that (34) (with $r = \rho_1, \eta = \eta_1$) forces

$$\min_{\eta_1\leq t\leq\rho_1}\lambda(t) < \epsilon.$$

Choose a particular $t_1 \in [\eta_1, \rho_1]$ so that $\lambda(t_1) < \epsilon$. I claim that $diam(u|_{D^*(t_1)}) \leq 5\epsilon$. To show this, it suffices to show that given any $\rho_2 \in (0, 1)$ we have $diam(u|_{A(\rho_2,t_1)}) < \epsilon$. But applying (34) once again with $r = \rho_2$ and $\eta = \eta_2 \ll \rho_2$ we find that (if we choose η_2 appropriately) $\min_{\eta_2\leq t\leq\rho_2}\lambda(t) < \epsilon$.

Choosing a particular $t_2 \in [\eta_2, \rho_2]$ so that $\lambda(t_2) < \epsilon$, we have $A(\rho_2, t_1) \subset A(t_2, t_1) \subset D^*(\rho_1)$. The second inclusion shows that $Area(u|_{A(t_2,t_1)}) < C'\epsilon^2$ by the choice of ρ_1, while we've arranged that the restriction of u to either boundary component of $A(t_2, t_1)$ has length less than ϵ. So by Proposition 8.10 we have $diam(u(A(t_2, t_1))) < 5\epsilon$, whence $diam(u(A(\rho_2, t_1))) < 5\epsilon$. The number $\rho_2 \in (0, t_1)$ was arbitrary, so this proves that $diam(u(D^*(t_1))) < 5\epsilon$.

In sum, we've shown that if $\epsilon > 0$ then there is $t_1 > 0$ such that $u(D^*(t_1))$ has diameter less than 5ϵ. But it immediately follows from this that if $z_n \in D^*(R)$ with $z_n \to 0$, then $u(z_n)$ is a Cauchy sequence. We have assumed that M is complete, so this Cauchy sequence has a limit u_0. Extending u by setting $u(0) = u_0$, we have arranged that if $\epsilon > 0$ then there is t_1 such that all points of $u(D(t_1))$ are within distance 5ϵ of $u(0)$, so u is continuous at 0. □

The analysis in Theorem 8.11 also helps to prove the following:

Proposition 8.12. *Where $\hbar$ is the constant of Propostion 8.9, if $\epsilon > 0$ there is a constant $\eta > 0$ such that if $u\colon D(1) \to M$ is a J-holomorphic map with $Area(u) < \hbar$, then*

$$u(D(\eta)) \subset B(u(0), \epsilon).$$

Proof. Given u as in the statement, for $0 \leq t \leq 1$ let $\alpha(t) = Area(u|_{D(t)})$ and $\lambda(t) = Length(u|_{\partial D(t)})$. By Proposition 8.9, since $Area(u) < \hbar$ we have $\alpha(t) \leq C\lambda(t)^2$ for some (newly redefined) constant C. Just as in the proof of Proposition 8.11, we have (by (34) with $\rho = 1$) for $0 < \eta < 1$

$$\hbar \geq \alpha(1) \geq \alpha(1) - \alpha(\eta) \geq \frac{1}{2\pi}\log\left(\frac{1}{\eta}\right)\min_{\eta\leq t\leq 1}\lambda(t)^2.$$

Given $\epsilon > 0$, we can then choose η independently of u such that for some $t_1 \in [\eta, 1]$ we'll have $\lambda(t_1) < \epsilon$. The same argument as in the proof of Theorem 8.11 then shows, using another application of (34), that for any $\rho_2 \in (0, t_1)$ there is is $t_2 < \rho_2$ with $\lambda(t_2) < \epsilon$. Also, $Area(u|_{A(t_2,t_1)}) \leq \alpha(t_1) < C\epsilon^2$, and so we get from Proposition 8.10 that (for some universal constant $\tilde{C}$ depending on C/C') $diam(u(A(\rho_2, t_1))) \leq diam(u(A(t_2, t_1))) < \tilde{C}\epsilon$. Since $\rho_2 < t_1$ was arbitrary this shows that $diam(u(D(t_1))) \leq \tilde{C}\epsilon$, and hence that (since $t_1 \geq \eta$) $diam(u(D(\eta))) \leq \tilde{C}\epsilon$. While t_1 and t_2 in principle depended on u, the values η and $\tilde{C}$ just depended on ϵ and $\hbar$, so the proposition follows upon renaming the parameters appropriately. □

Before applying Proposition 8.12, we record the following consequence of the proof of Theorem 6.1.

Proposition 8.13. *Fix $p > 2$, $k \geq 1$, and $r, N > 0$, and a coordinate neighborhood in M with compact closure, identified with a bounded open subset U of $\mathbb{R}^{2n}$. Then there are $\delta > 0$ and $N' > 0$ such that if $u\colon D(r) \to U$ is J-holomorphic and $\|u\|_{W^{1,p}(D(1))} \leq N$, then $\|u\|_{W^{k+1,p}(D(\delta r))} \leq N'$.*

The point of the proposition is that δ and N' depend only on the $W^{1,p}$ bound on u and not on the particular u obeying the bound.

Proof. After postcomposing the coordinate patch with a translation we assume that $u(0) = \vec{0}$ and that the almost complex structure J coincides with J_0 at $\vec{0}$. (The expression of the almost complex structure J in these new coordinates will depend on the initial value of $u(0)$, but by compactness of the neighborhood all expressions for J obtained in this way will obey a uniform C^k bound, and this will suffice for what follows.) As in Section 6.1, the equation for u to be J-holomorphic then takes the form

$$\partial_{\bar{z}} u + q(u(z))\partial_z u = 0$$

for some matrix valued function q with $q(0) = 0$ which obeys a uniform C^{k+1} bound. By induction on k, it's enough to prove the conclusion assuming that for some $N, r > 0$ we have a bound $\|u\|_{W^{k,p}(D(r))} \leq N$. The proof is essentially just a matter of examining carefully the proof of Theorem 6.1. Let $\alpha\colon D(r) \to [0, 1]$ be a smooth function equal to 1 on $D(3r/4)$ and to 0 outside $D(7r/8)$. The parameter δ shall be chosen so that the function $u_{2\delta}(z) = u(2\delta z)$ has the property that the operator on $W^{k,p}$ defined by multiplication by $q \circ \alpha u_{2\delta}$ has operator norm less than $\frac{1}{2\|T\|_{k,p}}$ where $\|T\|_{k,p}$ is the operator norm of the Calderón–Zygmund operator T on $W^{k,p}$. δ can be chosen in a way that depends only on N by the chain rule, the product rule, and the fact that $\|u_{2\delta}\|_{W^{k,p}(D(r))}$ is bounded by a constant times $\delta^{1-2/p}\|u\|_{W^{k,p}(D(r))}$.

With this δ chosen, let $\chi\colon D(r) \to [0, 1]$ be a smooth function (independent of u) supported in $D(3r/4)$ and equal to 1 on $D(r/2)$. We then have an equation

$$(I + (q \circ (\alpha u_{2\delta})) \cdot T)\partial_{\bar{z}}(\chi u_{2\delta}) = (\partial_{\bar{z}}\chi + q \circ (\alpha u_{2\delta})\partial_z \chi)u_{2\delta}.$$

The right hand side satisfies a $W^{k,p}$ bound which depends on N (and on δ, but δ just depends on N) but not on u, while the operator $I + (q \circ (\alpha u_{2\delta})) \cdot T$ is invertible as an operator on $W^{k,p}$, with inverse having norm no larger than 2. This gives us a $W^{k,p}$ bound on $\partial_{\bar{z}}(\chi u_{2\delta})$ that depends only on N, and hence a bound on $\|\chi u_{\delta/2}\|_{k+1,p}$ by Theorem 5.10. So since $\chi u_{\delta/2}|_{D(r/2)} = u|_{D(r/2)}$ we get a bound on $\|u_{2\delta}\|_{W^{k+1,p}(D(r/2))}$. Considering the behavior of the $W^{k+1,p}$-norm under rescaling by 2δ, this implies the desired bound on $\|u\|_{W^{k+1,p}(D(\delta r))}$. □

Corollary 8.14 (Gromov's Schwarz Lemma)**.** *For any compact $K \subset M$ there is a constant C such that if $u\colon D(1) \to M$ is a J-holomorphic map with Area$(u) < \hbar$ and $u(0) \in K$ then $\|du(0)\| \leq C$.*

In fact, we have the following stronger statement:

Corollary 8.15. *For any coordinate neighborhood in M with compact closure and any $p > 2, k \geq 1$ there are $C > 0$, $\delta > 0$ such that if $u\colon D(1) \to M$ is a J-holomorphic map with Area$(u) < \hbar$ we have $\|u\|_{W^{k,p}(D(\delta))} < C$.*

Proof. Proposition 8.13 implies that it's enough to prove the result for $k = 1$. Note also that the $k = 2$ version of the corollary implies Gromov's Schwarz Lemma, since Corollary 4.13 bounds the L^∞ norm of du on $D(\delta)$ as soon as the $W^{2,p}$ norm of u is bounded on $D(\delta)$.

The proof of the $k = 1$ case combines Proposition 8.12 with the strategy of the proof of Theorem 6.1. Namely, after changing coordinates on the target so that $u(0) = 0$, the equation for u to be J-holomorphic is

$$\partial_{\bar{z}} u + q(u)\partial_z u = 0$$

where the matrix-valued function q has $q(0) = 0$ and obeys a uniform C^1 bound. Let ϵ be so small that if $\|w\| < \epsilon$ then $|q(w)|$ is less than $\frac{1}{2\|T\|_p}$ where $\|T\|_p$ is the operator norm of the Calderón-Zygmund operator T on L^p. By Proposition 8.12, we may choose $\delta > 0$ so that any u as in the statement of the corollary will have $\|u(z)\| < \epsilon$

for all $|z| \le 2\delta$. Taking cutoffs χ equal to 1 on $D(\delta)$ and supported in $D(3\delta/2)$, and α equal to 1 on $D(3\delta/2)$ and supported in $D(2\delta)$, we have

$$(I + (q \circ (\alpha u)) \cdot T)\partial_{\bar{z}}(\chi u) = (\partial_{\bar{z}}\chi + q \circ (\alpha u)\partial_z \chi)u.$$

Since $(q \circ (\alpha u))$ has C^0 norm at most $\frac{1}{2\|T\|_p}$ by construction, $(I + (q \circ (\alpha u))$ is invertible on L^p with inverse having norm at most 2. Since the right hand side above obeys a u-independent L^p bound (as the cutoff functions are independent of u and $u|_{2\delta}$ is C^0-small), it follows that $\partial_{\bar{z}}(\chi u)$ obeys a u-independent L^p bound, and this gives the desired $W^{1,p}$ bound on χu (and hence also on $u|_{D(\delta)}$) by Theorem 5.10.

□

8.3. **Bubbling.** We now come to the heart of the derivation of the weak form of compactness that holds for J-holomorphic curves. The question to be addressed is: given a sequence $u_n \colon \Sigma \to M$ of J-holomorphic maps from a closed surface Σ, does there exist a subsequence converging in an appropriate sense to a J-holomorphic curve? Certainly we would need to assume that M is compact, or at least that the u_n all have image contained in a fixed compact set, since otherwise we could derive a counterexample just by considering a sequence of constant maps. It should also be plausible that we would need the energies (equivalently, the areas) of the u_n to be bounded—after all the area is presumably a continuous function with respect to our notion of convergence, and the limit should have a well-defined (and finite) area. Luckily, it is common to encounter sequences of J-holomorphic curves with area bounds: recall from the start of the course that for a J-holomorphic curve $u \colon \Sigma \to M$ the area is just the *topological* quantity $\langle [\omega], u_*[\Sigma] \rangle$, so if we assume the u_n to all represent the same homology class then an area bound comes for free.

So consider a sequence of J-holomorphic curves $u_n \colon \Sigma \to M$ such that there is a compact $K \subset M$ and a constant C such that for all n we have $Area(u_n) \le C$ and $U(\Sigma) \subset K$. Note that since $Area(u_n) = Energy(u_n) = \frac{1}{2}\int_\Sigma |du_n|^2$, so we are assuming an L^2 bound on the derivatives of the u_n. We'll see shortly (and fairly easily) based on the results at the end of the last subsection that if instead there were some $p > 2$ such that we had an L^p-bound on the du_n then the u_n would have a convergent subsequence—note that our area bound fails to give us this, but only by the slightest of margins. In particular, Corollary 8.15 implies compactness if our energy bound C less than the small but universal constant $\hbar$[16]. So what remains to be discussed is the situation where we have an energy bound which is not small, and we'll see that here there can be more complicated behavior, but also that very useful results can be obtained. First we'll justify the assertion that I just made about compactness under an L^p bound on the derivative for $p > 2$.

Theorem 8.16. *Let Σ be a fixed almost complex 2-manifold (perhaps with boundary) and let $u_n \colon \Sigma \to M$ be a sequence of J-holomorphic maps such that there is a compact subset of M (independent of n) containing the image of each u_n. Suppose that, for some $p > 2$, there is an open subset $int(U) \subset \Sigma$ and a constant C such that $\int_U |du_n|^p \le C$ for all n. Then for any compact subset $K \subset U$ and any $l \ge 1$ there is a subsequence $\{u_{n_k}\}_{k=1}^\infty$ and a J-holomorphic map $u \colon K \to M$ such that $u_{n_k} \to u$ in $W^{l,p}$.*

Proof. For any $x \in K$, applying Proposition 8.13 to a coordinate neighborhood of x with compact closure contained in U shows that there is a (smaller) neighborhood U_x of x such that for each n $\|u_n\|_{W^{l+1,p}(U_x)} \le N_x$, where N_x depends on on x but not on n because we have assumed an n-independent L^p bound on du_n. Since K is compact, it may be covered by finitely many of these coordinate charts $U_{x_1}, \dots U_{x_n}$, and then we have $\|u_n\|_{W^{l+1,p}(K)} \le N'$ where N' is the maximal value among the N_{x_i}. But then Theorem 4.16 shows that we may find a subsequence of the u_n whose restrictions to U_{x_1} converge in $W^{l,p}$, and then a subsubsequence of this subsequence which converges on restriction to U_{x_2}, and so on until we have produced a sub...subsubsequence $\{u_{n_k}\}_{k=1}^\infty$ converging in $W^{l,p}$ over all of K. Taking the limit of the equations $0 = \bar{\partial}_J u_{n_k} = du_{n_k} + J(u_{n_k}) \circ du \circ j$ shows that the limit $u = \lim_{k=1}^\infty u_{n_k}$ is J-holomorphic. (This uses that the fact that the convergence of $u_{n_k} \to u_n$ is with strength $W^{l,p}$ with $l \ge 1$ and $p > 2$ implies that $J \circ u_{n_k} \to J \circ u$ in C^0 and $du_{n_k} \to du$ in L^p). □

[16]Since I assumed that Σ was closed this isn't really saying anything since we earlier noted that any closed pseudoholomorphic curve with area less than $\hbar$ is constant; however the reasoning just given also applies if Σ has boundary

As noted earlier, the natural geometric assumption is that we have an L^2 bound on the derivatives, not an L^p bound with $p > 2$. Here is an example which both demonstrates that things can be more complicated in this context, and suggests that it should still be possible to salvage something.

Example 8.17. Consider initially the map $u_n\colon \mathbb{C}\setminus\{0\}\to\mathbb{C}$ given by $u_n(x) = \left(x, \frac{1}{nx}\right)$. Thus the "real cross-section" of the image of this map is a hyperbola, which in some sense (which you could make precise if you are so inclined) is converging to the union of the x and y axes as $n\to\infty$.

In order to work with compact domain and range, we can "projectivize" these map to the following (still denoted by u_n):

$$\begin{aligned} u_n\colon\ & \mathbb{C}P^1 \to \mathbb{C}P^2 \\ & [x:y] \mapsto \left[\frac{x}{y} : \frac{y}{nx} : 1\right] = [x^2 : y^2/n : xy], \end{aligned}$$

(where we use homogeneous coordinates throughout—the point is that taking the open subsets where the last coordinate of domain or range is 1 gives back the original map). Of course, while the maps as defined on $\mathbb{C}\setminus\{0\}$ failed to extend over zero, these maps $[x:y]\mapsto[x^2:y^2/n:xy]$ are well-defined on all of the closed manifold $\mathbb{C}P^1 = S^2$.

The maps $u_n\colon \mathbb{C}P^1\to\mathbb{C}P^2$ are clearly all homologous (indeed they are homotopic)[17] , so since they are holomorphic with respect to the standard complex structure J_0 they must all have the same energy, equal to the common value $\int_{S^2} u_n^*\omega$ where ω is the standard symplectic form on $\mathbb{C}P^2$ (we constructed this last semester and called it the Fubini-Study form). We thus have a sequence of J_0-holomorphic maps satisfying an energy bound.

Notice that the u_n are one-to-one maps, with

$$Im(u_n) = \left\{[z_0:z_1:z_2]\in\mathbb{C}P^2 \,\middle|\, z_0z_1 = \frac{z_2^2}{n}\right\}.$$

So as $n\to\infty$, it would appear that the images of the u_n converge to

$$\left\{[z_0:z_1:z_2]\in\mathbb{C}P^2\,|z_0z_1=0\right\} = \{[0:z_1:z_2]\}\cup\{[z_0:0:z_2]\} =: A_0\cup A_1,$$

i.e. to the union of the two "axes" A_0 and A_1. You should be able to convince yourself that there is no *holomorphic* map $S^2\to\mathbb{C}P^2$ having image $A_0\cup A_1$ (think about what the preimage of the single point $[0:0:1]$ of $A_0\cap A_1$ would be, and remember that nonconstant holomorphic functions have isolated zeros), which would seem to bode poorly for the prospect of the u_n converging to a J_0-holomorphic map.

Of course, since we have fairly simple formulas, we can directly check whether convergence happens. Observe that, for $x\neq 0$, and $n\to\infty$

$$u_n([x;y]) = [x^2:y^2/n:xy] \to [x^2:0;xy] = [x:0:y] \quad \text{for } x\neq 0.$$

(Since the coordinates of a point in projective space are not allowed to all be zero this isn't valid for $x=0$). Thus for all but the single point $[0:1]$ of $\mathbb{C}P^1$, the u_n converge to the function $u\colon [x:y]\mapsto[x:0:y]$. Of course, the singularity at $[0:1]$ is removable: the function $u([x:y]) = [x:0:y]$ is defined on *all* of $\mathbb{C}P^1$, with image equal to the "axis" A_1 referred to above.

So we have a J_0-holomorphic map $u\colon \mathbb{C}P^1\to\mathbb{C}P^2$ such that $u_n\to u$ on $\mathbb{C}P^1\setminus\{[0:1]\}$. Clearly u_n does *not* converge to u at $[0:1]$; indeed $u_n([0:1]) = [0:1/n:0] = [0:1:0]$ for all n whereas $u([0:1]) = [0:0:1]$. Related to this, we expected the images of the u_n to converge to the union of the two axes A_0 and A_1, but the image of u consists only of A_1; somehow, A_0 got lost.

Interestingly, we can recover A_0. Define $v_n\colon \mathbb{C}P^1\to\mathbb{C}P^2$ by $v_n([x:y]) = u_n([x:ny])$. In other words, where $\phi_n\colon \mathbb{C}P^1\to\mathbb{C}P^1$ is the holomorphic diffeomorphism $[x:y]\mapsto[x:ny]$ we are setting v_n equal to the reparametrization $v_n = u_n\circ\phi_n$. We see that

$$v_n([x:y]) = [x^2:ny^2:nxy] = [x^2/n:y^2:xy] \to [0:y^2:xy] = [0:y:x] \text{ for } y\neq 0$$

[17]It's not hard to see that the $(u_n)_*[S^2]$ is equal to 2 times the standard generator of $H_2(\mathbb{C}P^2;\mathbb{Z})\cong\mathbb{Z}$

Thus where $v([x : y]) = [0 : y : x]$, v is a holomorphic map on all of $\mathbb{C}P^1$ with image equal to A_0, such that $v_n \to v$ on $\mathbb{C}P^1 \setminus \{[0 : 1]\}$. Meanwhile, just like with the u_n, $v([0 : 1])$ has a well-defined but "wrong" value: $v([0 : 1]) = [0 : 0 : 1]$ whereas $v_n([0 : 1]) = [1 : 0 : 0]$ for all n.

Thus by reparametrizing the u_n, we were able to make them to converge on the complement of a point to the "other part" of the expected limit. The limit did extend as a holomorphic map on all of $\mathbb{C}P^1$ (just like the limit of the u_n did). It accordingly makes sense to think of the union of the two maps u and v (whose images, we note, meet at the point $[0 : 0 : 1]$ of $A_0 \cap A_1$) as a sort of generalized limit of the J_0-holomorphic curves u_n. The holomorphic sphere $v\colon \mathbb{C}P^1 \to \mathbb{C}P^2$ is considered to have "bubbled off" from the u_n.

Returning to the general situation, assume given a sequence $u_n\colon \Sigma \to M$ a sequence of J_n-holomorphic maps, where Σ is a fixed almost complex 2-manifold, (M, ω) is a symplectic manifold, and the ω-tame almost complex structures J_n converge (in C^2 norm) to the ω-tame almost complex structure J. We continue to impose the "bounded geometry" assumptions on M that were assumed at the start of this section (in particular these hold if M is closed), and we assume that the images of the u_n are all contained in some fixed compact subset of M. Moreover, assume that, for all n

$$Area(u_n) \le C.$$

Definition 8.18. *A point $z \in \Sigma$ is a* bubble point *of the sequence $\{u_n\}_{n=1}^{\infty}$ if, for every open $U \subset \Sigma$ such that $z \in U$, we have*

$$\liminf_{n\to\infty} Area(u_n|_U) \ge \hbar.$$

Clearly, if there is an open neighborhood U_0 of z on which we have an L^∞ bound $\|du_n\|_{L^\infty} < A$, then z is *not* a bubble point of $\{u_n\}_{n=1}^{\infty}$, as for open subsets $U \subset U_0$ we would have $Area(u_n|_U) < A^2 Area(U)$ which could be made arbitrarily small (in particular, smaller than $\hbar$) by taking U to be a very small disc around z.

If, on the other hand, there is no such neighborhood U_0, then after passing to a subsequence (still denoted by u_n) we could find $z_n \to z$ with $|du_n(z_n)| \to \infty$. One might think that one could have a situation where, say, for small open sets U around z the quantity $Area(u_n|_U)$ is on the order of $\hbar/2$ (or maybe even arbitrarily small), but the following important lemma shows that this is not the case.

Lemma 8.19. *With notation as above, suppose that there is a sequence $\{z_n\}_{n=1}^{\infty}$ in Σ with $z_n \to z$ and $|du_n(z_n)| \to \infty$. Then z is a bubble point of $\{u_n\}_{n=1}^{\infty}$, and there exists a nonconstant J-holomorphic sphere $v\colon S^2 \to M$.*

(The method of construction of the J-holomorphic sphere is more important than the mere existence statement, so you should pay attention to the proof.)

Proof. Fix a complex coordinate chart around z (so that z is identified with $0 \in \mathbb{C}$), and in general if w is in this coordinate chart write $D(\delta, w)$ for the disc of radius δ centered at w.

The argument will be made easier by replacing the sequence $\{z_n\}_{n=1}^{\infty}$ by a sequence $\{\zeta_n\}_{n=1}^{\infty}$, still obeying $\zeta_n \to z$ with the following property: Where $c_n = |du_n(\zeta_n)|$, there are positive $\epsilon_n \le \frac{1}{|du_n(z_n)|^{1/2}}$ such that $\epsilon_n c_n \ge |du_n(z_n)|^{1/2}$ and

$$|du_n(\zeta_n)| \ge \frac{1}{2} \sup_{D(\epsilon_n, \zeta_n)} |du_n|.$$

The idea is that the ζ_n should be something like local maxima of the function $|du_n|$, though as is reflected in the factor of $\frac{1}{2}$ above it may not be possible to arrange them to actually be local maxima. The proof that such ζ_n exist is deferred to the end of the proof of the lemma.

Note in particular that $\epsilon_n \to 0$ but $\epsilon_n c_n \to \infty$.

Now define

$$v_n\colon D(\epsilon_n c_n, 0) \to M$$
$$w \mapsto u_n\left(\zeta_n + \frac{w}{c_n}\right).$$

Several observations are in order regarding the v_n:

- If $K \subset \mathbb{C}$ is any compact subset, then for all n sufficiently large the domain of v_n contains K.
- For all n we have $|dv_n(0)| = \frac{1}{c_n}|du_n(\zeta_n)| = 1$.
- For all n and for all w in the domain $D(\epsilon_n c_n, 0)$ of v_n we have

$$|dv_n(w)| = \frac{1}{|du_n(\zeta_n)|}\left|du_n\left(\zeta_n + \frac{w}{c_n}\right)\right| \le 2.$$

The v_n have bounded area (their areas coincide with $Area(u_n|_{D(\epsilon_n,0)})$), which is at most C by our assumed area bound on u_n), and the third item above shows that their derivatives satisfy an L^∞ bound. Consequently, by Theorem 8.16, for any N we obtain a J-holomorphic map $v^N\colon D(N,0) \to M$ such that, after passing to a subsequence, we have $v_n \to v^N$ in any given $W^{l,p}$ norm on $D(N,0)$. Indeed, by a standard "diagonal" argument, we may pass to a further subsequence so that there is a J-holomorphic $v\colon \mathbb{C} \to M$ such that $v_n \to v$ in any given $W^{l,p}$ norm on any given compact subset of $\mathbb{C}$.[18]

Note that

$$Area(v) = \lim_{N\to\infty} Area(v|_{D(N,0)}) \le \lim_{N\to\infty}\left(\limsup_{n\to\infty} Area(v_n|_{D(N,0)})\right) \le C.$$

Now define

$$\tilde{v}\colon \mathbb{C}\setminus\{0\} \to M$$

by

$$\tilde{v}(z) = v(1/z).$$

The map $\tilde{v}$ is J-holomorphic and has finite area (equal to the area of v) and so by the Removal of Singularities Theorem 8.11 $\tilde{v}$ extends to a J-holomorphic map $\tilde{v}\colon \mathbb{C} \to M$. But then since $\tilde{v}(z) = v(1/z)$ wherever both are defined, $\tilde{v}$ and v now patch together to give a J-holomorphic map $v\colon S^2 \to M$ where as usual we identify S^2 with the Riemann sphere $\mathbb{C}\cup\{\infty\}$. This map v is the J-holomorphic map promised in the statement of the lemma; to see that it is nonconstant we simply note that every v_n obeys $|dv_n(0)| = 1$, and therefore $dv(0) \ne 0$.

Since $v\colon S^2 \to M$ is a nonconstant J-holomorphic map, its area is at least $\hbar$ by Proposition 8.9. So if $\epsilon > 0$ there is N so that when $n \ge N$ we have $Area(v_n) > \hbar - \epsilon$. Now $Area(v_n) = Area(u_n|_{D(\epsilon_n,\zeta_n)})$, and if U is any given open neighborhood of z then (since $\epsilon_n \to 0$ and $\zeta_n \to z$) for n sufficiently large we will have $D(\epsilon_n, \zeta_n) \subset U$. Thus for n sufficiently large $Area(u_n|_U) > \hbar - \epsilon$. Thus $\liminf_{n\to\infty} Area(u_n|_U) \ge \hbar$, which (since U was an arbitrary neighborhood of z) proves that z is a bubble point of $\{u_n\}_{n=1}^\infty$.

This proves the lemma except for the construction of the sequence of "quasi-maxima" ζ_n. This follows immediately from the following lemma (applied once for each value of n) about complete metric spaces, under the following dictionary: $z_n \leftrightarrow x$, $\zeta_n \leftrightarrow \xi$, $\frac{1}{|du_n(z_n)|^{1/2}} \leftrightarrow \delta$, $|du_n| \leftrightarrow f$, $\epsilon_n \leftrightarrow \epsilon$.

Lemma 8.20. *Let (X,d) be a complete metric space, $\delta > 0$, and $x \in X$, and $f\colon X \to [0,\infty)$ a continuous function. Then there are $\xi \in X$ and $\epsilon > 0$ with the following properties*

(i) $\epsilon \le \delta$
(ii) $d(x,\xi) < 2\delta$
(iii) $\epsilon f(\xi) \ge \delta f(x)$
(iv) $f(\xi) \ge \frac{1}{2}\sup_{B_\epsilon(\xi)} f$

Proof of Lemma 8.20. We will prove that, if the lemma failed, it would be possible to construct a sequence $\{x_k\}_{k=0}^\infty$ such that $x_0 = x$, $d(x_{k+1}, x_k) \le \frac{\delta}{2^k}$, and $f(x_{k+1}) > 2f(x_k)$. This would certainly give a contradiction, since f is assumed continuous and the x_k will converge by the completeness of the metric space, whereas the $f(x_k)$ are obviously divergent.

So assume the lemma is false and suppose inductively that we have constructed $x_0,\dots,x_k$ obeying the desired properties. It suffices to show how to construct x_{k+1}. Consider the properties (i)-(iv) of the lemma, applied with

[18] The diagonal argument goes as follows: Iteratively applying Theorem 8.16 allows us to find, for any m, a sequence $\{v_{n_{mk}}\}_{k=1}^\infty$ so that $\{v_{n_{(m+1)k}}\}_{k=1}^\infty$ is a subsequence of $\{v_{n_{mk}}\}_{k=1}^\infty$ and such that $v_{n_{mk}}$ converges to a J-holomorphic map on $D(m,0)$. The sequence $\{v_{n_{mm}}\}_{m=1}^\infty$ will then converge to a single J-holomorphic map $v\colon \mathbb{C} \to M$ in $W^{l,p}$ on any compact subset of $\mathbb{C}$.

$\xi = x_k$ and $\epsilon = \frac{\delta}{2^k}$ (our assumption that the lemma fails implies that not all of (i)-(iv) can hold). Obviously $\epsilon \le \delta$, so (i) holds. Since $\sum_{m=0}^{k} 2^{-m} < 2$, condition (ii) also holds. Since we have $f(x_{j+1}) > 2f(x_j)$ for all $0 \le j < k$, we have $f(x_k) > 2^k f(x_0)$, so since $\epsilon = \delta 2^{-k}$ (iii) holds. Hence (iv) does *not* hold, which is to say that there is $x_{k+1} \in B_{\delta 2^{-k}}(x_k)$ such that $f(x_{k+1}) > 2f(x_k)$. But these are precisely the desired properties for x_{k+1}.

This confirms that if the lemma is false we can produce a sequence $\{x_k\}_{k=1}^{\infty}$ with the stated (impossible) properties, thus producing the desired contradiction. □

□

This allows us to obtain a weak version of compactness, which we will later refine somewhat:

Theorem 8.21. *Let $u_n\colon \Sigma \to M$ be a sequence of J_n-holomorphic curves where $J_n \to J$, and assume that $Area(u_n) \le C$ for some constant C. Let $l \ge 2$ and $p > 2$. Then, after passing to a subsequence (still denoted $\{u_n\}_{n=1}^{\infty}$) the following holds. There is a finite collection of points $z^{(1)}, \dots, z^{(b)} \in \Sigma$ and a J-holomorphic map $u\colon \Sigma \to M$ such that*

- *For any compact subset $K \subset \Sigma \setminus \{z^{(1)}, \dots, z^{(b)}\}$, we have $u_n \to u$ in $W^{l,p}(K)$.*
- *For $i = 1, \dots, b$ the quantity*

$$m(z^{(i)}) = \lim_{\delta \to 0} \left(\lim_{n\to\infty} Area(u_n|_{D(\delta, z^{(i)})}) \right)$$

  is well-defined, and $m(z^{(i)}) \ge \hbar$.
-

$$\lim_{n\to\infty} Area(u_n) = Area(u) + \sum_{i=1}^{b} m(z^{(i)}) \tag{35}$$

Here $D(\delta, z)$ denotes the disc of radius δ around z (with respect to an arbitrary fixed metric on Σ).

Proof. Theorem 8.16 and Lemma 8.19 suggest how we should begin: If $\sup_n \|du_n\|_{L^\infty(\Sigma)} < \infty$, then Lemma 8.16 immediately gives the desired result with $b = 0$ (i.e. with no points $z^{(i)}$). On the other hand, if $\sup_n \|du_n\|_{L^\infty(\Sigma)} = \infty$, then after passing to a subsequence (and using the compactness of Σ) we may find a point $z^{(1)} \in \Sigma$ and a sequence $z_n \to z^{(1)}$ such that $|du_n(z_n)| \to \infty$. Then $z^{(1)}$ is a bubble point, so for all $\delta > 0$ we have

$$\liminf_{n\to\infty} Area(u_n|_{D(\delta, z^{(1)})}) \ge \hbar. \tag{36}$$

There are now two cases to consider. In the first, for every compact subset $K \subset \Sigma \setminus \{z^{(1)}\}$ there is an L^∞ bound $\|du_n\|_{L^\infty(K)} \le C_K$. Then appealing to Theorem 8.16 for each of the compact sets $\overline{K \setminus D(1/m, z^{(1)})}$ and then using a diagonal argument produces a J-holomorphic map $u\colon \Sigma \setminus \{z^{(1)}\}$ and a subsequence (still denoted u_n) such that $u_n \to u$ in $W^{l,p}$ on any compact subset of $\Sigma \setminus \{z^{(1)}\}$. Now since $\liminf Area(u_n|_{D(\delta, z^{(1)})}) \ge \hbar$ for all $\delta > 0$, the restriction of the u_n to any given compact subset of $\Sigma \setminus \{z^{(1)}\}$ has area at most $C - \hbar$. The same area bound hence holds for u. In particular, $u\colon \Sigma \setminus \{z^{(1)}\} \to M$ has finite area, so its singularity at $z^{(1)}$ may be removed by Theorem 8.11. Thus u extends to a J-holomorphic map $u\colon \Sigma \to M$. For some fixed small δ_0 we may pass to a subsequence such that the lim inf in (36) is a genuine limit for $\delta = \delta_0$; then since the u_n converge on any given $D(\delta_0, z^{(1)}) \setminus D(\delta, z^{(1)})$ for $0 < \delta < \delta_0$ the corresponding lim inf for δ will also be a limit. Now if $\epsilon > 0$, then for sufficiently small δ we will have $Area(u|_{D(\delta,z)}) \le \epsilon$, and therefore, once n is large enough, for any given $\eta \in (0, \delta)$,

$$Area(u_n|_{\{\eta \le |z| \le \delta\}}) < 2\epsilon.$$

This proves that the limits in the definition of $m(z^{(1)})$ in the statement of the lemma do exist, and (36) implies that they are at least $\hbar$. By taking $\delta > 0$ small and N large, for $n \ge N$ we can make $Area(u_n|_{\Sigma \setminus \{z^{(1)}\}})$ arbitrarily close to $Area(u)$, while $Area(u_n|_{D(\delta, z^{(1)})})$ is arbitrarily close to $m(z^{(1)})$. This proves the last statement of the theorem in this case.

This concludes the proof in the case where for every compact $K \subset \Sigma \setminus \{z^{(1)}\}$ there is an L^∞ bound on $|du_n|$. The remaining case is that in which there exists some compact $K \subset \Sigma \setminus \{z^{(1)}\}$ on which no L^∞ bound holds. Then after passing to a subsequence there are $z_n \in K$ so that $|du_n(z_n)| \to \infty$. Passing to a further subsequence and using the compactness of K, we find that $z_n \to z^{(2)}$ for some $z^{(2)} \in K$ (in particular $z^{(2)} \neq z^{(1)}$). Then Lemma 8.19 gives that $\liminf_{n\to\infty} Area(u_n|_{D(\delta,z^{(2)})}) \geq \hbar$ for all $\delta > 0$.

Proceeding inductively, assume that we have found $z^{(1)}, \ldots, z^{(k)}$ such that for each i and each $\delta > 0$ we have $\liminf_{n\to\infty} Area(u_n|_{D(\delta,z^{(i)})}) \geq \hbar$. Thus on every compact subset of $\Sigma \setminus \{z^{(1)}, \ldots, z^{(k)}\}$, for n large enough the u_n have area bounded by $C - k\hbar$. (In particular, this puts an a priori bound on k: it can't be larger than $C/\hbar$). If there is some compact subset K of $\Sigma \setminus \{z^{(1)}, \ldots, z^{(k)}\}$ on which the u_n are not bounded in L^∞, then applying Lemma 8.19 again produces a point $z^{(k+1)} \in K$ obeying the same property as the other $z^{(i)}$. However, after finitely many (at most $C/\hbar$) steps, we will be unable to produce such a point, and so it will necessarily hold that, for every compact $K \subset \Sigma \setminus \{z^{(1)}, \ldots, z^{(k)}\}$, the du_n are bounded in $L^\infty(K)$. We then apply the same reasoning as before: after passing to a suitable diagonal subsequence, we obtain a J-holomorphic map $u\colon \Sigma \setminus \{z^{(1)}, \ldots, z^{(b)}\} \to M$ such that $u_n \to u$ in $W^{l,p}$ on all compact subsets. Theorem 8.11 shows that the singularities of u may be removed to produce a J-holomorphic map $u\colon \Sigma \to M$. Using, as earlier, that the u_n converge (at least in C^1 since $l \geq 2, p > 2$) to the C^1 function u on annuli around the $z^{(i)}$, we obtain that the limits defining the quantities $m(z^{(i)})$ exist, and they are at least $\hbar$ by construction. And finally, the fact that u extends in C^1 fashion over the $z^{(i)}$ can be used to prove Equation 35: for small δ and large n it holds that $Area(u_n|_{\Sigma\setminus\cup D(\delta,z^{(i)})})$ is approximately equal to $Area(u)$, while $Area(u_n|_{\cup D(\delta,z^{(i)})})$ is approximately equal to $\sum m(z^{(i)})$. □

Lemma 8.19 and Theorem 8.21 show that any sequence $u_n\colon \Sigma \to M$ of J_n-holomorphic curves with bounded area has a subsequence which "converges modulo bubbling" (in particular, converges genuinely on compact subsets of the complement of finitely many points in Σ) to a J-holomorphic curve $u\colon \Sigma \to M$, and moreover that J-holomorphic spheres can be produced by studying the failure of convergence at any of the finitely many "bubble points." We have yet to relate these "bubbles" to the curve u. Our intention now is to argue that the combination of u with a collection of bubbles similar to those are produced in the proof of Lemma 8.19 forms a "bubble tree" which serves as an appropriate generalized limit of a subsequence of u_n.

Accordingly let $\{u_n\}_{n=1}^\infty$ be a (sub)sequence as in the conclusion of Theorem 8.21. For some fixed $i = 1, \ldots, b$ choose local coordinates around the bubble point $z^{(i)}$ (with $z^{(i)}$ identified with 0); we may scale these so that $D(1,0)$ does not contain any other $z^{(j)}$ and

$$Area(u_n|_{D(1,0)}) \leq m(z^{(i)}) + \frac{\hbar}{3}.$$

By the definition of $m(z^{(i)})$ we may, for sufficiently large n) choose a number $\delta_n > 0$ such that

$$Area(u_n|_{D(\delta_n,0)}) = m(z^{(i)}) - \frac{\hbar}{2},$$

and moreover it holds that $\lim_{n\to\infty} \delta_n = 0$. Define

$$v_n\colon\ D\left(\frac{1}{\delta_n}, 0\right) \to M$$

by

$$v_n(z) = u_n(\delta_n z).$$

Where as before we write $A(r,R)$ for the annulus $\{z \in \mathbb{C} | r \leq |z| \leq R\}$, we have

$$Area(v_n|_{D(1,0)}) \leq m(z^{(i)}) - \frac{\hbar}{2}, \quad Area(v_n|_{A(1,\delta_n^{-1})}) < \hbar.$$

For any compact $K \subset \mathbb{C} \setminus \overline{D(1,0)}$, for n sufficiently large K will be contained in the domain of v_n, and there will be $r_K > 0$ such that any $z \in K$ has $D(r_K, z) \subset \mathbb{C} \setminus \overline{D(1,0)}$. By applying Gromov's Schwarz Lemma (Corollary 8.14) to the function $w \mapsto v_n((w - z)/r_K)$ we thus see that the v_n obey a uniform (and n-independent) bound $\|dv_n\|_{L^\infty(K)} \leq C'/r_K$. So by Theorem 8.16 the v_n have a subsequence (still denoted by v_n) which converges to some limit v in any given $W^{l,p}(K)$ where K is any given compact subset of $\mathbb{C} \setminus \overline{D(1,0)}$. By writing $\mathbb{C} \setminus \overline{D(1,0)}$ as a countable union of compact subsets and using a diagonal argument we can arrange that a single subsequence of the v_n has such a limit simultaneously on each such compact subset.

At the same time, Theorem 8.21 and its proof establish that, after passing to a further subsequence, the v_n will "converge modulo bubbling" within $\overline{D(1,0)}$ as well. Thus we obtain a limiting J-holomorphic map $v\colon \mathbb{C} \to M$ such that $v_n \to v$ in $W^{l,p}$ (for arbitrary l, p) on any compact subset of the complement of a finite set of points, with each of these (new) bubble points contained in the unit disc $\overline{D(1,0)}$. The map v has area at most equal to $\limsup_{n\to\infty} Area(v_n)$, which by assumption is finite, so by considering the map $z \mapsto v(1/z)$ we see as in the proof of Lemma 8.19 that v extends to a J-holomorphic map $v\colon\ S^2 \to M$, where we make the usual identification $S^2 \cong \mathbb{C} \cup \{\infty\}$. We view this $v\colon\ S^2 \to M$ as the "main bubble" that forms from the u_n at the bubble point $z^{(i)}$; as our discussion should suggest, there may be additional bubbles that form "off the side of" this main bubble v at points in the southern hemisphere $D(1,0) \subset S^2$.

Claim 8.22. We have

$$v(\infty) = u(0) \quad \text{and} \quad Area(v|_{S^2\setminus D(1,0)}) = \frac{\hbar}{2}.$$

Proof. Note that by construction, if $\epsilon > 0$ then for any $T > 1$ we will have, for n sufficiently large,

$$Area(v_n|_{A(1,T)}) = Area(u_n|_{A(\delta_n,T\delta_n)}) \le \frac{\hbar}{2} + \epsilon$$

(using that $Area(u_n|_{D(\delta_n,0)}) = m(z^{(1)}) - \frac{\hbar}{2}$ and that $T\delta_n \to 0$). Consequently for any T, ϵ we have $Area(v|_{A(1,T)}) \le \frac{\hbar}{2} + \epsilon$. Since this holds for all T, ϵ it follows that $Area(v|_{S^2\setminus D(1,0)}) \le \frac{\hbar}{2}$. Thus for the second equation in the claim we only need to prove the inequality '$\ge$' (which in particular will imply the non-obvious fact that v is nonconstant).

Let $\epsilon > 0$. Now u_n converges in C^1 norm to the C^1 function u on any given compact subset of $D(1,0) \setminus \{0\}$. Consequently there is $\eta < 1$ such that, for all sufficiently large n, we have

$$Length(u_n|_{|z|=\eta}) < \epsilon \quad \text{and} \quad u_n(\{|z| = \eta\}) \subset B_\epsilon(u(0)).$$

By the same token, the v_n converge in C^1 norm on any given compact subset of $\mathbb{C} \setminus \overline{D(1,0)}$ to the function v, which extends to a C^1 function on $\mathbb{C} \cup \{\infty\}$. Consequently there is $\rho > 1$ such that, for all sufficiently large n, we have

$$Length(v_n|_{|z|=\rho}) < \epsilon \quad \text{and} \quad v_n(\{|z| = \rho\}) \subset B_\epsilon(v(\infty)).$$

Recall that $v_n(z) = u_n(\delta_n z)$. Consider the restrictions $v_n|_{A(\rho,\eta/\delta_n)}$. By construction (since $\eta < 1$ and $\rho > 1$) these J_n-holomorphic annuli have energy at most $\frac{\hbar}{2} + \frac{\hbar}{3} < \hbar$. Consequently Proposition 8.9 shows that $Area(v_n|_{A(\rho,\eta/\delta_n)}) \le C''\epsilon^2$ for an appropriate constant C''. But then

$$\begin{aligned} Area(v_n|_{A(1,\rho)}) &= Area(v_n|_{D(\eta/\delta_n,0)}) - Area(v_n|_{D(1,0)}) - Area(v_n|_{A(\rho,\eta/\delta_n)}) \ge Area(u_n|_{D(\eta,0)}) - \left(m(z^{(1)}) - \frac{\hbar}{2}\right) - C''\epsilon^2 \\ &\ge \frac{\hbar}{2} - 2C''\epsilon^2 \end{aligned}$$

for sufficiently large n. Thus $Area(v|_{A(1,\rho)}) \ge \frac{\hbar}{2} - 2C''\epsilon^2$, implying that (since ρ and ϵ are both independent of n) $Area(v) \ge \frac{\hbar}{2}$.

Meanwhile, since the $v_n|_{A(\rho,\eta/\delta_n)}$ have area bounded by a constant times ϵ^2 and their boundaries have length bounded by 2ϵ, Proposition 8.10 shows that their diameters are bounded by a constant times ϵ. But by construction $u(0)$ has distance at most ϵ from the image under v_n of one boundary component of $A(\rho, \eta/\delta_n)$, while $v(\infty)$ has distance at most ϵ from the image of the other boundary component. Hence $dist(v(\infty), u(0)) \le \tilde{C}\epsilon$ for an appropriate constant $\tilde{C}$. Since ϵ was arbitrary this proves that $v(\infty) = u(0)$. □

The discussion in the proof of the above claim shows moreover that, for a sufficiently small choice of the parameter ϵ, the J_n-holomorphic annuli $v_n|_{A(\rho,\eta\delta_n^{-1})}$ are homotopic to the map formed by taking the "connected sum" of the J-holomorphic discs $v|_{|z|\ge\rho}$ and $u|_{|z|\le\eta}$ (a precise formulation of this statement is left to the reader;

to prove it, construct the homotopy br travelling along geodesics contained in a small ball around the common value $u(0) = v(\infty)$. In the simple case where the v_n genuinely converge (with no bubbling) to v and where $z^{(i)}$ is the only bubble point of the u_n, this implies that, for large n, the J_n-holomorphic curves u_n are homotopic to the connected sum of $u\colon \Sigma \to M$ and $v\colon S^2 \to M$.

As discussed earlier, the sequence v_n may have bubble points, though these are all contained in the southern hemisphere $D(1,0)$. This adds some complication (especially in notation), but the bubble points of the v_n can be analyzed in much the same way as the bubble points of the u_n, in particular producing a "main bubble" at each of the bubble points (analagous to v), which again will be nonconstant by Claim 8.22 and off of which additional bubbles might form. This process must eventually terminate, since all the bubbles that arise have area at least $\hbar$ and the sum of their areas can be no larger than our bound C on the areas of the u_n. Making repeated use of the analysis of the proof of Claim 8.22 one can prove the following embellishment of Theorem 8.21.

Theorem 8.23. *Let $u_n\colon \Sigma \to M$ be a sequence of J_n-holomorphic curves where $J_n \to J$, and assume that $Area(u_n) \leq C$ for some constant C. Let $l \geq 2$ and $p > 2$. Then, after passing to a subsequence (still denoted $\{u_n\}_{n=1}^{\infty}$) the following holds. There is a finite (possibly empty) collection of points $z^{(1)}, \ldots, z^{(b)} \in \Sigma$ and a J-holomorphic map $u\colon \Sigma \to M$ such that*

- *For any compact subset $K \subset \Sigma \setminus \{z^{(1)}, \ldots, z^{(b)}\}$, we have $u_n \to u$ in $W^{l,p}(K)$.*
- *There are nonconstant J-holomorphic spheres $v^{(1)}, \ldots, v^{(b)}\colon S^2 \to M$ such that $v^{(i)}(\infty) = u(z^{(i)})$.*
- *For some finite (possibly empty) collection of nonconstant J-holomorphic spheres $w_i\colon S^2 \to M$ such that, for each i, $w_i(\infty)$ is in the image either of some other w_j or of some $v^{(j)}$, we have*

$$\lim_{n\to\infty} (u_n)_*[\Sigma] = u_*[\Sigma] + \sum_{j=1}^{b} v_*^{(j)}[S^2] + \sum_i (w_i)_*[S^2].$$

The picture to have here is that the u_n are converging to a "bubble tree" consisting of u, the $v^{(i)}$ (which are attached to u), and the w_j (which themselves could be subdivided into stages, of which the first consists of spheres attached to the $v^{(i)}$ and each successive one consists of spheres attached to spheres in the previous stage). Importantly, the total homology class of this bubble tree is equal to the limit of the homology classes of the u_n. (Of course, since $H_2(M;\mathbb{Z})$ is a discrete set this means that the $(u_n)_*[\Sigma]$ are eventually constant, which itself is a notable result since our assumptions on the u_n don't obviously imply it). As follows from this remark about the homology class, or more directly from repeated application of Claim 8.22, it is also true that the total area of the bubble tree is equal to the limit of the areas of the u_n.

There are ways of being more precise about the definition of a bubble tree and the exact meaning of the statement that a sequence of maps converges to a bubble tree, but these tend to be rather notationally cumbersome and we won't generally need them. If you're interested, one approach to this (at least where Σ is a sphere) is developed in some detail in [MS2, Chapter 5].

9. Nonsqueezing revisited

The significant amount of machinery that we've built up now allows us to fill in the missing piece (namely the proof of Lemma 3.4) of the proof of the Gromov Non-Squeezing Theorem 3.1. Below the symplectic manifold (M,Ω) has M equal to a product $S^2 \times T^{2n-2}$, and Ω equal to a "split" symplectic form $\omega_{S^2} \oplus \omega_{T^{2n-2}}$, where ω_{S^2} is any area form on S^2 (recall from last semester that these are classified up to symplectomorphism by their total area, which can be an arbitrary real number) and $\omega_{T^{2n-2}}$ is obtained by viewing T^{2n-2} as quotient $\frac{\mathbb{R}^{2n-2}}{N\mathbb{Z}^{2n-2}}$ for some real number N and setting $\omega_{T^{2n-2}}$ equal to the form induced on the quotient by the standard symplectic form $\omega_0 = \sum dx^i \wedge dy^i$ on $\mathbb{R}^{2n-2}$. Of course this Ω depends on two parameters (the area of the sphere and the parameter N in the identification of T^{2n-2} as a quotient) but the following result holds independently of those parameters. We identify S^2 with the Riemann sphere $\mathbb{C} \cup \{\infty\}$ in the usual way.

Theorem 9.1. *For any $J \in \mathcal{J}_\tau(M,\Omega)$ and any $p \in T^{2n-2}$ there is a J-holomorphic map $u\colon S^2 \to M$ such that $u(0) = (0,p)$ and $u_*[S^2] = [S^2 \times \{p\}] \in H_2(M;\mathbb{Z})$.*

Proof. Write the projections onto either factor as

$$\begin{array}{ccc} & M = S^2 \times T^{2n-2} & \\ {}^{\pi_1}\swarrow & & \searrow^{\pi_2} \\ S^2 & & T^{2n-2} \end{array}$$

For any $J \in \mathcal{J}_\tau(M, \Omega)$ define

$$\mathcal{M}_J = \left\{ u\colon\ S^2 \to M \left| \bar{\partial}_J u = 0,\ u_*[S^2] = [S^2 \times \{p\}],\ u(0) = (0, p), \pi_1(u(1)) = 1, \pi_1(u(\infty)) = \infty \right. \right\}.$$

(The last two conditions are included because, without them, for any $u \in \mathcal{M}_J$ it would hold that also $u \circ \phi \in \mathcal{M}_J$ for a continuous family of Möbius transformations ϕ, and it's techinically useful to eliminate this redundancy.)

Consider first $\mathcal{M}_J$ for J equal to the standard, split almost complex structure $J_0 = J_{S^2} \oplus J_{T^{2n-2}}$, given by operating on tangent vectors in the S^2 direction by the usual complex structure on $S^2 = \mathbb{C}P^1$, and on tangent vectors in the T^{2n-2} direction by the complex structure induced on the quotient from the standard one on $\mathbb{R}^{2n-2} = \mathbb{C}^{n-1}$. The map $u_0(z) = (z, p)$ clearly belongs to $\mathcal{M}_{J_0}$. Conversely, suppose that $u \in \mathcal{M}_{J_0}$. Now by our choice of J_0 we have $\pi_{1*} \circ J_0 = J_{S^2} \circ \pi_{1*}$ and $\pi_{2*} \circ J_0 = J_{T^{2n-2}} \circ \pi_{2*}$. Consequently the fact that $\bar{\partial}_{J_0} u = 0$ implies that $\pi_1 \circ u\colon\ S^2 \to S^2$ and $\pi_2 \circ u\colon\ S^2 \to T^{2n-2}$ are holomorphic (with respect to the standard complex structures on source and target). Now the condition on the homology class represented by u shows that $\pi_2 \circ u$ is nullhomologous, and hence constant (as it has zero energy). So since $\pi_2(u(0)) = p$, $\pi_2 \circ u$ is equal to the constant map to p. Meanwhile $\pi_1 \circ u$ is a degree-1 holomorphic map from S^2 to S^2 and hence is a Möbius transformation. Also, $\pi_1 \circ u$ takes each of the points $0, 1, \infty$ to themselves; the only Möbius transformation that acts on $0, 1, \infty$ in this fashion is the identity, so $\pi_1 \circ u(z) = z$. This proves that our arbitrary element $u \in \mathcal{M}_{J_0}$ is equal to the map $u_0(z) = (z, p)$. Thus

$$\mathcal{M}_{J_0} = \{u_0\}.$$

Let us call an almost complex structure J *regular* if, for every $u \in \mathcal{M}_J$ the linearization at u of the restriction of the Cauchy-Riemann operator $\bar{\partial}_J$ to maps $v\colon\ S^2 \to M$ having $v(0) = (0, p)$, $\pi_1(v(1)) = 1$, $\pi_1(v(\infty)) = \infty$ is surjective. The results of Section 7 (appropriately adjusted as in Section 7.4) show that the space $\mathcal{J}^{reg}$ of regular almost complex structures contains a countable intersection of open, dense subsets of $\mathcal{J}_\tau(M, \Omega)$, and in particular is dense. For any $J \in \mathcal{J}^{reg}$, $\mathcal{M}_J$ is a manifold of dimension equal to the index of the aforementioned linearization of the Cauchy-Riemann operator, which we see is

$$2n + 2\langle c_1(S^2 \times T^{2n-2}), [S^2 \times \{p\}]\rangle - 2n - 2 - 2$$

(the last three terms arise from the conditions on the values of u at (respectively) 0, 1, and ∞). Now $c_1(S^2 \times T^{2n-2}) = \pi_1^* c_1(S^2) + \pi_2^* c_1(T^{2n-2})$, so

$$\langle c_1(S^2 \times T^{2n-2}), [S^2 \times \{p\}]\rangle = \langle c_1(S^2), [S^2]\rangle = 2,$$

so the above formula for the dimension of $\mathcal{M}_J$ works out to give 0. Thus for $J \in \mathcal{J}^{reg}$, $\mathcal{M}_J$ is a 0-dimensional manifold.

We will postpone the proof of the following lemma to the end of the proof of this theorem:

Lemma 9.2. *The standard complex structure J_0 on $S^2 \times T^{2n-2}$ is regular,*

Now as explained in Section 7.4, if J_1 is another almost complex structure belonging to $\mathcal{J}^{reg}$, then for a dense set of paths $\{J_t | 0 \leq t \leq 1\}$ connecting the regular almost complex structure J_0 to J_1, the *parametrized moduli space*

$$\mathcal{M}_{\{J_t\}} := \cup_{0 \leq t \leq 1} \{t\} \times \mathcal{M}_{J_t}$$

is a 1-dimensional manifold, with boundary $\mathcal{M}_{J_0} \cup \mathcal{M}_{J_1}$.

I now claim that $\mathcal{M}_{\{J_t\}}$ is *compact*. Let $\{(t_n, u_n)\}_{n=1}^\infty$ be a sequence in $\mathcal{M}_{\{J_t\}}$. Since $[0, 1]$ is compact, we may pass to a subsequence so that $t_n \to t^*$ for some $t^* \in [0, 1]$. The u_n are thus J_{t_n} holomorphic spheres, all representing the same homology class $[S^2 \times \{p\}]$ (hence all having the same area), where $J_{t_n} \to J_{t^*}$. Thus by Theorem 8.23, after passing to a further subsequence the u_n converge modulo bubbling to some J_{t^*}-holomorphic $u\colon S^2 \to M$, producing in the limit a J_{t^*}-holomorphic bubble tree, consisting of several J_{t^*}-holomorphic spheres (including u), the sum of whose homology classes is equal to $[S^2 \times \{p\}]$. I will now argue that no bubbling in fact takes place, so that $u_n \to u$ genuinely.

Denote the spheres in the bubble tree (including u) by $\{u_\alpha\}_{\alpha\in A}$. If $u_\alpha\colon S^2 \to M$ is any one of these spheres, then $\pi_2 \circ u_\alpha\colon S^2 \to T^2$ necessarily has degree zero by topological considerations, in view of which $(u_\alpha)_*[S^2] = k_\alpha[S^2 \times \{p\}]$ for some $k_\alpha \in \mathbb{Z}$. Since $Area(u_\alpha) = \langle[\Omega], (u_\alpha)_*[S^2]\rangle$, we have $k_\alpha \geq 0$, with equality only if u_α is constant. On the other hand since the total homology class of the bubble tree equals that of the u_n, namely $[S^2 \times \{p\}]$, we must have $\sum_\alpha k_\alpha = 1$. So only one of the u_α can be nonconstant, and this nonconstant u_α must represent the class $[S^2 \times \{p\}]$. Consulting Theorem 8.23, note that all of the bubbles in the bubble tree are nonconstant. This leaves just two possibilities: either there is just one bubble and the limit-modulo-bubbling u is a constant map, or else there are no bubbles and the u_n genuinely converge to u. We now rule out the first possibility.[19] Indeed, if u were a constant map and if there were just one bubble, then in particular there would be just one bubble point for the sequence u_n. So since $u_n \to u$ uniformly on compact subsets of the complement of the bubble point, at least two of the three points $p = 0, 1, \infty$ are in a region where $\lim_{n\to\infty} u_n(p) = u(p)$. So by the conditions on $\pi_1(u_n(p))$ for $p = 0, 1, \infty$ coming from the definition of $\mathcal{M}_{J_n}$, we will have $\pi_1(u(p)) = p$ for at least two of the three points $p = 0, 1, \infty$. But obviously this implies that u is not constant, a contradiction.

Thus no bubbling occurs in the sequence u_n, and so we have genuine convergence $u_n \to u$. In other words, for an arbitrary sequence $\{(t_n, u_n)\}_{n=1}^\infty$ in $\mathcal{M}_{\{J_t\}}$ we have produced a subsequence converging to some $(t^*, u) \in \mathcal{M}_{\{J_t\}}$. Thus $\mathcal{M}_{\{J_t\}}$ is *compact*. Now we have already determined $\mathcal{M}_{\{J_t\}}$ to be a 1-manifold with boundary $\mathcal{M}_{J_0} \cup \mathcal{M}_{J_1}$. The only *compact* 1-manifolds with boundary are disjoint unions of closed intervals and circles; in particular they always have an even number of boundary points. So since $\mathcal{M}_{J_0}$ consists of the single point u_0, it follows that $\mathcal{M}_{J_1}$ must be nonempty, as otherwise $\mathcal{M}_{\{J_t\}}$ would be a compact manifold with an odd number of boundary points, which is impossible (of course if we didn't have a compactness statement this wouldn't have worked, since then $\mathcal{M}_{\{J_t\}}$ could have been a half-open interval). So we have shown that

$$\text{If } J_1 \in \mathcal{J}^{reg}, \text{ then } \mathcal{M}_{J_1} \neq \emptyset.$$

The theorem called for J to be an arbitrary element of $\mathcal{J}_\tau(M, \Omega)$, not necessarily a regular one. But since $\mathcal{J}^{reg}$ is dense, we may choose regular almost complex structures J_n ($n \in \mathbb{N}$) so that $J_n \to J$, and elements $u_n \in \mathcal{M}_{J_n}$. Since the u_n all represent the same homology class and so have the same area, the u_n converge modulo bubbling to a J-holomorphic map $u\colon S^2 \to M$. But the exact same argument used earlier shows that, essentially by topological considerations, no bubbling can in fact occur, and so $u_n \to u$ genuinely (in any Sobolev norm). In particular, $u(0) = (0, p)$ and $u_*[S^2] = [S^2 \times \{p\}]$, as desired. This completes the proof of the Theorem except for the proof of Lemma 9.2. □

Proof of Lemma 9.2. Since $\mathcal{M}_{J_0} = \{u_0\}$, we need to show that the linearization of the Cauchy-Riemann operator (associated to the standard split complex structure on $S^2 \times T^{2n-2}$) at u_0 is surjective. Denote this linearization by D. Recall that we are restricting attention to maps $u\colon S^2 \to S^2 \times T^{2n-2}$ such that $u(0) = (0, p)$, $\pi_1(u(1)) = 1$, and $\pi_1(u(\infty)) = \infty$; thus the domain of the linearization D consists of sections ξ of u_0^*TM with the property that these conditions are preserved upon moving in the direction ξ.

We have

$$u_0^*TM = (\pi_1 \circ u_0)^*TS^2 \oplus (\pi_2 \circ u_0)^*TT^{2n-2}.$$

Now $\pi_2 \circ u_0$ is just the constant map to p, while $\pi_1 \circ u_0$ is the identity, so in fact

$$u_0^*TM = TS^2 \oplus T_pT^{2n-2}.$$

[19] I sort of elided this possibility in class, but in a more general context it can happen: for instance consider $v_n\colon \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}$ defined by $v_n(z) = nz$. This sequence converges modulo bubbling to the constant map $v(z) = \infty$, with the unique "bubble point" of the sequence being zero. In our context, the fact that we fixed the values of $\pi_1 \circ u$ at 0, 1, and ∞ prevents this sort of behavior.

The tangent space T_pT^{2n-2} at the point p can just be identified with $\mathbb{C}^{n-1}$ (using the standard complex structure $J_{T^{2n-2}}$). Thus taking into account the conditions at 0, 1, ∞, the linearization D is a map

$$D\colon\ W^{1,p}(S^2, TS^2) \times W^{1,p}(S^2, \mathbb{C}^{n-1}) \to L^p(\overline{Hom}_{J_0}(TS^2 u_0^* TM)).$$

Now as noted in the proof of the theorem, D is a Fredholm operator of index zero. Thus $\dim(\ker D) = \dim(\operatorname{coker} D)$, so to show that D is surjective it is enough to show that D is injective. Consider then an arbitrary element $\xi = (\xi_1, \xi_2)$ of $\ker D$, so that ξ_1 is a section of TS^2 (i.e. a vector field on S^2) and $\xi_2\colon\ S^2 \to \mathbb{C}^{n-1}$ is some function. The conditions on the maps u that we are considering at $0, 1, \infty$ show that $\xi_1(0) = \xi_2(0) = 0$ and that $\xi_1(1) = \xi_1(\infty) = 0$. Since J_0 is the standard split integrable almost complex structure on $S^2 \times T^{2n-2}$, the operator $\bar{\partial}_{J_0}$ is just the standard Cauchy-Riemann operator acting on maps from the complex manifold $S^2 = \mathbb{C}P^1$ to the complex manifold $\mathbb{C}P^1 \times T^{2n-2}$. Note that when this standard Cauchy–Riemann operator is written out in local coordinates, it is a linear operator, and so the linearization in local coordinates coincides with this local-coordinate expression for the operator. Hence to say that $D(\xi_1, \xi_2) = 0$ is to say that ξ_1 is a holomorphic vector field on $S^2 = \mathbb{C}P^1$ and that $\xi_2\colon\ \mathbb{C}P^1 \to \mathbb{C}^{n-1}$ is a holomorphic function.

Now any holomorphic function from $\mathbb{C}P^1$ to $\mathbb{C}^{n-1}$ is constant: indeed, since $\mathbb{C}P^1$ is compact any such function attains it maximum modulus somewhere, and the maximum principle from complex analysis quickly implies that the set on which it attains its maximum is open, so since this set is also closed it is all of $\mathbb{C}P^1$, and so the open mapping theorem forces the map to be constant. But then since $\xi_2(0) = 0$ it follows that $\xi_2 = 0$ identically.

As for ξ_1, note that ξ_1 is a holomorphic vector field on $\mathbb{C}P^1$ which vanishes at 0, 1, and ∞. I claim that any holomorphic vector field on $\mathbb{C}P^1$ which vanishes at 3 points in fact vanishes everywhere. There are various ways of seeing this (you might try to think of other ways yourself), but here is an elementary argument. Suppose that ξ_1 is a holomorphic vector field vanishing at the 3 points 0, 1, and ∞ (if it vanished at a different set of 3 points we could apply a Möbius transformation to reduce to this case). Consider the expression for ξ_1 in local coordinates on $\mathbb{C}$: we will have, for $z \in \mathbb{C}$, $\xi_1(z) = f(z)\partial_z$ for some entire function f, where $\partial_z = \frac{1}{2}(\partial_x - i\partial_y)$ as usual. Since $\xi_1(0) = \xi_1(1) = 0$ we can write, for some entire function g, $f(z) = z(z-1)g(z)$.

Now let $w = \frac{1}{z}$ be the standard local holomorphic coordinate near $\infty \in \mathbb{C}P^1$. We have

$$dw = -\frac{1}{z^2}dz, \text{ and so } \partial_w = -z^2\partial_z = -\frac{1}{w^2}\partial_z.$$

Hence, on $\mathbb{C} \setminus \{0\}$ (i.e. on the intersection of the two standard coordinate patches),

$$\xi = z(z-1)g(z)\partial_z = -\frac{1}{w}\left(\frac{1}{w} - 1\right)w^2 g\left(\frac{1}{w}\right)\partial_w = (1-w)g\left(\frac{1}{w}\right)\partial_w.$$

Hence the fact that $\xi_1(\infty) = 0$ (where the point ∞ corresponds to $w = 0$) shows that $g(z) \to 0$ as $|z| \to \infty$. But g is an entire function, and the only entire function with this property is the zero function. Thus ξ_1 is identically zero.

We have thus shown that the only element of $\ker D$ has both its entries ξ_1 and ξ_2 identically equal to 0. Thus $\dim\ker D = 0$, and so since D has Fredholm index zero it also holds that $\dim\operatorname{coker} D = 0$. In other words, D is surjective. Since D is the linearization of the Cauchy-Riemann operator at the only point of $\mathcal{M}_{J_0}$, this proves that J_0 is regular in the sense of the proof of Theorem 9.1.

□

References

[AL] M. Audin and J. Lafontaine, editors. *Holomorphic curves in symplectic geometry*. Progr. Math. **117**, Birkhäuser, 1994.

[Ev] L. C. Evans. *Partial Differential Equations*. Grad. Stud. Math. **19**, AMS, Providence, 1998.

[CaZy] A. Calderón and A. Zygmund. *On the existence of certain singular integrals*. Acta Math. **88** (1952), 85–139.

[Gr] M. Gromov. *Pseudo-holomorphic curves in symplectic manifolds*. Invent. Math. **82** (1985), 307–347.

[Lan] S. Lang. *Introduction to Differentiable Manifolds*. Springer, 2002.

[Law] H. B. Lawson. *Lectures on Minimal Submanifolds*. Publish or Perish, 1980.

[MS1] D. McDuff and D. Salamon. *Introduction to Symplectic Topology*, 2nd ed. Oxford Mathematical Monographs, 1998.

[MS2] D. McDuff and D. Salamon. *J-holomorphic Curves and Symplectic Topology* AMS, Providence, 2004.

[Meg] R. Megginson. *An introduction to Banach Space Theory*. Springer, GTM **183**, 1998.
[MW] M. Micallef and B. White. *The structure of branch points in area-minimizing surfaces and in pseudo-holomorphic curves*. Ann. Math. **141** (1995), 35–85.
[Mr] T. Mrowka. *A Brief Introduction to Linear Analysis: Fredholm Operators*, Lectures 16-17, 18.965, Fall 2004, MIT OpenCourseWare
[Ra] J. Ratzkin. *The montonicity formula*, http://www.math.uga.edu/˜jratzkin/papers/monotonicity-formula.pdf
[Sik] J.-C. Sikorav. *Singularities of J-holomorphic curves*. Math. Zeitschrift **226** (1997), 359–373.
[Y] F. Yao. *A new approach to L^p estimates for Calderón–Zygmund type singular integrals*. Arch. Math. (Basel) **92** (2009), 137–146.