J. DIFFERENTIAL GEOMETRY
25 (1987) 167–197

THE YAMABE PROBLEM ON CR MANIFOLDS

DAVID JERISON & JOHN M. LEE

1. Introduction

The geometry of CR manifolds, the abstract models of real hypersurfaces in complex manifolds, has recently attracted much attention. This geometry is richest when the CR manifold is "strictly pseudoconvex," in which case there are many parallels with Riemannian geometry. (See the recent survey article by M. Beals, C. Fefferman, and R. Grossman [2] for a nice overview of these parallels.)

There are two complementary approaches to the study of CR geometry. The first is via the *Levi form*, a hermitian metric on complex tangent vectors; the second is via the *Fefferman metric*, a Lorentz metric on a natural circle bundle over the manifold.

Both of these geometric structures are determined only up to a conformal multiple by the CR structure. A choice of multiple of the Levi form is called a *pseudohermitian structure* on the manifold; such a choice also determines the multiple of the Fefferman metric.

The state of affairs suggests that, in order to find CR-invariant information, we proceed by analogy with conformal Riemannian geometry, in which a Riemannian metric is given only up to a conformal factor. A common strategy in conformal geometry is to choose a particular conformal representative for the metric which is normalized so as to simplify some aspect of the geometry. For example, the *Yamabe problem* on a conformal Riemannian manifold is to find a conformal representative for the metric that has constant scalar curvature. It is this problem that we generalize to CR manifolds in this paper.

An obvious analogue of the Yamabe problem for a CR manifold would be to find a pseudohermitian structure for which the associated Fefferman metric

Received June 20, 1985. The research of both authors was supported in part by National Science Foundation research grants MCS-82-02127 and MSC-84-04131 respectively.

has constant scalar curvature. Alternatively, S. Webster [26] has defined a scalar curvature associated with a pseudohermitian structure, and it is shown in [16] that these two notions of scalar curvature coincide.

Thus we are led to the following *CR Yamabe problem*: On a compact, strictly pseudoconvex CR manifold, find a choice of pseudohermitian structure with constant (Webster or Fefferman) scalar curvature.

Our main result is Theorem 3.4, which can be summarized as follows: There is a numerical CR invariant $\lambda(N)$ associated with every compact, orientable, strictly pseudoconvex $2n+1$ dimensional CR manifold N, which is always less than or equal to the value corresponding to the sphere S^{2n+1} in $\mathbf{C}^{n+1}$. If $\lambda(N)$ is strictly less than $\lambda(S^{2n+1})$, then N admits a pseudohermitian structure with constant scalar curvature.

This result was announced in [14]. S. S. Chern and R. Hamilton [5], studying contact structures on 3-manifolds, have independently obtained a result which is equivalent to our existence assertion in the case $\lambda(N) \leqslant 0$ and $n = 1$.

The proof of the main theorem in many respects parallels that of the analogous theorem for conformal Riemannian manifolds, due to H. Yamabe [27], N. Trudinger [24], and T. Aubin [1]. In §2 we describe the Riemannian theorem and sketch its proof, as a way of charting our course. At the end of the section, we explain a technical difficulty in the CR case, which makes our proof longer.

§3 contains the definitions and facts about CR and pseudohermitian structures we will need, and the proof of the CR invariance of $\lambda(N)$.

In §4, we describe normal coordinates due to G. Folland and E. Stein [9] which closely approximate the given pseudohermitian structure of N near a point with that of the Heisenberg group, and use these to prove that $\lambda(N) \leqslant \lambda(S^{2n+1})$.

In §5, we summarize some Sobolev-type inequalities and regularity estimates for CR manifolds due to Folland and Stein, and use these to prove various regularity theorems for the Yamabe equation (3.2). §6 contains the proof of existence of solutions under the assumption $\lambda(N) < \lambda(S^{2n+1})$.

In §7 we describe our progress to date on the question of uniqueness of solutions to the CR Yamabe problem. In the case of the sphere, this is the problem of identifying the extremals for the Heisenberg group analogue of the classical Sobolev lemma.

We would like to thank Karen Uhlenbeck for first introducing us to the Yamabe problem, and Sigurdur Helgason, who introduced us to the conformally invariant Laplacian on a Lorentz manifold, in connection with his work on Huygens' phenomenon [10].

It will become apparent throughout the rest of this paper that there is a far-reaching analogy between conformal and CR geometries. The following table summarizes the most important parallels that will be discussed below.

Conformal Geometry	CR Geometry
Riemannian manifold (M, g)	Pseudohermitian manifold (N, θ)
Euclidean space $\mathbf{R}^m$	Heisenberg group $\mathbf{H}^n$
m-sphere S^m in $\mathbf{R}^{m+1}$	$2n+1$-sphere S^{2n+1} in $\mathbf{C}^{n+1}$
Stereographic projection	Cayley transform
Riemannian normal coordinates	Folland-Stein normal coordinates
Scalar curvature K	Webster scalar curvature R
Laplace-Beltrami operator Δ	Sublaplacian $\Delta_b(\mathrm{Re}\,\Box_b$ on functions)
Sobolev spaces L_k^r	Folland-Stein spaces S_k^r
Sobolev embedding $L_1^2 \subset L^q$, $\frac{1}{q} = \frac{1}{2} - \frac{1}{m}$	Sobolev embedding $S_1^2 \subset L^p$, $\frac{1}{p} = \frac{1}{2} - \frac{1}{2n+2}$
Conformal change $\tilde{g} = \phi^{q-2} g$	Change of contact form $\tilde{\theta} = u^{p-2}\theta$
Conformal invariant $\mu(M)$	CR invariant $\lambda(N)$
Yamabe equation: $a_m \Delta\phi + K\phi = \mu\phi^{q-1}$	CR Yamabe equation: $b_n \Delta_b u + Ru = \lambda u^{p-1}$

2. The Riemannian Yamabe problem

Let (M, g) be a Riemannian manifold of dimension $m \geqslant 3$. If $\tilde{g} = \phi^{q-2} g$ (with $q = 2m/(m-2)$) is a new metric conformal to g, the scalar curvature $\tilde{K}$ of $\tilde{g}$ is given by

$$\tilde{K} = \phi^{1-q}(a_m \Delta\phi + K\phi), \qquad a_m = 4(m-1)/(m-2),$$

in which Δ is the Laplace-Beltrami operator of g and K its scalar curvature (see, e.g., [1]). Thus the problem of finding a conformal metric with constant scalar curvature $\tilde{K} \equiv \mu$ is equivalent to finding a positive, C^∞ solution ϕ to the *Yamabe equation*:

$$a_m \Delta\phi + K\phi = \mu\phi^{q-1}. \tag{2.1}$$

This problem has the following nice variational formulation. Consider the constrained variational problem

$$\mu(M) = \inf\left\{ \int_M \left(a_m |d\phi|^2 + K\phi^2\right) dV_g : \int_M |\phi|^q \, dV_g = 1 \right\}. \tag{2.2}$$

One computes readily that the Euler-Lagrange equation for (2.2) is the Yamabe equation, provided $\phi \geqslant 0$. Thus one is led to search for extremals for (2.2).

One of the major milestones in the solution of the Yamabe problem was the following theorem, due to H. Yamabe [27], N. Trudinger [24], and T. Aubin [1].

Theorem 2.3. *Let (M, g) be a compact Riemannian manifold of dimension $m \geqslant 3$.*

(a) *$\mu(M)$, defined by* (2.2), *depends only on the conformal class of g.*

(b) *$\mu(M) \leqslant \mu(S^m)$, in which the sphere S^m has the standard metric.*

(c) *If $\mu(M) < \mu(S^m)$, then the infimum in* (2.2) *is attained by a positive, C^∞ solution to* (2.1). *Thus the metric $\tilde{g} = \phi^{q-2}g$ has constant scalar curvature $\mu(M)$.*

Aubin also proved that $\mu(M) < \mu(S^m)$ in all cases in which M is not locally conformally flat and $m \geqslant 6$. More recently, R. Schoen [21] has completed the solution of the Yamabe problem by proving that $\mu(M) < \mu(S^m)$ unless M is the sphere.

The proof of Theorem 2.3(a) consists of the fundamental observation that problem (2.2) is conformally invariant in the following sense. Under the conformal change of metric $\tilde{g} = t^{q-2}g$, if we let $\tilde{\Delta}$ and $\tilde{K}$ denote the Laplacian and scalar curvature of $\tilde{g}$, then we have the transformation law (cf. [1]):

$$(a_m\tilde{\Delta} + \tilde{K})\tilde{\phi} = t^{1-q}(a_m\Delta + K)\phi, \quad \text{with } \tilde{\phi} = t^{-1}\phi. \tag{2.4}$$

It follows that the integral in (2.2) is unchanged if we replace g by $\tilde{g}$ and ϕ by $\tilde{\phi}$, and thus $\mu(M)$ is a conformal invariant.

We remark that the transformation law (2.4) can be interpreted as saying that the operator $(a_m\Delta + K)$ (the "conformally invariant Laplacian") acts naturally as an operator on certain bundles of densities on M, and that the functional in (2.2) is really a conformally invariant functional on densities. We will elaborate on this point of view in the context of CR manifolds in §3.

The analysis of (2.2) begins with a thorough understanding of the special case of the sphere S^m in $\mathbf{R}^{m+1}$. The conformal change of variables given by stereographic projection coupled with the transformation law (2.4) converts the variational problem on S^m to the more familiar problem on $\mathbf{R}^m$:

$$\mu(S^m) = \inf\left\{ a_m\int_{\mathbf{R}^m} |df|^2\,dx : \int_{\mathbf{R}^m} |f|^q\,dx = 1 \right\}. \tag{2.5}$$

This is just the problem of finding the best constant and extremal functions for Sobolev's inequality on $\mathbf{R}^m$:

$$\mu(S^m)\left(\int_{\mathbf{R}^m} |f|^q\,dx\right)^{2/q} \leqslant a_m\int_{\mathbf{R}^m} |df|^2\,dx.$$

Aubin proved that the extremals exist and have the form

$$\left(a + b|x - x_0|^2\right)^{-(m-2)/2}$$

(see also Talenti [23]). On a compact Riemannian manifold M, using Riemannian normal coordinates and the dilation invariance of problem (2.5), one can transplant an approximate extremal function for (2.5) from $\mathbf{R}^m$ to a small neighborhood on M and deduce that $\mu(M) \leqslant \mu(S^m)$.

The proof of Theorem 2.3(c) uses the Sobolev lemma for compact Riemannian manifolds. Consider the Sobolev space $L_1^2(M)$ with norm

$$\|f\|_{L_1^2(M)}^2 = \int_M \left(|df|^2 + f^2\right) dV_g.$$

The Sobolev lemma asserts, in part, that for $1/s \geqslant 1/2 - 1/m$, $L_1^2(M)$ is continuously embedded in the Lebesgue space $L^s(M)$, with compact inclusion if $1/s > 1/2 - 1/m$. If we choose a minimizing sequence $\phi_i \in L_1^2(M)$ for problem (2.2), the Sobolev lemma implies that $\{\phi_i\}$ is uniformly bounded in $L_1^2(M)$, and so a subsequence converges weakly to $\phi \in L_1^2(M)$. The main difficulty is that the exponent q is exactly the critical value for which the inclusion $L_1^2(M) \subset L^q(M)$ is *not* compact. Thus we cannot guarantee that the constraint $\int_M |\phi|^q \, dV_g = 1$ is preserved in the limit. On the other hand, if we consider the perturbed problem

$$a_m \Delta\phi_{(s)} + K\phi_{(s)} = \mu_s \phi_{(s)}^{s-1}, \qquad 2 \leqslant s < q,$$

the compactness of $L_1^2(M) \subset L^s(M)$ guarantees that a subsequence converges strongly in the L^s norm to $\phi_{(s)} \in L_1^2(M)$, so the constraint is preserved. Iteration of standard L^p estimates for the Laplace-Beltrami operator and the L^p version of the Sobolev lemma shows that $\phi_{(s)}$ is smooth; the strong maximum principle implies that $\phi_{(s)}$ is strictly positive.

The remaining step is to show that, as s tends to q, $\phi_{(s)}$ tends to a smooth, positive function ϕ. Aubin completed the proof with the help of the observation that the best constant in the Sobolev inequality is the same for all compact manifolds in the following sense: if $\mu = \mu(S^m)$ is defined by (2.5), then for any M and any $\varepsilon > 0$, there exists $C_{M,\varepsilon}$ such that

$$(2.6) \qquad (\mu - \varepsilon)\left(\int_M |f|^q \, dV_g\right)^{2/q} \leqslant a_m \int_M |df|^2 \, dV_g + C_{M,\varepsilon} \int_M |f|^2 \, dV_g$$

for all $f \in L_1^2(M)$. Inequality (2.6) is proved by transferring the inequality from Euclidean space to the manifold via Riemannian normal coordinates and a partition of unity.

Applying (2.6) to $\phi_{(s)}$, with ε chosen so that $\mu - \varepsilon > \mu_s$ for s sufficiently close to q, one can show that $\|\phi_{(s)}\|_2^2$ is bounded away from zero as $s \to q$, thus completing the proof.

The main technical difficulty in the CR case is that we have been unable to prove the analogue of (2.6). The problem is that in normal coordinates, the CR analogue of the gradient on the manifold is not comparable to that on the flat model.

An alternative proof of Theorem 2.3(c) has been given by K. Uhlenbeck [25], which does not require the result that the Sobolev constant is independent of M. Instead, assuming $\phi_{(s)}$ does not converge, she used Riemannian normal coordinates to transplant $\phi_{(s)}$ to $\mathbf{R}^n$ in such a way that the transplanted functions converge in $C^1(\mathbf{R}^n)$. The limit function ϕ then is shown to contradict Sobolev's inequality on $\mathbf{R}^n$ if $\mu(M) < \mu(S^m)$. It is this method that we shall adapt to the CR case in §6. The technical difficulty is overcome by obtaining uniform estimates for a family of nonequivalent "gradients".

3. CR manifolds

Let N be an orientable, real, $(2n+1)$-dimensional manifold. A *CR structure* on N is given by a complex n-dimensional subbundle $T_{1,0}$ of the complexified tangent bundle $\mathbf{C}TN$ of N, satisfying $T_{1,0} \cap T_{0,1} = \{0\}$, where $T_{0,1} = \overline{T}_{1,0}$. We will assume throughout that the CR structure is *integrable*; that is, it satisfies the formal Frobenius condition $[T_{1,0}, T_{1,0}] \subset T_{1,0}$. We set $G = \mathrm{Re}(T_{1,0} + T_{0,1})$, so that G is a real $2n$-dimensional subbundle of TN. G carries a natural complex structure map $J: G \to G$ given by $J(V + \overline{V}) = i(V - \overline{V})$ for $V \in T_{1,0}$.

Let $E \subset T^*M$ denote the real line bundle $G^{\perp}$. Because we assume N is orientable, and G is oriented by its complex structure, E has a global nonvanishing section. Associated with each such section θ is the real symmetric bilinear form L_θ on G:

$$L_\theta(V, W) = 2\langle d\theta, V \wedge JW \rangle, \qquad V, W \in G,$$

called the *Levi form* of θ. L_θ extends by complex linearity to $\mathbf{C}G$, and induces a hermitian form on $T_{1,0}$, which we write

$$L_\theta(V, \overline{W}) = \langle -2i\, d\theta, V \wedge \overline{W} \rangle, \qquad V, W \in T_{1,0}.$$

If θ is replaced by $\tilde{\theta} = f\theta$, L_θ changes conformally by $L_{\tilde{\theta}} = fL_\theta$. We will assume that N is *strictly pseudoconvex*, that is, that L_θ is positive definite for a suitable choice of θ. In this case, θ defines a contact structure on M, and we call θ a *contact form*. We denote by E^+ the $\mathbf{R}^+$-bundle of positive multiples of such a contact form.

The most important example of an integrable CR structure is of course that induced by an embedding of N in a complex manifold Ω, in which case $T_{1,0} = T_{1,0}\Omega \cap \mathbf{C}TN$. If ρ is a defining function for N, then one choice for the contact form is $\theta = i(\overline{\partial} - \partial)\rho$.

A *pseudohermitian structure* on N is a CR structure together with a given contact form θ. With this choice, N is equipped with a natural volume form $\theta \wedge d\theta^n$ (nonzero because N is strictly pseudoconvex). The inner product L_θ determines an isomorphism $G \cong G^*$, which in turn determines a dual form L_θ^*

on G^*, which extends naturally to T^*N. This defines a norm $|\omega|_\theta$ on real 1-forms ω, which satisfies

$$|\omega|_\theta^2 = L_\theta^*(\omega, \omega) = 2\sum_{j=1}^{n} \left|\omega(Z_j)\right|^2,$$

whenever $Z_1, \cdots, Z_n$ form an orthonormal basis for $T_{1,0}$ with respect to the Levi form (see [16]). (Note that this normalization of $|\omega|_\theta$ differs from that given in [14] by a factor of 2. The definition we have chosen here, in terms of the dual metric L_θ^*, is the more natural one.)

The *subplacian operator* Δ_b is defined on real functions $u \in C^\infty(N)$ (cf. [16]) by

$$\int_N (\Delta_b u) v\theta \wedge d\theta^n = \int_N L_\theta^*(du, dv)\theta \wedge d\theta^n \quad \text{for all } v \in C_0^\infty(N).$$

Since evidently $|\theta|_\theta = 0$, L_θ^* is degenerate on T^*N, and so the operator Δ_b is subelliptic rather than elliptic. It is shown in [16] that $\Delta_b = \operatorname{Re}\Box_b$, where $\Box_b$ is the Kohn-Spencer Laplacian [15] acting on functions.

The *Fefferman metric* of (N, θ) is a pseudo-Riemannian metric g of Lorentz signature, defined on the total space of a certain circle bundle C over N. It was first introduced by C. Fefferman [8] in the case of an embedded hypersurface in $\mathbf{C}^{n+1}$; various intrinsic characterizations of g on an abstract CR manifold are known ([4], [7], [16]).

If θ is replaced by $\tilde{\theta} = r^{p-2}\theta$, with $p = 2 + 2/n$, then g goes over to $\tilde{g} = r^{p-2}g$, so the conformal class of the Fefferman metric is a CR invariant of N. (The reason for representing the conformal factor in this strange way is that it simplifies the transformation laws below.) As a consequence of (2.4), if $\Box$ denotes the (Laplace-Beltrami) wave operator of g, and K its scalar curvature, then we have the transformation law

$$(a_{2n+2}\tilde{\Box} + \tilde{K})\tilde{\phi} = r^{1-p}(a_{2n+2}\Box + K)\phi,$$

with $\tilde{\phi} = r^{-1}\phi$.

Because the metric g is invariant under the action of S^1 on C, the operator $\Box$ pushes forward under projection π: $C \to N$ to an operator $\pi_*\Box$ on N. It is easy to verify (see [16]) that $\pi_*\Box = 2\Delta_b$. Moreover, K is constant on the fibers of C by S^1-invariance, so it projects to a function π_*K on N. It is shown in [16] that $\pi_*K = (2(2n+1)/(n+1))R$, where R is the Webster scalar curvature of the pseudohermitian structure θ. It follows that the operator $(b_n\Delta_b + R)$ on N, with $b_n = ((n+1)/2(2n+1))a_n = 2 + 2/n$, satisfies the transformation law

$$(3.1) \qquad (b_n\tilde{\Delta}_b + \tilde{R})\tilde{u} = r^{1-p}(b_n\Delta_b + R)u,$$

with $\tilde{u} = r^{-1}u$.

If we substitute $r = u$ in (3.1), we obtain the transformation law for the Webster scalar curvature R:

$$\tilde{R} = u^{1-p}(b_n\Delta_b + R)u$$

when $\tilde{\theta} = u^{p-2}\theta$. Thus if θ is a given contact form and u a positive C^∞ function on N, a necessary and sufficient condition for the contact form $\tilde{\theta} = u^{p-2}\theta$ to have constant Webster scalar curvature $\tilde{R} \equiv \lambda$ is that u satisfy

$$b_n\Delta_b u + Ru = \lambda u^{p-1}. \tag{3.2}$$

This is the *CR Yamabe equation.*

As with the Riemannian Yamabe equation, (3.2) is the Euler-Lagrange equation for the constrained variational problem

$$\lambda(N) = \inf\{A_\theta(u): B_\theta(u) = 1\} \tag{3.3}$$

in which

$$A_\theta(u) = \int_N \left(b_n|du|_\theta^2 + Ru^2\right)\theta \wedge d\theta^n, \qquad B_\theta(u) = \int_N |u|^p\,\theta \wedge d\theta^n.$$

(If N is compact, Hölder's inequality shows that $\lambda(N) > -\infty$.)

Our main theorem is:

Theorem 3.4. *Let N be a compact, orientable, strictly pseudoconvex, integrable CR manifold of dimension $2n+1$, θ any contact form on N, and define $\lambda(N)$ by* (3.3).

(a) *$\lambda(N)$ depends only on the CR structure of N, not the choice of θ.*

(b) *$\lambda(N) \leqslant \lambda(S^{2n+1})$, in which $S^{2n+1} \subset \mathbf{C}^{n+1}$ is the sphere with its standard CR structure.*

(c) *If $\lambda(N) < \lambda(S^{2n+1})$, then the infimum in* (3.3) *is attained by a positive C^∞ solution to* (3.2). *Thus the contact form $\tilde{\theta} = u^{p-2}\theta$ has constant Webster scalar curvature $R \equiv \lambda(N)$.*

Part (a) follows immediately if we observe that with the change of contact form $\tilde{\theta} = r^{p-2}\theta$ and the substitution $\tilde{u} = r^{-1}u$, $\tilde{\theta} \wedge d\tilde{\theta}^n = r^p\theta \wedge d\theta^n$, and so as a consequence of the transformation law (3.1), $B_{\tilde{\theta}}(\tilde{u}) = B_\theta(u)$ and $A_{\tilde{\theta}}(\tilde{u}) = A_\theta(u)$.

Part (b) will be proved in §4, and part (c) in §6 (Theorem 6.5).

To conclude this section, we would like to observe that the transformation law (3.1) can be expressed more invariantly in terms of densities. We introduce density bundles E^α on N, with fiber E_x^α at $x \in N$ given by

$$E_x^\alpha = \{\mu : E_x^+ \to \mathbf{R} : \mu(\lambda\theta) = \lambda^{-\alpha}\mu(\theta) \text{ for all } \lambda > 0\}.$$

E^α will be called the *bundle of densities of CR weight* α on N. Observe that if θ is a contact form (section of E), then E^1 is spanned by μ_θ, given by

$$\mu_\theta(\theta') = \theta/\theta', \qquad \theta' \in E^+.$$

The correspondence $\theta \mapsto \mu_\theta$ gives a natural (CR invariant) isomorphism $E^1 \cong E$; similarly, $\theta \wedge d\theta^n \mapsto \mu_\theta^{n+1}$ is a linear isomorphism between E^{n+1} and the bundle Ω_N of ordinary densities on N.

Once a contact form θ is chosen on N, a section of E^α can be represented as $u\mu_\theta^\alpha$, where u transforms by $\tilde{u} = r^{-\alpha}u$ when $\tilde{\theta} = r\theta$ and $\tilde{u}\mu_{\tilde{\theta}}^\alpha = u\mu_\theta^\alpha$.

As an immediate consequence of (3.1), therefore, we obtain the following proposition.

Proposition 3.5. *The linear operator Δ_b^c: $E^{n/2} \to E^{n/2+1}$, given by*

$$\Delta_b^c\left(u\mu_\theta^{n/2}\right) = \left(b_n\Delta_b u + Ru\right)\mu_\theta^{n/2+1},$$

is well defined, independently of the choice of θ.

We call Δ_b^c the *CR invariant Laplacian of* N. The CR invariance of $\lambda(N)$ can also be seen from the easily verified fact that

$$\lambda(N) = \inf\left\{\int_N (\Delta_b^c\phi) \otimes \phi : \phi \text{ a positive } C^\infty \text{ section of } E^{n/2} \text{ with } \int_N \phi^p = 1\right\}.$$

4. The Heisenberg group and normal coordinates

The Heisenberg group $\mathbf{H}^n$ is the Lie group whose underlying manifold is $\mathbf{C}^n \times \mathbf{R}$ with coordinates $(z,t) = (z^1,\cdots,z^n,t)$ and whose group law is given by

$$(z,t)(z',t') = (z + z', t + t' + 2\operatorname{Im} z \cdot \bar{z}'),$$

where $z \cdot \bar{z}' = \sum_{j=1}^n z^j\bar{z}^{j\prime}$. We will also denote elements of $\mathbf{H}^n$ by x and y and Lebesgue measure on $\mathbf{C}^n \times \mathbf{R}$ by dx or dy. Convolution in $\mathbf{H}^n$ is given by

$$f * g(x) = \int_{\mathbf{H}^n} f(xy^{-1})g(y)\,dy,$$

defined, for instance, for $f \in C_0^\infty(\mathbf{H}^n)$ and g locally integrable.

Define a norm on $\mathbf{H}^n$ by $|x| = |(z,t)| = (|z|^4 + t^2)^{1/4}$ and dilations by

$$x = (z,t) \mapsto \delta x = (\delta z, \delta^2 t), \qquad \delta > 0.$$

The dilations preserve the group law: $\delta(xy) = (\delta x)(\delta y)$. With respect to these dilations the norm is homogeneous of degree 1, i.e. $|\delta x| = \delta|x|$. The vector fields $Z_j = \partial/\partial z^j + i\bar{z}^j\partial/\partial t$, $j = 1,\cdots,n$, are invariant with respect to group multiplication on the left and homogeneous of degree -1 with respect to the dilations. Then $T_{1,0} = \operatorname{span}\{Z_1,\cdots,Z_n\}$ gives a left-invariant CR structure on $\mathbf{H}^n$. The real 1-form

$$\theta_0 = dt + \sum_{j=1}^n \left(iz^j d\bar{z}^j - i\bar{z}^j dz^j\right)$$

is left invariant and homogeneous of degree 2. Since θ_0 annihilates $T_{1,0}$ we may take it to be the contact form for the CR structure. The Levi form is then given by

$$L_{\theta_0}\left(Z_j, \bar{Z}_k\right) = \langle -2id\theta_0, Z_j \wedge \bar{Z}_k\rangle = 2\delta_{jk}.$$

Also, for $u \in C^1(\mathbf{H}^n)$,

$$du = \left(\frac{\partial u}{\partial t}\right)\theta_0 + \sum_{j=1}^{n}\left(\left(Z_j u\right) dz^j + \left(\bar{Z}_j u\right) d\bar{z}^j\right).$$

Therefore,

$$|du|^2_{\theta_0} = \sum_{j=1}^{n} |Z_j u|^2$$

if u is real-valued.

If we write

$$\mathscr{L}_0 = -\tfrac{1}{2}\sum_{j=1}^{n}\left(Z_j\bar{Z}_j + \bar{Z}_j Z_j\right),$$

the operator Δ_b associated to the contact form θ_0 is $\mathscr{L}_0$.

The scalar curvature of $\mathbf{H}^n$ with pseudohermitian structure θ_0 is identically zero. Hence the extremal problem (3.3) in $\mathbf{H}^n$ is

$$\text{(4.1)} \quad \lambda(\mathbf{H}^n) = \inf\left\{\int_{\mathbf{H}^n}\left(b_n \sum_{j=1}^{n} |Z_j u|^2\right)\theta_0 \wedge d\theta_0^n : \int_{\mathbf{H}^n} |u|^p \theta_0 \wedge d\theta_0^n = 1\right\},$$

with $p = b_n = 2 + 2/n$. Note that

$$\begin{aligned}\theta_0 \wedge d\theta_0^n &= n!(2i)^n dt \wedge dz^1 \wedge d\bar{z}^1 \wedge \cdots \wedge dz^n \wedge d\bar{z}^n \\ &= n! 2^{2n} dx, \qquad x = (z, t).\end{aligned}$$

The Cayley transform is a biholomorphism between the unit ball in $\mathbf{C}^{n+1}$ and the Siegel upper half space $\mathscr{D} = \{(z, w) \in \mathbf{C}^n \times \mathbf{C} : \operatorname{Im} w > |z|^2\}$, given by

$$\text{(4.2)} \qquad w = i\left(\frac{1 - \zeta^{n+1}}{1 + \zeta^{n+1}}\right), \qquad z^k = \frac{\zeta^k}{1 + \zeta^{n+1}}, \quad k = 1, \cdots, n,$$

where $\zeta \in \mathbf{C}^{n+1}$, $|\zeta| < 1$. When restricted to the boundary, this transformation gives a CR equivalence between S^{2n+1} minus a point and $\partial\mathscr{D}$. The Heisenberg group is identified with $\partial\mathscr{D}$ by $(z, t) \leftrightarrow (z, t + i|z|^2) = (z, w)$. Denote by F: $S^{2n+1} \to \mathbf{H}^n$ the mapping given by (4.2) followed by this correspondence $\partial\mathscr{D} = \mathbf{H}^n$. Write

$$\theta_1 = i(\bar{\partial} - \partial)|\zeta|^2 = i\sum_{j=1}^{n+1}\left(\zeta^j d\bar{\zeta}^j - \bar{\zeta}^j d\zeta^j\right),$$

the standard contact form for the sphere. Then

$$\theta_1 = F^*\left(\frac{4}{|i+w|^2}\theta_0\right).$$

The conformal factor r of (3.1) is given by $r = 2^n|i+w|^{-n}$, and for $v(\zeta) = |1+\zeta^{n+1}|^{-n}u \circ F(\zeta)$,

$$\int_{S^{2n+1}} \left(b_n|dv|^2_{\theta_1} + R_n v^2\right)\theta_1 \wedge d\theta_1^n = \int_{\mathbf{H}^n} b_n \sum_{j=1}^n |Z_j u|^2 \theta_0 \wedge d\theta_0^n,$$

$$\int_{S^{2n+1}} v^p \theta_1 \wedge d\theta_1^n = \int_{\mathbf{H}^n} u^p \theta_0 \wedge \theta_0^n,$$

where u is a nonnegative function on $\mathbf{H}^n$ and $R_n = n(n+1)/2$ is the scalar curvature associated to θ_1. Thus the extremal problems (4.1) for $\mathbf{H}^n$ and (3.3) for S^{2n+1} are the same. In particular, $\lambda(\mathbf{H}^n) = \lambda(S^{2n+1})$.

Folland and Stein constructed normal coordinates which show how closely the Heisenberg group approximates a general strictly pseudoconvex pseudohermitian structure. If $(W_1, \cdots, W_n)$ is a frame for $T_{1,0}$ over some open set $V \subset N$ which is orthonormal with respect to the given pseudohermitian structure on N, we will call $(W_1, \cdots, W_n)$ a *pseudohermitian frame*. The unique real vector field T defined by $\theta(T) = 1$, and $d\theta(T, X) = 0$ for all X, is transverse to G, and $(W_1, \cdots, W_n, \overline{W}_1, \cdots, \overline{W}_n, T)$ forms a local frame for $\mathbf{C}TN$.

Theorem 4.3 ([9], 14.1, 14.9, 14.10, 16.1). *Let N be a strictly pseudoconvex pseudohermitian manifold of dimension $2n+1$ with contact form θ, and let $V \subset N$ be an open set on which there is given a pseudohermitian frame $(W_1, \cdots, W_n)$. There is a neighborhood of the diagonal $\Omega \subset V \times V$ and a C^∞ mapping $\Theta : \Omega \to \mathbf{H}^n$ satisfying*:

(a) $\Theta(\xi, \eta) = -\Theta(\eta, \xi) = \Theta(\eta, \xi)^{-1}$. (*In particular*, $\Theta(\xi, \xi) = 0$.)

(b) *Denote* $\Theta_\xi(\eta) = \Theta(\xi, \eta)$. Θ_ξ *is thus a diffeomorphism of a neighborhood Ω_ξ of ξ onto a neighborhood of the origin in $\mathbf{H}^n$. Denote by $y = (z, t) = \Theta(\xi, \eta)$ the coordinates of $\mathbf{H}^n$. Denote by O^k, $k = 1, 2, \cdots$, a C^∞ function f of ξ and y such that for each compact set $K \subset\subset V$ there is a constant C_K, with $|f(\xi, y)| \leqslant C_K|y|^k$ (Heisenberg norm) for $\xi \in K$. Then, writing $\Theta_{\xi*} = (\Theta_\xi^{-1})^*$,*

$$\Theta_{\xi*}\theta = \theta_0 + O^1 dt + \sum_{j=1}^n \left(O^2 dz^j + O^2 d\bar{z}^j\right),$$

$$\Theta_{\xi*}(\theta \wedge d\theta^n) = (1 + O^1)\theta_0 \wedge d\theta_0^n.$$

(c)

$$\Theta_{\xi*}W_j = Z_j + O^1\mathscr{E}(\partial_z) + O^2\mathscr{E}(\partial_t), \qquad \Theta_{\xi*}T = \partial/\partial_t + O^1\mathscr{E}(\partial_z, \partial_t),$$

$$\begin{aligned}\Theta_{\xi*}\Delta_b &= \Theta_{\xi*}(\mathbf{Re}\,\square_b)\\ &= \mathscr{L}_0 + \mathscr{E}(\partial_z) + O^1\mathscr{E}(\partial_t, \partial_z^2) + O^2\mathscr{E}(\partial_z\partial_t) + O^3\mathscr{E}(\partial_t^2),\end{aligned}$$

in which $O^k\mathscr{E}$ indicates an operator involving linear combinations of the indicated derivatives with coefficients in O^k, and we have used ∂_z to denote any of the derivatives $\partial/\partial z^j$, $\partial/\partial \bar{z}^j$. (The uniformity with respect to ξ of bounds on functions in O^k is not stated explicitly in [9], *but follows immediately from the fact that the coefficients are C^∞.)*

In what follows, we will use the term *frame constants* to mean bounds on finitely many derivatives of the coefficients in the $O^k\mathscr{E}$ terms in Theorem 4.3.

The function Θ is an approximate group multiplication in the following sense. In the case $N = \mathbf{H}^n$, $\theta = \theta_0$, we can take $\Theta(\xi, \eta) = \xi^{-1}\eta$ and the terms with coefficients in O^k all vanish. In the general case, these extra terms have a higher homogeneity with respect to the dilations $(z, t) \to (\delta z, \delta^2 t)$. Hence they can be viewed as error terms. More precisely, we can rephrase (b) and (c) as:

Remark 4.4. Let $T^\delta(z, t) = (\delta^{-1}z, \delta^{-2}t)$, $K \subset\subset V$, and let r be fixed. With the notation of Theorem 4.3 and $B_r = \{y \in \mathbf{H}^n : |y| \leqslant r\}$, $T^\delta \circ \Theta_\xi(\Omega_\xi) \supset B_r$ for sufficiently small δ and all $\xi \in K$. Moreover, for $\xi \in K$ and $y \in B_r$,

$$\begin{aligned}
&\left(T^\delta \circ \Theta_\xi\right)_* \theta = \delta^2(1 + \delta O^1)\theta_0,\\
&\left(T^\delta \circ \Theta_\xi\right)_*(\theta \wedge d\theta^n) = \delta^{2n+2}(1 + \delta O^1)\theta_0 \wedge d\theta_0^n,\\
&\left(T^\delta \circ \Theta_\xi\right)_* W_j = \delta^{-1}\left(Z_j + \delta O^1\mathscr{E}(\partial_z) + \delta^2 O^2\mathscr{E}(\partial_t)\right),\\
&\left(T^\delta \circ \Theta_\xi\right)_* \Delta_b = \delta^{-2}\left(\mathscr{L}_0 + \mathscr{E}(\partial_z) + \delta O^1\mathscr{E}\left(\partial_t, \partial_z^2\right)\right.\\
&\qquad\qquad\left. + \delta^2 O^2\mathscr{E}(\partial_z\partial_t) + \delta^3 O^3\mathscr{E}\left(\partial_t^2\right)\right).
\end{aligned}$$

(Here O^k may depend also on δ, but its derivatives are bounded by multiplies of the frame constants, uniformly as $\delta \to 0$. Recall that $T_*^\delta Z_j = \delta^{-1}Z_j$, and $T_*^\delta \theta_0 = \delta^2\theta_0$.)

The simplest illustration of rescaling is the proof of Theorem 3.4(b), which we will now carry out.

Lemma 4.5. *The class of test functions defining $\lambda(\mathbf{H}^n)$ can be restricted further to C^∞ functions with compact support.*

Proof. Let $\Psi \in C_0^\infty(\mathbf{H}^n)$ satisfy $\Psi \geqslant 0$, $\int_{\mathbf{H}^n} \Psi(y)\,dy = 1$. Denote $\Psi_\delta(x) = \delta^{-(2n+2)}\Psi(\delta^{-1}x)$. Consider a test function u satisfying $\int_{\mathbf{H}^n} |u|^p \theta_0 \wedge d\theta_0^n = 1$ and $Z_j u \in L^2(\mathbf{H}^n)$, $j = 1, \cdots, n$. The left-invariance of Z_j implies $Z_j(\Psi_{\delta *} u) = \Psi_{\delta *}(Z_j u)$. It is easy to show that $\Psi_{\delta *} u \in C^\infty(\mathbf{H}^n)$, $\Psi_{\delta *} u \to u$ in $L^p(\mathbf{H}^n)$, and $\Psi_{\delta *} Z_j u \to Z_j u$ in $L^2(\mathbf{H}^n)$ as $\delta \to 0$. Hence we can restrict the class of test functions to $u \in C^\infty(\mathbf{H}^n)$.

To see that u can be taken to have compact support, consider $\phi \in C_0^\infty(\mathbf{H}^n)$ such that $\phi(x) = 1$ for $|x| < 1$, $\phi(x) = 0$ for $|x| > 2$, and $0 \leqslant \phi(x) \leqslant 1$ for all x. Denote $\phi^\delta(x) = \phi(\delta x)$. Notice that $Z_j\phi^\delta$ is supported in the "annulus" $\delta^{-1} \leqslant |x| \leqslant 2\delta^{-1}$, and that there is a constant C such that $|Z_j\phi^\delta| \leqslant C\delta$.

Therefore,

$$\int_{\mathbf{H}^n} \left|Z_j(\phi^\delta u)\right|^2 \theta_0 \wedge d\theta_0^n = \int_{\mathbf{H}^n} |(Z_j\phi^\delta)u + \phi^\delta Z_j u|^2 \theta_0 \wedge d\theta_0^n$$

$$\leqslant \int_{\mathbf{H}^n} \left((1+S^{-1})|Z_j\phi^\delta|^2|u|^2 + (1+S)|\phi^\delta|^2|Z_j u|^2\right) \theta_0 \wedge d\theta_0^n$$

$$\leqslant C^2(1+S^{-1})\int_{\mathbf{H}^n} \delta^2\chi^\delta|u|^2 \theta_0 \wedge d\theta_0^n + (1+S)\int_{\mathbf{H}^n} |Z_j u|^2 \theta_0 \wedge d\theta_0^n$$

for any $S > 0$, with

$$\chi^\delta(x) = \begin{cases} 1 & \delta^{-1} \leqslant |x| \leqslant 2\delta^{-1}, \\ 0 & \text{elsewhere.} \end{cases}$$

Note that $\int_{\mathbf{H}^n} \chi^\delta(x)\,dx = C_n\delta^{-(2n+2)}$. Hence by Hölder's inequality and the relation $(2n+2)(1-2/p) = 2$,

$$\int_{\mathbf{H}^n} \delta^2\chi^\delta|u|^2\,dx \leqslant \left(\int_{\mathbf{H}^n} |u|^p\chi^\delta\,dx\right)^{2/p} \delta^2\left(\int_{\mathbf{H}^n} \chi^\delta(x)\,dx\right)^{1-2/p}$$

$$= C_n^{1-2/p}\left(\int_{\mathbf{H}^n} |u|^p\chi^\delta\,dx\right)^{2/p}.$$

This last integral tends to zero as $\delta \to 0$ since $u \in L^p(\mathbf{H}^n)$. Choosing S and then δ sufficiently small we see that

$$\overline{\lim_{\delta\to 0}} \int_{\mathbf{H}^n} \sum_{j=1}^n |Z_j(\phi^\delta u)|^2 \theta_0 \wedge d\theta_0^n \leqslant \int_{\mathbf{H}^n} \sum_{j=1}^n |Z_j u|^2 \theta_0 \wedge d\theta_0^n.$$

Also, clearly,

$$\lim_{\delta\to 0} \int_{\mathbf{H}^n} |\phi^\delta u|^p \theta_0 \wedge d\theta_0^n = \int_{\mathbf{H}^n} |u|^p \theta_0 \wedge d\theta_0^n.$$

Hence we can also restrict the class of test functions to functions of compact support.

We are now ready to prove that $\lambda(N) \leqslant \lambda(\mathbf{H}^n)$. Choose $u \in C_0^\infty(\mathbf{H}^n)$ such that $B_{\theta_0}(u) = 1$, $A_{\theta_0}(u) < \lambda(\mathbf{H}^n) + \varepsilon$. Denote $u_{(\delta)}(x) = \delta^{-n}u(\delta^{-1}x)$. Choose any point $\xi \in N$ and a Folland-Stein coordinate chart Θ_ξ as in Theorem 4.3. Define $v_{(\delta)}(\eta) = u_{(\delta)}(\Theta_\xi(\eta))$. For δ sufficiently small, the support of $u_{(\delta)}$ is contained in $\Theta_\xi(\Omega_\xi)$. Thus $v_{(\delta)}$ has compact support in Ω_ξ and can be extended by zero outside Ω_ξ to a function in $C^\infty(N)$. Note that $B_{\theta_0}(u_{(\delta)}) = B_{\theta_0}(u) = 1$ and $A_{\theta_0}(u_{(\delta)}) = A_{\theta_0}(u) < \lambda(\mathbf{H}^n) + \varepsilon$. Also

$$\int_{\mathbf{H}^n} |u_{(\delta)}|^2 \theta_0 \wedge d\theta_0^n = \delta^2\int_{\mathbf{H}^n} |u|^2 \theta_0 \wedge d\theta_0^n \to 0 \quad \text{as } \delta \to 0.$$

It now follows from Remark 4.4 that $\lim_{\delta\to 0} B_\theta(v_{(\delta)}) = 1$ and $\lim_{\delta\to 0} A_\theta(v_{(\delta)}) = A_{\theta_0}(u) < \lambda(\mathbf{H}^n) + \varepsilon$. Since ε was an arbitrary positive number, we can conclude $\lambda(N) \leqslant \lambda(S^{2n+1})$, which is Theorem 3.4(b).

5. Folland-Stein spaces and estimates for Δ_b

In this section we will define the function spaces that are best suited to regularity properties of the operator Δ_b. These spaces were introduced by Folland and Stein [9], [22] and Propositions 5.1, 5.5, 5.7, and 5.9 are due to them.

We begin by proving the analogue of the classical Sobolev lemma.

Proposition 5.1. *Let $X_j = \operatorname{Re} Z_j$ and $X_{j+n} = \operatorname{Im} Z_j$, $j = 1,\cdots,n$. There exists a constant C_n such that with $p = 2 + 2/n$,*

$$\left(\int_{\mathbf{H}^n} |\phi|^p \theta_0 \wedge d\theta_0^n\right)^{2/p} \leqslant C_n \int_{\mathbf{H}^n} \sum_{j=1}^{2n} |X_j\phi|^2 \theta_0 \wedge d\theta_0^n$$

for every $\phi \in C_0^\infty(\mathbf{H}^n)$.

Proof. The key tool is the fundamental solution

$$F(z,t) = a_n |(z,t)|^{-2n}, \qquad a_n = 2^{2-2n}\pi^{n+1}/\Gamma(n/2)^2,$$

to the operator $\mathscr{L}_0 = -\frac{1}{2}\sum_{j=1}^n (Z_j\bar{Z}_j + \bar{Z}_j Z_j) = -\sum_{j=1}^{2n} X_j^2$. For $\phi \in C_0^\infty(\mathbf{H}^n)$, $(\mathscr{L}_0\phi) * F = \phi$ [9, Proposition 7.1]. Note that by left invariance of X_j, $(X_j h) * F = h * X_j F$ for $h \in C_0^\infty(\mathbf{H}^n)$. Hence,

$$\phi = (\mathscr{L}_0\phi) * F = -\sum_{j=1}^{2n} (X_j\phi) * (X_j F). \tag{5.2}$$

$X_j F$ is homogeneous of degree $-2n-1$. In particular, $|X_j F(z,t)| \leqslant C|(z,t)|^{-2n-1}$.

Lemma 5.3 [9, *Proposition* 8.7]. *If $0 < \alpha < 2n + 2$ and $|H(z,t)| \leqslant C|(z,t)|^{-2n-2+\alpha}$, then the mapping $g \mapsto g * H$ extends to a bounded mapping $L^r(\mathbf{H}^n) \to L^s(\mathbf{H}^n)$, where $s^{-1} = r^{-1} - \alpha/(2n+2)$ and $1 < r < s < \infty$.*

The lemma (applied with $\alpha = 1$, $r = 2$, $s = p$) yields the proposition.

If we consider the inequality of Proposition 5.1 for real-valued functions ϕ, in light of Lemma 4.5, finding the smallest possible constant C_n in Proposition 5.1 is equivalent to finding $\lambda(\mathbf{H}^n)$. In particular, Proposition 5.1 is equivalent to

Proposition 5.4. $\lambda(\mathbf{H}^n) > 0$.

Now let U be a relatively compact open subset of a normal coordinate neighborhood $\Omega_\xi \subset N$ as in Theorem 4.2, with contact form θ and pseudohermitian frame $(W_1,\cdots,W_n)$. Let $X_j = \operatorname{Re} W_j$ and $X_{j+n} = \operatorname{Im} W_j$ for $j = 1,\cdots,n$. Denote $X^\alpha = X_{\alpha_1}\cdot\cdots\cdot X_{\alpha_k}$, where $\alpha = (\alpha_1,\cdots,\alpha_k)$, each α_j an integer $1 \leqslant \alpha_j \leqslant 2n$, and denote $l(\alpha) = k$. Define the norms

$$\|f\|_{S_k^p(U)} = \sup_{l(\alpha)\leqslant k} \|X^\alpha f\|_{L^p(U)},$$

where

$$\|g\|_{L^p(U)} = \left(\int_U |g|^p \theta \wedge d\theta^n\right)^{1/p}.$$

The Folland-Stein space $S_k^p(U)$ is defined as the completion of $C_0^\infty(U)$ with respect to the norm $\|\cdot\|_{S_k^p(U)}$.

Folland and Stein also defined Hölder spaces suited to Δ_b. The function $\rho(\xi, \eta) = |\Theta(\xi, \eta)|$ (Heisenberg norm) is the natural distance function on U. For $0 < \beta < 1$ define

$$\Gamma_\beta(U) = \left\{f \in C^0(\bar{U}) : |f(x) - f(y)| \leqslant C\rho(x, y)^\beta\right\}$$

with norm

$$\|f\|_{\Gamma_\beta(U)} = \sup_{x\in U} |f(x)| + \sup_{x, y\in U} \frac{|f(x) - f(y)|}{\rho(x, y)^\beta}.$$

For any integer $k \geqslant 1$ and $k < \beta < k + 1$ define

$$\Gamma_\beta(U) = \left\{f \in C^0(\bar{U}) : X^\alpha f \in \Gamma_{\beta-k}(U) \text{ for } l(\alpha) \leqslant k\right\}$$

with norm

$$\|f\|_{\Gamma_\beta(U)} = \sup_{x\in U} |f(x)| + \sup_{\substack{x,y\in U\\ l(\alpha)\leqslant k}} \frac{|X^\alpha f(x) - X^\alpha f(y)|}{\rho(x, y)^{\beta-k}}.$$

(The definition of Γ_β for integer values of β involves second differences (cf. [9], [19]). We will not need to use the integer case.) Notice that the norms above depend on the choice of pseudohermitian frame.

Now for a compact strictly pseudoconvex pseudohermitian manifold N, choose a finite open covering $U_1, \cdots, U_m$ for which each U_j has the properties of U above. Choose a C^∞ partition of unity ϕ_i subordinate to this covering, and define

$$S_k^p(N) = \left\{f \in L^1(N) : \phi_j f \in S_k^p(U_j) \text{ for all } j\right\},$$

$$\Gamma_\beta(N) = \left\{f \in C^0(N) : \phi_j f \in \Gamma_\beta(U_j) \text{ for all } j\right\}.$$

Proposition 5.5. *With the notations above, $S_k^r(N) \subset L^s(N)$ for $1/s = 1/r - k/(2n + 2)$ and $1 < r < s < \infty$.*

Proof. According to a fundamental theorem of Folland and Stein [9, Theorem 15.5] extending (5.2), there exist operators A_j, $j = 0, \cdots, 2n$, given by

$$A_j f(x) = \int_N K_j(x, y) f(y)\, dV(y)$$

with

$$|K_j(x, y)| \leqslant \begin{cases} C\rho(x, y)^{-2n-1}, & (x, y) \in \Omega, \\ C & \text{elsewhere on } N \times N, \end{cases}$$

such that $f = \sum_{j=1}^{2n} A_j X_j f + A_0 f$ for every $f \in S_1^r(N)$. Since f and $X_j f$ belong to $L^r(N)$, we conclude from the analogue of Lemma 5.3 on N that $f \in L^s(N)$. The case $k > 1$ follows easily by induction.

Proposition 5.6. *If N is as above, $1 < r < s < \infty$, and $1/s > 1/r - 1/(2n+2)$, then the unit ball in the space $S_1^r(N)$ is compact in $L^s(N)$.*

The proof of this proposition requires the theory of pseudodifferential operators associated to the subelliptic structure of Δ_b as developed by Nagel and Stein [19] and a calculus [11] that permits one to define Folland-Stein spaces S_k^p for fractional values of k. These ingredients would take us too far afield so the proof will appear elsewhere [13].

Let U be a relatively compact open set in a normal coordinate neighborhood as above. We will fix local coordinates to be those given by $(z, t) = \Theta_\xi$ for a fixed point $\xi \in U$. The standard Hölder space $\Lambda_\beta(U)$ is defined for $0 < \beta < 1$ by

$$\Lambda_\beta(U) = \{ f \in C^0(\overline{U}) : |f(x) - f(y)| \leqslant C\|x - y\|^\beta \}$$

with norm

$$\|f\|_{\Lambda_\beta(U)} = \sup_{x \in U} |f(x)| + \sup_{x, y \in U} \frac{|f(x) - f(y)|}{\|x - y\|^\beta}.$$

For $k < \beta < k + 1$, k an integer $\geqslant 1$,

$$\Lambda_\beta(U) = \left\{ f \in C^0(\overline{U}) : (\partial/\partial x)^\alpha f \in \Lambda_{\beta-k}(U) \text{ for } l(\alpha) \leqslant k \right\}$$

with the obvious norm. Then the following fundamental estimates are due to Folland and Stein.

Proposition 5.7. *For each positive noninteger β, each r, $1 < r < \infty$, and each integer $k \geqslant 1$, there exists a constant C such that for every $f \in C_0^\infty(U)$,*

(a) $\|f\|_{\Gamma_\beta(U)} \leqslant C\|f\|_{S_k^r(U)}$, *where* $1/r = (k - \beta)/(2n+2)$,

(b) $\|f\|_{\Lambda_{\beta/2}(U)} \leqslant C\|f\|_{\Gamma_\beta(U)}$,

(c) $\|f\|_{S_2^r(U)} \leqslant C(\|\Delta_b f\|_{L^r(U)} + \|f\|_{L^r(U)})$,

(d) $\|f\|_{\Gamma_{\beta+2}(U)} \leqslant C(\|\Delta_b f\|_{\Gamma_\beta(U)} + \|f\|_{\Gamma_\beta(U)})$.

The constant C depends only on the frame constants.

Folland and Stein proved Proposition 5.7 with $\Box_b$ in place of Δ_b (see [9, Theorems 21.1, 20.1, 16.6, and 15.20]). Their arguments apply verbatim to Δ_b, since it is modelled on the operator $\mathscr{L}_0$, which has a fundamental solution.

Applying a partition of unity, we conclude:

Proposition 5.8. *The estimates in Proposition* 5.7 *hold with U replaced by a compact strictly pseudoconvex CR manifold N.*

The following regularity result follows from these estimates just as in [9, Theorem 16.7].

Proposition 5.9. *If $u, v \in L^1_{\mathrm{loc}}(U)$, and $\Delta_b u = v$ in the distribution sense on U, then for any $\eta \in C^\infty_c(U)$ the following hold.*

(a) *If $v \in L^r(U)$, $n + 1 < r \leqslant \infty$, then $\eta u \in \Gamma_\beta(U)$ where $\beta = 2 - (2n + 2)/r$.*

(b) *If $v \in S^r_k(U)$, $1 < r < \infty$, $k = 0, 1, 2, \cdots$, then $\eta u \in S^r_{k+2}(U)$.*

(c) *If $v \in \Gamma_\beta(U)$, β a noninteger > 0, then $\eta u \in \Gamma_{\beta+2}(U)$.*

We will also need the following regularity result involving critical exponents, which will get an iterative regularity proof started.

Proposition 5.10. *Let U be as in Proposition* 5.7. *Suppose that $f \in L^{n+1}(U)$, $u \in L^p(U)$ (where $p = 2 + 2/n$), $u \geqslant 0$, and $(\Delta_b + f)u = 0$ in the distribution sense on U. Then, for any $\eta \in C^\infty_0(U)$, $\eta u \in L^s(U)$ for every $s < \infty$.*

This proposition is a variant of results of Yamabe [27], Trudinger [24], and Brezis and Kato [3]. A proof is given in the Appendix.

Proposition 5.11. *With the hypotheses of Proposition* 5.10 *and the additional assumption $f \in L^s(U)$ for some $s > n + 1$, we have that u is Hölder continuous in U, and for some $\beta > 0$ and any $K \subset\subset U$,*

$$\|u\|_{\Gamma_\beta(K)} \leqslant C$$

for a constant C depending only on K, $\|f\|_{L^s(U)}$, $\|u\|_{L^p(U)}$, and the frame constants.

Proof. Consider a nested sequence of cutoff functions $\eta_j \in C^\infty_0(U)$ such that $\eta_j = 1$ on K and the support of η_{j+1} is contained in the set on which $\eta_j = 1$. By Hölder's inequality $fu \in L^q(U)$ for $1/q = 1/p + 1/s$. Proposition 5.9(b) implies that $\eta_1 u \in S^q_2(U)$, and thus by Proposition 5.5, $\eta_1 u \in L^{p_1}(U)$ for $1/p_1 = 1/q - 2/(2n + 2) = 1/p - (1/(n + 1) - 1/s)$. Repeating this argument we can conclude that $\eta_k u \in L^{p_k}(U)$ for $1/p_k = 1/p - k(1/(n + 1) - 1/s)$, and every k for which $1/p_k > 0$. Suppose k is the largest possible. Then $p_k > n + 1$, and so Proposition 5.9(a) gives Hölder regularity $\eta_{k+1} u \in \Gamma_\beta(U)$ for $\beta = 2 - (2n + 2)/p_k$. The bound on $\|u\|_{\Gamma_\beta(K)}$ follows from Proposition 5.7.

Proposition 5.12. *With the hypotheses and notation of Proposition* 5.11 *and the additional hypothesis $f \in L^\infty(U)$, we have that*

$$\max_{x \in K} u(x) \leqslant C \min_{x \in K} u(x)$$

for a constant C *depending on the same bounds as in* 5.11 *and in addition* $\|f\|_{L^\infty(U)}$.

The additional hypothesis $f \in L^\infty(U)$ is not necessary (for the classical version see Trudinger [24]). However, we only need the case $f \in L^\infty(U)$, and in this case the proof is practically a verbatim transcription of Moser's proof [18] of the Harnack inequality for uniformly elliptic operators. Instead of considering balls in the ordinary Euclidean sense one has to use balls with respect to the distance function ρ. The appropriate notion of functions of bounded mean oscillation (BMO) relative to this distance and the analogous John-Nirenberg inequality are discussed in [6]. There is only one ingredient of Moser's proof that requires a more detailed discussion, namely the following Poincaré-type inequality.

Proposition 5.13. *Let* U *be as above. There is a constant* C *depending only on the frame constants such that if* $B_r \subset U$ *is a ball of radius* r *with respect to the distance* ρ, *then for every* f *such that* $|df|_\theta \in L^q(B_r)$, $1 < q < \infty$,

$$\int_{B_r} |f - f_{B_r}|^q \theta \wedge d\theta^n \leqslant Cr^q \int_{B_r} |df|_\theta^q \theta \wedge d\theta^n,$$

in which $f_A = (\int_A f\theta \wedge d\theta^n)/(\int_A \theta \wedge d\theta^n)$ *denotes the average value of* f.

This inequality was first proved by A. Greenleaf and D. Jerison (unpublished). A different proof will appear in a forthcoming paper [12].

We note in passing that this implies the following interpolation inequality for the spaces S_1^q.

Proposition 5.14. *If* $u \in L^1(U)$ *and* $|du|_\theta \in L^q(U)$ *with* $1 < q < \infty$, *then* $u \in S_1^q(U)$ *and*

$$\|u\|_{S_1^q(U)} \leqslant C\big(\|\,|du|_\theta\|_{L^q(U)} + \|u\|_{L^1(U)}\big),$$

where C *depends only on the frame constants.*

Proof. From the definition of S_1^q, it suffices to estimate $\|u\|_{L^q(U)}$. We note that

$$\|u\|_{L^q(U)} \leqslant C\big(\|u - u_U\|_{L^q(U)} + \|u_U\|_{L^q(U)}\big),$$

$$\|u_U\|_{L^q(U)} = C\|u_U\|_{L^1(U)} \leqslant C\big(\|u - u_U\|_{L^1(U)} + \|u\|_{L^1(U)}\big),$$

$$\leqslant C\big(\|u - u_U\|_{L^q(U)} + \|u\|_{L^1(U)}\big).$$

Proposition 5.13 completes the proof.

Finally, we are ready to prove regularity results for the Yamabe equation.

Theorem 5.15. *Let* U *be a relatively compact open set in a normal coordinate neighborhood as above. Suppose that* $f, g \in C^\infty(U)$, $u \geqslant 0$ *on* U, $u \in L^r(U)$ *for some* $r > p$, *and* $\Delta_b u + gu = fu^{q-1}$ *in the distribution sense on* U *for some* q, $2 \leqslant q \leqslant p$. *Then* $u \in C^\infty(U)$, $u > 0$, *and if* $K \subset\subset U$, $\|u\|_{C^k(K)}$ *depends only on* K, $\|u\|_{L^r(U)}$, $\|f\|_{C^k(K)}$, $\|g\|_{C^k(K)}$, *and the frame constants, but not on* q.

Proof. Let $h = fu^{q-2} - g \in L^{r/(q-2)}(U)$. By Hölder's inequality, $h \in L^s(U)$, where $s = r/(p-2) > n+1$, and $\|h\|_{L^s(U)}$ depends only on the stated bounds. Then choosing K_1 with $K \subset\subset K_1 \subset\subset U$, it follows from Proposition 5.11 that $u \in \Gamma_\beta(K_1)$ for some $\beta > 0$, and from Proposition 5.12 that u is bounded away from zero by a constant depending on the same bounds. The spaces Γ_β are algebras, and since u is bounded away from zero, $u^\alpha \in \Gamma_\beta(K_1)$ for any real α. Thus, replacing K_1 with a smaller set that we still denote K_1, $h \in \Gamma_\beta(K_1)$ and we conclude from Proposition 5.7(d) that $u \in \Gamma_{\beta+2}(K_1)$. Repeating this argument by induction we see that $u \in C^k(K)$ for any k (see Proposition 5.7(b)).

Corollary 5.16. *Let U, f, g, and u be as above, but assume only that $r = p$ instead of $r > p$. Then we still have $u > 0$ on U and $u \in C^\infty(U)$.*

Proof. Again write $h = g - fu^{q-2}$. With K_1 as above, we conclude successively that $h \in L^{n+1}(K_1)$; $u \in L^s(K_1)$ for all $s < \infty$ (Proposition 5.10); and u is positive and C^∞ (Theorem 5.15).

Finally, we prove the following removable singularities result, which we will use in §7.

Proposition 5.17. *Suppose U is as above, $\xi \in U$, $u \in L^r(U)$ for $r > p/2$, $u \geqslant 0$, $f \in L^{n+1}(U)$, and $(\Delta_b + f)u = 0$ in the distribution sense on $U - \{\xi\}$. Then $(\Delta_b + f)u = 0$ in the distribution sense on U.*

Proof. The hypothesis means that for all $\phi \in C_0^\infty(U - \{\xi\})$,

$$\int_U (u\Delta_b\phi + fu\phi)\,\theta \wedge d\theta^n = 0. \tag{5.18}$$

We need to show this holds for all $\phi \in C_0^\infty(U)$.

Let Θ_ξ be Folland-Stein normal coordinates centered at ξ, with respect to the pseudohermitian frame $(W_1, \cdots, W_n)$. We may assume that $\Theta_\xi(U) = B_R = \{(z,t) : |(z,t)| < R\}$. Choose $\psi \in C_0^\infty(B_R)$ with $0 \leqslant \psi \leqslant 1$ and $\psi \equiv 1$ in $B_{R/2}$, and set $\psi_\delta(z,t) = \psi(\delta^{-1}z, \delta^{-2}t)$. Then $(1 - \psi_\delta)\phi \in C_0^\infty(B_R - \{0\})$ for $\phi \in C_0^\infty(B_R)$ and $0 < \delta < 1$, and so from (5.18)

$$\int_{B_R} (u\Delta_b\phi + fu\phi)\theta \wedge d\theta^n = \int_{B_R} (u\Delta_b(\phi\psi_\delta) + fu\phi\psi_\delta)\,\theta \wedge d\theta^n.$$

We will show the right-hand sides goes to zero as $\delta \to 0$.

By Hölder's inequality, with $r^{-1} + s^{-1} = 1$,

$$\int_{B_\delta} (fu\phi\psi_\delta)\theta \wedge d\theta^n \leqslant C\|f\|_{L^s(B_\delta)}\|\phi u\|_{L^r(B_\delta)},$$

and since $s < n+1$, this expression goes to zero as $\delta \to 0$.

From the definition of Δ_b, we have

$$\Delta_b(\phi\psi_\delta) = \phi\Delta_b\psi_\delta - 2L_\theta^*(d\phi, d\psi_\delta) + \psi_\delta\Delta_b\phi.$$

The term $\int u\psi_\delta \Delta_b \phi$ goes to zero by the same argument as before. Referring to Remark 4.4, and considering the homogeneity of each term in W_j or Δ_b, we see that in B_δ,

$$|W_j \psi_\delta| \leqslant C\delta^{-1}, \qquad |\Delta_b \psi_\delta| \leqslant C\delta^{-2},$$

where C depends only on ψ and the choice of normal coordinates. Then, noting that

$$L_\theta^*(d\phi, d\psi_\delta) = \sum_{j=1}^{n} \left(W_j \phi \overline{W}_j \psi_\delta + W_j \psi_\delta \overline{W}_j \phi \right)$$

and integrating over B_δ,

$$\left| \int_{B_\delta} \left(u\phi \Delta_b \psi_\delta - 2uL_\theta^*(d\phi, d\psi_\delta) \right) \theta \wedge d\theta^n \right|$$

$$\leqslant C\delta^{-2} \|u\|_{L^1(B_\delta)} \leqslant C\delta^{-2} \|u\|_{L^r(B_\delta)} \left(\int_{B_\delta} \theta \wedge d\theta^n \right)^{1/s}$$

by Hölder's inequality. But observe that

$$\theta \wedge d\theta^n = (1 + O(\delta))(\theta_0 \wedge d\theta_0^n) \quad \text{and} \quad \int_{B_\delta} \theta_0 \wedge d\theta_0^n = C\delta^{2n+2}.$$

Thus the last expression above goes to zero provided $(2n + 2)/s > 2$, that is, provided $r > n + 1/n = p/2$.

6. Existence of extremals

In this section we will prove Theorem 3.4(c). As we indicated in §2, we will do so by first considering a perturbed variational problem.

Fix a compact strictly pseudoconvex CR manifold N with contact form θ, and consider for each q, $2 \leqslant q \leqslant p$, the extremal problem

$$\lambda_q = \inf\left\{ A_\theta(\phi) : \phi \in S_1^2(N),\ B_{\theta,q}(\phi) = 1 \right\}, \tag{6.1}$$

in which A_θ is as in (3.3) and

$$B_{\theta,q}(\phi) = \int_N |\phi|^q \theta \wedge d\theta^n.$$

Theorem 6.2. *For $2 \leqslant q < p$, there exists a positive C^∞ solution u_q to the equation*

$$b_n \Delta_b u_q + R u_q = \lambda_q u_q^{q-1} \tag{6.3}$$

satisfying $A_\theta(u_q) = \lambda_q$ and $B_{\theta,q}(u_q) = 1$.

Proof. Consider a minimizing sequence ϕ_j for (6.1), that is, a sequence such that $A_\theta(\phi_j) \to \lambda_q$ and $B_{\theta,q}(\phi_j) = 1$. After replacing ϕ_j by $|\phi_j|$, we can suppose that $\phi_j \geqslant 0$. Since $\{A_\theta(\phi_j)\}$ and $\{B_{\theta,q}(\phi_j)\}$ are bounded, $\{\phi_j\}$ is bounded in S_1^2, and so there is a subsequence converging weakly in S_1^2 to $\phi \in S_1^2(N)$. By the compactness result, Proposition 5.6, the subsequence converges in L^q norm, so $B_{\theta,q}(\phi) = 1$. By Hölder's inequality, $\int R\phi_j^2 \to \int R\phi^2$, and so $A_\theta(\phi) \leqslant \lambda_q$. But since λ_q is an infimum we necessarily have $A_\theta(\phi) = \lambda_q$. Moreover, $\phi \geqslant 0$, and by a standard variational argument ϕ satisfies (6.3) in the distribution sense. Finally $\phi \in L^p(U)$ by Proposition 5.5, and so ϕ is strictly positive and C^∞ by Corollary 5.16.

Next we examine what happens to u_q as $q \to p$. First we consider the behavior of λ_q.

Lemma 6.4. *Suppose θ is chosen so that $\int_N \theta \wedge d\theta^n = 1$. Then*

(a) *If $\lambda_q < 0$ for some q, then $\lambda_q < 0$ for all $q \geqslant 2$ and λ_q is a nondecreasing function of q.*

(b) *If $\lambda_q \geqslant 0$ for some (hence all) $q \geqslant 2$, then λ_q is a nonincreasing function of q, and is continuous from the left.*

Proof. Suppose $\lambda_q < 0$ for some q, and let $q' \geqslant 2$ be arbitrary. Given $\varepsilon > 0$ sufficiently small, choose a C^∞ function ϕ with $B_{\theta,q'}(\phi) = 1$ and $A_\theta(\phi) < \lambda_q + \varepsilon < 0$. With $\phi' = \alpha\phi$ for $\alpha \in \mathbf{R}$, we have $B_{\theta,q}(\phi') = \alpha^{q'}B_{\theta,q'}(\phi)$ and $A_\theta(\phi') = \alpha^2 A_\theta(\phi)$. We set $\alpha = (B_{\theta,q'}(\phi))^{-1/q'}$ so that $B_{\theta,q'}(\phi') = 1$ and $A_\theta(\phi') < 0$. Thus $\lambda_{q'} < 0$. If $q' \leqslant q$, then $\alpha \geqslant 1$ by Hölder's inequality and our normalization of θ. Consequently $A_\theta(\phi') \leqslant \lambda_q + \varepsilon$, which proves that λ_q is nondecreasing.

On the other hand, if $\lambda_q \geqslant 0$ the same argument shows $\lambda_{q'} \leqslant \lambda_q$ if $q' \geqslant q$. Since we can force α to be close to 1 by choosing q' close to q, we also see that λ_q is continuous on the left.

From now on, replacing θ by a constant multiple of itself, we will assume that θ has been normalized so that $\int_N \theta \wedge d\theta^n = 1$.

Theorem 6.5. *If $\lambda(N) < \lambda(S^{2n+1})$, then there exists a sequence q_j tending to p from below such that u_{q_j} converges in $C^k(N)$ for any k to a function $u \in C^\infty(N)$ such that $u > 0$, $b_n\Delta_b u + Ru = \lambda(N)u^{p-1}$, $A_\theta(u) = \lambda(N)$, and $B_{\theta,p}(u) = 1$.*

Proof. *Case* 1. $\lambda(N) < 0$. For $2 \leqslant q < p$ and any $\phi \in S_1^2(N)$, we have

$$\int_N \left(L_\theta^*(du_q, d\phi) + Ru_q\phi\right) = \int_N \lambda_q u_q^{q-1}\phi.$$

Let $\phi = u_q^{q-1}$. Then since $\lambda_q < 0$ by Lemma 6.4,

$$\int_N \frac{q-1}{2} u_q^{q-2}\left|du_q\right|_\theta^2 \leqslant \int_N \left|Ru_q^q\right|.$$

Denote $w_q = u_q^{q/2}$. Then

$$\int_N |dw_q|_\theta^2 \leqslant C\int_N w_q^2 = C\int_N u_q^q = C.$$

Also, by Proposition 5.5, $\int_N w_q^p \leqslant C\int_N(|dw_q|_\theta^2 + w_q^2)$. Hence $\int_N w_q^p \leqslant C$. Now let $q_0 > 2$ and set $r = (q_0/2)p > p$. Then for $q \geqslant q_0$ we have that $\|u_q\|_{L^r(N)}$ is uniformly bounded as $q \to p$. It follows from Theorem 5.15 that $\{u_q\}$ is uniformly bounded in $C^k(N)$, and so a subsequence u_{q_j} converges in C^k for every k. Hence, the limit u satisfies $b_n\Delta_b u + Ru = \lambda u^{p-1}$, $A_\theta(u) = \lambda$, $B_{\theta,p}(u) = 1$, $u > 0$, and $u \in C^\infty(N)$, where $\lambda = \lim_{j\to\infty}\lambda_{q_j}$. By Lemma 6.4, $\lambda \leqslant \lambda(N)$, and so by definition of $\lambda(N)$ we have $\lambda = \lambda(N)$.

Case 2. $\lambda(N) \geqslant 0$. In this case Lemma 6.4 shows that $\lim_{q\to p}\lambda_q = \lambda_p = \lambda(N)$.

Case 2a: For some sequence $q_j \to p$, $\sup_N|du_{q_j}|_\theta$ is uniformly bounded. By Proposition 5.14, $\{u_{q_j}\}$ is uniformly bounded in $S_1^q(N)$ for any q, and hence in $L^r(N)$ for every r. The theorem is concluded in the same way as in Case 1.

Case 2b: $\sup_N|du_q|_\theta \to \infty$ as $q \to p$. We will show that this case never arises.

Choose a point $\xi_q \in N$ such that $\sup_N|du_q|_\theta = |du_q(\xi_q)|_\theta$. Let Θ_{ξ_q} be normal coordinates as in Theorem 4.3. We can assume there is a fixed neighborhood U of the origin in $\mathbf{H}^n$ contained in the image of Θ_{ξ_q} for all q, and for each q we will use Θ_{ξ_q} to identify U with a neighborhood of ξ_q, with coordinates $(z,t) = \Theta_{\xi_q}$.

Now consider the change of coordinates $(\tilde z, \tilde t) = T^\delta(z,t) = (\delta^{-1}z, \delta^{-2}t)$ on $\mathbf{H}^n$, as in Remark 4.4, and set

$$\tilde\theta_0 = d\tilde t + \sum_{j=1}^n \left(i\tilde z^j d\bar{\tilde z}^j - i\bar{\tilde z}^j d\tilde z^j\right) = \delta^{-2}T_*^\delta\theta_0.$$

(Here, as in §4, we write $T_*^\delta = ((T^\delta)^{-1})^*$.) On the set $\delta^{-1}U$ with coordinates $(\tilde z, \tilde t)$ define $h_q(\tilde z, \tilde t) = \delta^{2/(q-2)}u_q(\delta\tilde z, \delta^2\tilde t)$ with $\delta = \delta_q > 0$ chosen so that $|dh_q(0)|_{\tilde\theta_0} = 1$. Observe that $\theta = \theta_0$ at 0, and $|w|_{\delta^{-2}\theta}^2 = \delta^2|w|_\theta^2$ for any 1-form w, and so

$$|dh_q(0)|_{\tilde\theta_0} = |T^{\delta*}dh_q(0)|_{\delta^{-2}\theta_0} = \delta^{1+2/(q-2)}|du_q(\xi_q)|_\theta.$$

In particular, $\delta \to 0$ as $q \to p$, and hence $\delta^{-1}U$ tends to the full space $\mathbf{H}^n$ as $q \to p$.

Now define the contact form $\theta_q = \delta^{-2}T_*^\delta\,\theta$ in coordinates $(\tilde z, \tilde t)$ on the region $\delta^{-1}U$, and set $\mathscr{L}_q = \Delta_b^{(\theta_q)} = \delta^2\Delta_b^{(T_*^\delta\theta)}$. The equation for h_q can then be written

$$b_n\mathscr{L}_q h_q + R_q\delta_q^2 h_q = \lambda_q h_q^{q-1},$$

in which R_q is the scalar curvature of θ expressed in coordinates $(\tilde{z}, \tilde{t})$. Observe that $|R_q| \leqslant \|R\|_{L^\infty(N)}$.

By compactness of N, passing to a subsequence if necessary, we may assume that ξ_q converges to $\xi \in N$, and if we denote by $(W_1^q, \cdots, W_n^q)$ the pseudohermitian frame used to define Θ_{ξ_q} we may assume that $(W_1^q, \cdots, W_n^q)$ converges in C^k for all k to a frame $(W_1, \cdots, W_n)$. Now set $Z_j^q = \delta T_*^\delta W_j^q$, so that $(Z_1^q, \cdots, Z_n^q)$ is a pseudohermitian frame for θ_q. By examining the error terms in the expression for W_j^q in Remark 4.4, it is easy to show that, for any $R > 0$, Z_j^q converges in $C^k(B_R)$ to Z_j for every k. Similarly, θ_q and $\mathscr{L}_q$ converge uniformly in $C^k(B_R)$ to θ_0 and $\mathscr{L}_0$, respectively.

Now fix a radius $R > 0$. Suppose q is sufficiently close to p that $B_{3R} \subset \delta_q^{-1} U$. Let $\eta \in C_0^\infty(B_{2R})$ be equal to 1 on B_R. Then

$$\begin{aligned}(6.6)\quad \mathscr{L}_q(\eta h_q) &= \eta \mathscr{L}_q h_q - 2L_{\theta_q}^*(d\eta, dh_q) + (\mathscr{L}_q \eta) h_q \\ &= \eta\left(-R_q \delta_q^2 h_q + \lambda_q h_q^{q-1}\right) - 2L_{\theta_q}^*(d\eta, dh_q) + (\mathscr{L}_q \eta) h_q.\end{aligned}$$

First, $|dh_q|_{\theta_q}$ is bounded by 1 in B_{2R} because it attains its maximum value of 1 at the origin. Note that

$$(6.7)\quad \int_{|(\tilde{z},\tilde{t})|<R} |h_q(\tilde{z},\tilde{t})|^q \, d\tilde{z}\, d\tilde{t} = \delta_q^{2q/(q-2)-(2n+2)} \int_{|(z,t)|<\delta_q R} |u_q(z,t)|^q \, dz\, dt.$$

For $q < p$, we have $2q/(q-2) > (2n+2)$, and so the coefficient of the right-hand integral is bounded by 1 as $q \to p$. Moreover, the volume element $dz\,dt$ is equal to $C_n(1 + \delta O^1)\theta \wedge d\theta^n$ on $B_{2\delta R}$ by Remark 4.4. Therefore, $h_q \in L^q(B_{2R}, d\tilde{z}\, d\tilde{t})$ with uniform bounds on the norm. In particular, $h_q \in L^1(B_{2R}, d\tilde{z}\, d\tilde{t})$ uniformly as $q \to p$. Combined with the uniform bound on $|dh_q|_{\theta_q}$, this gives $h_q \in S_1^r(B_{2R}, \theta_q)$ for every $r < \infty$ with uniform bounds on the norm, by Proposition 5.14. Thus by Proposition 5.5, ηh^q is uniformly bounded in $L^r(B_{2R})$ for every r, and by Theorem 5.15, uniformly bounded in $C^k(B_R)$ for every k.

Now we can take a subsequence $q_j \to p$ for which h_{q_j} converges, say, in $C^1(B_R)$. Define a function u on all of $\mathbf{H}^n$ by first choosing a subsequence h_{q_j} converging in $C^1(B_1)$, and then a subsequence converging in $C^1(B_2)$, etc. Notice that $u \geqslant 0$, $u \in C^1(\mathbf{H}^n)$, and $u \not\equiv 0$ because $|du(0)|_{\theta_0} = 1$. For $\phi \in C_0^\infty(\mathbf{H}^n)$ we have, since θ_{q_j} converges to θ_0,

$$(6.8)\quad \int_{\mathbf{H}^n} \left(b_n L_{\theta_0}^*(du, d\phi) - \lambda(N) u^{p-1}\phi\right) \theta_0 \wedge d\theta_0^n = 0.$$

Denote $\|u\|_p^p = \int_{\mathbf{H}^n} u^p \theta_0 \wedge d\theta_0^n$. We claim first that

$$(6.9)\quad \|u\|_p \leqslant 1.$$

Since $\theta_q \wedge d\theta_q^n$ approaches $\theta_0 \wedge d\theta_0^n$ uniformly on compact sets, the constraint $\int_N u_q^q \theta \wedge d\theta^n = 1$ and equation (6.7) imply that $\int_{B_R} u^p \theta_0 \wedge d\theta_0^n \leqslant 1$. Because R is arbitrary, (6.9) is proved.

Next we verify

$$\int_{\mathbf{H}^n} |du|^2_{\theta_0} \theta_0 \wedge d\theta_0^n \leqslant C < \infty. \tag{6.10}$$

In fact,

$$\begin{aligned}\int_{B_R} |du|^2_{\theta_0} \theta_0 \wedge d\theta_0^n &= \lim_{j\to\infty} \int_{B_R} |dh_{q_j}|^2_{\theta_{q_j}} \theta_{q_j} \wedge d\theta^n_{q_j} \\ &= \lim_{j\to\infty} \int_{B_{\delta R}} \delta^{2q/(q-2)} |du_{q_j}|^2_\theta \delta^{-2}\theta \wedge (\delta^{-2} d\theta)^n \\ &\leqslant \overline{\lim_{j\to\infty}} \int_N |du_{q_j}|^2_\theta \theta \wedge d\theta^n,\end{aligned}$$

which is bounded. (Here $\delta = \delta_{q_j}$, and we use once again $2q/(q-2) > 2n+2$.)

We can now conclude the proof. Because of the estimates (6.9) and (6.10) we can take a sequence $\phi_j \in C_0^\infty(\mathbf{R}^n)$ approximating u in the norms associated to (6.9) and (6.10). Hence we conclude from (6.8) that

$$b_n \int_{\mathbf{H}^n} |du|^2_{\theta_0} \theta_0 \wedge d\theta_0^n = \lambda(N) \|u\|_p^p.$$

The function $\tilde{u} = u/\|u\|_p$ satisfies the constraint $\|\tilde{u}\|_p = 1$, but using (6.9) and the fact that $p \geqslant 2$ we find that

$$b_n \int_{\mathbf{H}^n} |d\tilde{u}|^2_{\theta_0} \theta_0 \wedge d\theta_0^n = \lambda(N) \|u\|_p^p / \|u\|_p^2 \leqslant \lambda(N) < \lambda(\mathbf{H}^n).$$

This contradicts the definition of $\lambda(\mathbf{H}^n)$. Thus Case 2(b) is impossible, and the proof is concluded.

7. Uniqueness

It would be interesting to know under what circumstances a contact form with constant scalar curvature is unique. As is the case with the Riemannian Yamabe problem, the answer depends on the sign of $\lambda(N)$.

Theorem 7.1. *If $\lambda(N) \leqslant 0$, then any two choices of θ with constant scalar curvature are constant multiples of each other.*

Proof. We note first that the *sign* of any constant scalar curvature is a CR invariant of N. Suppose θ and $\tilde{\theta} = u^{p-2}\theta$ both have constant scalar curvature R and $\tilde{R}$, respectively. Then

$$b_n \Delta_b u + Ru = \tilde{R} u^{p-1}.$$

Integrating this over N and noting that $\int_N \Delta_b u = \int_N L_\theta^*(d1, du) = 0$, we conclude that either $R = \tilde{R} = 0$ or $R/\tilde{R} = \int u^{p-1}/\int u > 0$.

Now suppose $\lambda(N) < 0$. By Theorem 6.5, there exists θ with scalar curvature $R \equiv \lambda(N)$. Suppose $\tilde{\theta} = u^{p-2}\theta$ has constant scalar curvature $\tilde{R}$. The preceding observation shows that $\tilde{R} < 0$, so after multiplying u by a constant we may assume $R = \tilde{R}$. Then u satisfies

$$b_n \Delta_b u + Ru = Ru^{p-1}.$$

It suffices to show $u \equiv 1$.

Since Δ_b is degenerate elliptic, it satisfies a weak maximum principle. At a point $x \in N$ where u is maximum, $\Delta_b u(x) \geqslant 0$, and so $u^{p-1}(x) - u(x) \leqslant 0$, which implies $u \leqslant 1$. Similarly, at a point y where u is minimum, we conclude $u(y) \geqslant 1$. Thus $u \equiv 1$.

Now suppose $\lambda(N) = 0$. By Theorem 6.5, there exists θ with scalar curvature $R \equiv 0$, and by the remarks at the beginning of the proof any other choice $\tilde{\theta} = u^{p-2}\theta$ with constant scalar curvature has $\tilde{R} = 0$. Thus u satisfies $b_n \Delta_b u = 0$, which implies $\int_N |du|_\theta^2 = 0$. Therefore, du is a multiple of θ, say $du = f\theta$ for some $f \in C^\infty(N)$. Differentiating this, we see that $0 = df \wedge \theta + f\, d\theta$. Restricting to G, $f d\theta = 0$ which implies $f \equiv 0$. Thus u is constant.

On the other hand, if $\lambda(N) > 0$, the solution to the Yamabe problem may not be unique. In particular, on the sphere S^{2n+1}, there are many obvious solutions: if we start with the standard contact form θ_1 (cf. §4), and subject S^{2n+1} to a CR automorphism $\Phi: S^{2n+1} \to S^{2n+1}$, then $\Phi^*\theta_1$ will also have constant scalar curvature. In general, $\Phi^*\theta_1 \neq \theta_1$.

It is important to know whether these solutions are extremal for problem (3.3) on S^{2n+1}. We note first that the extremals exist.

Theorem 7.2. *There exists a positive C^∞ contact form $\theta = u^{p-2}\theta_1$ on S^{2n+1} for which the infimum $\lambda(S^{2n+1})$ in* (3.3) *is attained.*

Proof. For $2 \leqslant q < p$, let u_q be the solution to $b_n \Delta_b u_q + Ru_q = \lambda_q u_q^{q-1}$ given by Theorem 6.2. If $|du_q|_\theta^2$ is uniformly bounded as $q \to p$, then u_q converges to a solution u_0 of $b_n \Delta_b u_0 + Ru_0 = \lambda(S^{2n+1}) u_0^{p-1}$ as in the proof of Theorem 6.5, Case 2(a). On the other hand, if $|du_q|_\theta^2$ is unbounded, then as in Theorem 6.5, Case 2(b), we can construct a function $\tilde{u}$ on $\mathbf{H}^n$ satisfying

$$\|\tilde{u}\|_p = 1 \quad \text{and} \quad b_n \int_{\mathbf{H}^n} |d\tilde{u}|_{\theta_0}^2 \theta_0 \wedge d\theta_0^n = \lambda \leqslant \lambda(S^{2n+1}).$$

But since $\lambda(S^{2n+1}) = \lambda(\mathbf{H}^n)$ as defined by (4.1), we must have $\lambda = \lambda(S^{2n+1})$. Now, setting $\theta = F^*(\tilde{u}^{p-2}\theta_0)$ on S^{2n+1} (with $F: S^{2n+1} \to \mathbf{H}^n$ as in §4), we have a contact form that can be written $\theta = u^{p-2}\theta_1$, with $u \in L^p(S^{2n+1})$ satisfying

$$b_n \Delta_b u + Ru = \lambda(S^{2n+1}) u^{p-1} \tag{7.3}$$

on S^{2n+1} minus a point. By Proposition 5.17, this equation holds on all of S^{2n+1}. Finally, by Corollary 5.16, u is positive and C^∞. (This theorem can also be proved using the method of P.-L. Lions [17].)

With this theorem, it is natural to conjecture the following in view of the analogous result of Obata [20] in the Riemannian case.

Conjecture 7.4. *The contact forms* $\theta = \Phi^*\theta_1$, *for* $\Phi \in \operatorname{Aut}(S^{2n+1})$, *are the only ones on the sphere which have constant scalar curvature. Thus*

$$\lambda(S^{2n+1}) = \frac{1}{2}n(n+1)\left(\int_{S^{2n+1}} \theta_1 \wedge d\theta_1^n\right)^{2/p}.$$

To understand this conjecture, we may use the mapping $F: S^{2n+1} \to \mathbf{H}^n$ given by the Cayley transform as in §4 to transfer the problem to the Heisenberg group. If θ is any contact form on S^{2n+1} with constant scalar curvature R, then $\tilde{\theta} = F_*\theta$ is a contact form on $\mathbf{H}^n$ with constant scalar curvature R. For some positive C^∞ function u we can write $\tilde{\theta} = u^{p-2}\theta_0$. Since

$$\int_{\mathbf{H}^n} u^p \theta_0 \wedge d\theta_0^n = \int_{\mathbf{H}^n} \tilde{\theta} \wedge d\tilde{\theta}^n = \int_{S^{2n+1}} \theta \wedge d\theta^n < \infty,$$

we have $u \in L^p(\mathbf{H}^n)$. As in the proof of Theorem 7.1, we may multiply u by a constant to achieve $R = n(n+1)/2$. Then, on $\mathbf{H}^n$, u satisfies

$$4\Delta_b u = n^2 u^{p-1}. \tag{7.5}$$

A routine computation shows that for $\theta = \Phi^*\theta_1$ with Φ a CR automorphism of S^{2n+1}, u has the form

$$u(z,t) = C\left|t + i|z|^2 + z \cdot \bar{\mu} + \lambda\right|^{-n} \tag{7.6}$$

with $C > 0$, $\lambda \in \mathbf{C}$, $\operatorname{Im}\lambda > 0$, and $\mu \in \mathbf{C}^n$. So Conjecture 7.4 is implied by

Conjecture 7.7. *If* $u \in L^p(\mathbf{H}^n)$ *is a positive* C^∞ *solution to* (7.5), *then u is of the form* (7.6).

So far, we have only been able to prove the following weak version of Conjecture 7.7:

Theorem 7.8. *If* $u \in L^p(\mathbf{H}^n)$ *is a positive* C^∞ *solution to* (7.5) *which is radial in the z variable, then u is of the form* (7.6) (*with* $\mu = 0$). (*The other solutions are obtained by left translations on the Heisenberg group.*)

Proof. Introduce the function $w = t + i|z|^2$ on $\mathbf{H}^n$, and write $y = |z|^2$, $x = t$. The hypothesis on u means that $u(z,t) = v(w)$,. where v is a smooth function of the complex variable $w = x + iy$.

We first examine the behavior of v near infinity. Consider the CR inversion $\mathscr{I}: (\mathbf{H}^n - \{0\}) \to (\mathbf{H}^n - \{0\})$ given by $(\tilde{z}, \tilde{t}) = \mathscr{I}(z,t) = (z/w, -t/|w|^2)$. $\mathscr{I}$ satisfes $\mathscr{I}^*\theta_0 = |w|^{-2}\theta_0$. Note that $\mathscr{I}^*u(z,t) = v(-1/w)$. Since

$$\tilde{\theta} = \mathscr{I}^*\left(u^{p-2}\theta_0\right) = |w|^{-2}u\left(z/w, -t/|w|^2\right)^{p-2}\theta_0$$

also has constant scalar curvature $n(n+1)/2$, we have that $\tilde{u}(z,t) = |w|^{-n}u(z/w, -t/|w|^2)$ also satisfies (7.5) on $\mathbf{H}^n - \{0\}$. By Proposition 5.17, (7.5) holds on $\mathbf{H}^n$, so by Corollary 5.16, $\tilde{u}$ is positive and C^∞ near the origin, and so is $\tilde{v}(w) = |w|^{-n}v(-1/w)$. In particular, this means that, as $|w| \to \infty$, $C^{-1}|w|^{-n} \leqslant |v(w)| \leqslant C|w|^{-n}$ for some constant C. Differentiating $\tilde{v}$ with respect to w or $\overline{w}$, we find that

$$|v_{\overline{w}}(w)|,\ |v_w(w)| \leqslant C|w|^{-n-1},$$
$$|v_{ww}(w)|,\ |v_{w\overline{w}}(w)| \leqslant C|w|^{-n-2}.$$

Now consider the function $\phi(w) = (v(w))^{-2/n} = (u(z,t))^{-2/n}$. Observe that $Z_j u = 2i\bar{z}^j \partial v/\partial w$, and so

$$\begin{aligned}\Delta_b u &= -\tfrac{1}{2}\sum_j \left(Z_j\overline{Z}_j + \overline{Z}_j Z_j\right)u = -4yv_{w\overline{w}} - nv_y \\ &= \phi^{-(n+2)/2}\left(2ny\phi_{w\overline{w}} - n(n+2)y\frac{\phi_w\phi_{\overline{w}}}{\phi} + \frac{n^2}{2}\phi_y\right),\end{aligned}$$

and thus (7.5) becomes

$$\phi_{w\overline{w}} = \frac{n+2}{2}\frac{\phi_w\phi_{\overline{w}}}{\phi} - \frac{n}{4}\frac{\phi_y}{y} + \frac{n}{8y}.$$

A rather long computation using this last equation shows that, with $\psi = \phi - y/2$,

$$\begin{aligned}&\operatorname{Re}\partial_{\overline{w}}\left(y^n\phi^{-(n+1)}\left(\psi_{\overline{w}}\phi_{ww} + \psi_w\left(\phi_{w\overline{w}} - \frac{\phi_w\phi_{\overline{w}}}{\phi}\right)\right)\right) \\ &\qquad = y^n\phi^{-(n+1)}\left(\phi_{ww}\phi_{\overline{w}\overline{w}} + \frac{n+2}{n}\left(\phi_{w\overline{w}} - \frac{\phi_w\phi_{\overline{w}}}{\phi}\right)^2\right).\end{aligned} \tag{7.9}$$

When integrated over the upper half-plane $\{(x,y): y > 0\}$, we claim the left-hand side vanishes. It suffices to show that

$$A(w) = y^n\phi^{-(n+1)}\left(\psi_{\overline{w}}\phi_{ww} + \psi_w\left(\phi_{w\overline{w}} - \frac{\phi_w\phi_{\overline{w}}}{\phi}\right)\right)$$

goes to zero sufficiently rapidly as $|w| \to \infty$. But our estimates on the decay of v and its derivatives imply

$$C^{-1}|w|^2 \leqslant |\phi| \leqslant C|w|^2, \quad |\phi_w|, |\psi_w| \leqslant C|w|, \quad |\phi_{ww}|, |\phi_{w\overline{w}}| \leqslant C.$$

Thus $|A(w)| \leqslant C|w|^{-(n+1)}$, and so

$$\int_{y>0} \partial_{\overline{w}}A\,dx\,dy = \lim_{R\to\infty}\int_{\substack{y>0\\|w|<R}} \partial_{\overline{w}}A\,dx\,dy$$

and since $|\int_{|w|=R,\,y>0} A| \leqslant \mathrm{C}R^{-n}$, the last term goes to zero as $R \to \infty$.

Since the integrand on the right-hand side of (7.9) is the positive function $y^n\phi^{-(n+1)}$ multiplied by a sum of squares, we conclude that

$$\phi_{ww} \equiv \phi_{w\bar{w}} - \frac{\phi_w \phi_{\bar{w}}}{\phi} \equiv 0.$$

The second equality implies that $\log \phi$ is harmonic, so $\phi = |f|^2$ for some function f which is holomorphic in w. But then $\phi_{ww} = 0$ implies f is a linear polynomial in w, so ϕ is of the form $\phi(w) = C|w + \lambda|^2$ for $C \in \mathbf{R}$, $\lambda \in \mathbf{C}$. This implies immediately that u is of the form (7.6) (with $\mu = 0$).

Appendix

In this appendix we will prove Proposition 5.10.

Lemma A.1. *Let U_1 be an open set such that $U_1 \subset\subset U$. With the hypotheses of Proposition* 5.10, $u \in S_1^2(U_1)$.

Proof. For $f \in C_0^\infty(U)$, define

$$Pf(\xi) = \int_U f(\eta) F(\Theta(\eta, \xi))\, \theta(\eta) \wedge d\theta(\eta)^n / n! 2^{2n},$$

where F is defined in Proposition 5.1 and Θ in Theorem 4.3. The arguments of [9, Proposition 16.5], show that P is a parametrix for Δ_b in the sense that for $h \in C_0^\infty(U)$, $P(\Delta_b h) = h + Rh$ in U, where R is smoothing of order 1 and P is smoothing of order 2 in the sense of the Folland-Stein spaces $S_k^p(U)$. In particular, if $W_1, \cdots, W_n$ is a pseudohermitian frame for U and $1 < q < r < \infty$, $1/r = 1/q - 1/(2n+2)$, then

(A.2) $\qquad W_j P$ is bounded: $L^q(U) \to L^r(U)$, $j = 1, \cdots, n$;

(A.3) $\qquad W_j R$ is bounded: $L^q(U) \to L^q(U)$, $\qquad j = 1, \cdots, n$.

Let $\psi \in C_0^\infty(U)$ be a real-valued function such that $\psi \equiv 1$ in a neighborhood of $\overline{U}_1$. Then we have in the distribution sense

$$\Delta_b(\psi u) = (\Delta_b \psi) u - f\psi u - 2L_\theta^*(d\psi, du).$$

Therefore, applying a routine limiting argument and the properties of P, we have

$$\psi u + R(\psi u) = P\big((\Delta_b \psi) u - f\psi u - 2L_\theta^*(d\psi, du)\big)$$

in the distribution sense on U_1. Consequently, on U_1,

$$W_j(\psi u) = -W_j R(\psi u) + W_j P\big((\Delta_b \psi) u - f\psi u - 2L_\theta^*(d\psi, du)\big).$$

Because ψ is constant in a neighborhood of $\overline{U}_1$, $d\psi$ and $\Delta_b\psi$ vanish there. The kernel representing the operator P is C^∞ away from the diagonal $\eta = \xi$ in U, and thus

$$W_jP\big((\Delta_b\psi)u - 2L_\theta^*(d\psi, du)\big) \in C^\infty(\overline{U}_1).$$

By (A.3), $W_jR(\psi u) \in L^p(U) \subset L^2(U)$ (with $p = 2 + 2/n$ as usual). Finally, $f \in L^{n+1}(U)$, $u \in L^p(U)$, and Hölder's inequality imply $f\psi u \in L^{p'}(U)$, where $1/p' = 1/p + 1/(n+1)$. Consequently, using (A.2) with $q = p'$, $W_jP(f\psi u) \in L^2(U)$. This yields $W_j(\psi u) \in L^2(U)$, or $W_ju \in L^2(U_1)$, so Lemma A.1 is proved.

It follows from Lemma A.1 that we can use functions in $S_1^2(U_1)$ as test functions: Let $\phi_j \in C_0^\infty(U_1)$ tend to $\phi \in S_1^2(U_1)$ in the S_1^2 norm. Then

$$\int_U u\Delta_b\phi_j = \int_{U_1} L_\theta^*(du, d\phi_j) \to \int_{U_1} L_\theta^*(du, d\phi).$$

Also, $\phi_j \to \phi$ in $L^p(U_1)$, so by Hölder's inequality

$$\int_{U_1} fu\phi_j \to \int_{U_1} fu\phi.$$

Since $\int_{U_1}(u\Delta_b\phi_j + fu\phi_j) = 0$ we conclude that

$$\text{(A.4)} \qquad \int L_\theta^*(du, d\phi) + fu\phi = 0 \quad \text{for every } \phi \in S_1^2(U_1).$$

(This and all subsequent integrations are with respect to $\theta \wedge d\theta^n$.)

Now choose $\beta > 1$ and $N > 0$. Define

$$G(t) = \begin{cases} t^\beta & \text{for } 0 \leqslant t \leqslant N, \\ N^{\beta-1}t & \text{for } t \geqslant N, \end{cases} \qquad F(t) = \begin{cases} t^{(\beta+1)/2} & \text{for } 0 \leqslant t \leqslant N, \\ N^{(\beta-1)/2}t & \text{for } t \geqslant N. \end{cases}$$

Notice that for all $t \geqslant 0$ except $t = N$,

$$\text{(A.5)} \qquad F'(t)^2 \leqslant \beta G'(t), \qquad F(t)^2 = tG(t),$$

$$\text{(A.6)} \qquad G(t) \leqslant F(t)F'(t).$$

Let $\psi \in C_0^\infty(U_1)$, $\psi \geqslant 0$. Because ψ has compact support and $G(t)$ is a Lipschitz function uniformly in t, the function $\phi = \psi^2G(u)$ belongs to $S_1^2(U_1)$. Hence by (A.4)

$$\text{(A.7)} \qquad \int \psi^2G'(u)|du|_\theta^2 + 2\psi L_\theta^*(du, d\psi)G(u) + fu\psi^2G(u) = 0.$$

From (A.6) we have

$$\begin{aligned}\left|\int \psi L_\theta^*(du, d\psi)G(u)\right| &\leqslant \left(\int \psi^2|du|_\theta^2F'(u)^2\right)^{1/2}\left(\int |d\psi|_\theta^2F(u)^2\right)^{1/2} \\ &\leqslant \frac{1}{4\beta}\int \psi^2|du|_\theta^2F'(u)^2 + \beta\int |d\psi|_\theta^2F(u)^2.\end{aligned}$$

Combining this inequality with (A.5) and (A.7) we find that

$$\frac{1}{2\beta}\int \psi^2 |du|_\theta^2 F'(u)^2 \leqslant 2\beta\int |d\psi|_\theta^2 F(u)^2 + \int |f|\psi^2 F(u)^2.$$

Denote $w = \psi F(u)$. Then $dw = F(u)d\psi + \psi F'(u)\,du$. The Sobolev inequality (Proposition 5.5) implies there is a constant C such that

$$\begin{aligned}\left(\int w^p\right)^{2/p} &\leqslant C\int |dw|_\theta^2 + C\int w^2 \\ &\leqslant C_\beta\int |d\psi|_\theta^2 F(u)^2 + C_\beta\int |f|w^2 + C\int w^2.\end{aligned}$$

If E denotes the set where $\psi \neq 0$, then

$$\int |f|w^2 \leqslant \left(\int_E |f|^{n+1}\right)^{1/(n+1)}\left(\int w^p\right)^{2/p}.$$

For E sufficiently small that $(\int_E |f|^{n+1})^{1/(n+1)} < \frac{1}{2}C_\beta$, we conclude that

$$\left(\int w^p\right)^{2/p} \leqslant 2C_\beta\int |d\psi|_\theta^2 F(u)^2 + 2C\int w^2.$$

Taking the limit as $N \to \infty$, we have

$$\left(\int \psi^p u^{(\beta+1)p/2}\right)^{2/p} \leqslant C'_\beta\int \left(|d\psi|_\theta^2 + \psi^2\right)u^{\beta+1}.$$

Now by choosing a suitable collection of cutoff functions ψ we can show that if $u \in L^{\beta+1}(U_2)$ for some $U_2 \subset\subset U_1$, then $u \in L^{(\beta+1)p/2}(U_3)$ for all $U_3 \subset\subset U_2$. Thus, since $p/2 > 1$, we conclude by induction that $u \in L^s(U_2)$ for any $U_2 \subset\subset U_1$ and any $s < \infty$.

References

[1] T. Aubin, *Équations différentielles non linéaires et problème de Yamabe concernant la courbure scalaire*, J. Math. Pures Appl. **55** (1976) 269–296.

[2] M. Beals, C. Fefferman & R. Grossman, *Strictly pseudoconvex domains in C^n*, Bull. Amer. Math. Soc. (N.S.) **8** (1983) 125–322.

[3] H. Brezis & T. Kato, *Remarks on the Schroedinger operator with singular complex potentials*, J. Math. Pures Appl. **58** (1979) 137–151.

[4] D. Burns, K. Diederich & S. Shnider, *Distinguished curves in pseudoconvex boundaries*, Duke Math. J. **44** (1977) 407–431.

[5] S. S. Chern & R. Hamilton, *On Riemannian metrics adapted to three-dimensional contact manifolds*, preprint.

[6] R. R. Coifman & G. Weiss, *Extensions of Hardy spaces and their use in analysis*, Bull. Amer. Math. Soc. **83** (1977) 569–645.

[7] F. Farris, *An intrinsic construction of Fefferman's CR metric*, Pacific J. Math. **123** (1986) 33–45.

[8] C. Fefferman, *Monge-Ampère equations, the Bergman kernel, and geometry of pseudoconvex domains*, Ann. of Math. **103** (1976) 395–416; *Correction*, **104** (1976) 393–394.

[9] G. B. Folland & E. M. Stein, *Estimates for the $\bar{\partial}_b$-complex and analysis on the Heisenberg group*, Comm. Pure Appl. Math. **27** (1974) 429–522

[10] S. Helgason, *Solvability questions for invariant differential operators*, Fifth Internat. Colloq. on Group-Theoretical Methods in Physics, Montreal, 1976.

[11] D. Jerison, *The Dirichlet problem for the Kohn Laplacian on the Heisenberg group*. I, J. Functional Analysis **43** (1981) 97–142.

[12] ______, *The Poincaré inequality for vector fields satisfying Hörmander's condition*, in preparation.

[13] ______, *Remarks on non-isotropic Sobolev spaces*, in preparation.

[14] D. Jerison & J. M. Lee, *A subelliptic, nonlinear eigenvalue problem and scalar curvature on CR manifolds*, Contemporary Math. **27** (1984) 57–63.

[15] J. J. Kohn, *Boundaries of complex manifolds*, Proc. Conf. on Complex Analysis (Minneapolis, 1964), Springer, Berlin, 1965, 81–94.

[16] J. M. Lee, *The Fefferman metric and pseudohermitian invariants*, Trans. Amer. Math. Soc. **296** (1986) 411–429.

[17] P.-L. Lions, *Applications de la méthode de concentration-compacité à l'existence de fonctions extremales*, C. R. Acad. Sci. Paris Sér. I Math. **296** (1983) 645–648.

[18] J. Moser, *On Harnack's theorem for elliptic differential equations*, Comm. Pure Appl. Math. **14** (1961) 577–591.

[19] A. Nagel & E. M. Stein, *Lectures on pseudo-differential operators*, Princeton University Press, Princeton, NJ, 1979.

[20] M. Obata, *The conjectures on conformal transformations of Riemannian manifolds*, J. Differential Geometry **6** (1971) 247–258.

[21] R. Schoen, *Conformal deformation of a Riemannian metric to constant scalar curvature*, J. Differential Geometry **20** (1984) 479–495.

[22] E. M. Stein, *Singular integrals and estimates for the Cauchy-Riemann equations*, Bull. Amer. Math. Soc. **79** (1973) 440–445.

[23] G. Talenti, *Best constant in Sobolev inequality*, Ann. Mat. Pura Appl. **110** (1976) 353–372.

[24] N. Trudinger, *Remarks concerning the conformal deformation of Riemannian structures on compact manifolds*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **22** (1968) 265–274.

[25] K. Uhlenbeck, *Elliptic theory and calculus of variations on manifolds*, Lectures given at Harvard University, 1983.

[26] S. M. Webster, *Pseudohermitian structures on a real hypersurface*, J. Differential Geometry **13** (1978) 25–41.

[27] H. Yamabe, *On a deformation of Riemannian structures on compact manifolds*, Osaka Math. J. **12** (1960) 21–37.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
HARVARD UNIVERSITY