Revista Matemática Iberoamericana
Vol. 4, n.º 1, 1988

José Luis Rubio de Francia (1949-88). Semblanza de su vida y obra

Antonio Córdoba

Es una amarga experiencia, apenas creíble, encontrarse a uno mismo hablando en tiempo pasado de un entrañable y querido compañero con quien se ha compartido, durante los últimos catorce años, esta profesión tan absorbente y tan fascinante de las Matemáticas.

José Luis nació en el año 1949 en Miedes de Aragón, un pueblecito próximo a Calatayud, aunque su familia se trasladó enseguida a Zaragoza, que es la ciudad de su infancia. Tenía la edad de moda entre el profesorado de nuestro departamento.

En el umbral de los cuarenta, para la generalidad de las personas, la vida ya ha presentado sus asperezas y algunos de los árboles más queridos han caído en el entorno de nuestra existencia. Sin embargo, sus amigos no estábamos preparados para esta pérdida que es como hielo crudo que desdora el otoño y como rayo que raja el tronco de la primavera. En palabras de otro de nuestros poetas:

Temprano levantó la muerte el vuelo,
Temprano madrugó la madrugada.

José Luis pertenece por derecho propio a la ciudad de las ideas y a uno de sus barrios más bellos y exquisitos, aquel en cuyo paisaje predominan las flores de las matemáticas. Cuando un matemático muere joven, sus trabajos postreros suelen resultar los más interesantes. En la obra matemática de José Luis

se observa con nitidez una línea ascendente en la importancia de sus problemas, en la calidad de sus soluciones y en la riqueza de su repertorio. Es por esta razón que nadie puede imaginar el territorio que hubiera explorado en una existencia más dilatada.

Su vida ha sido corta, pero creo que ha sido una vida buena y fructífera. Al evocar mis recuerdos y vivencias de estos años, en los que hemos compartido proyectos científicos y universitarios, aficiones y trabajo, predominan sobre todo los innumerables ratos dedicados a nuestros problemas del Análisis Armónico; las horas agradables dedicadas a estimar integrales singulares especialmente aviesas; a imaginar lo que pueda ser el espacio euclídeo de dimensión fraccionaria, proyecto que nos divertía mucho y cuya realización nos permitiría interpolar las dimensiones en el problema de Bochner-Riesz. O cuando allá en Princeton, por el año 1974, nos convencíamos de que las acotaciones L^p de las integrales singulares resultaban más convenientes consideradas como acotaciones L^2 con respecto a los pesos adecuados. Idea que precisó brillantemente años después con su notable «teorema de extrapolación».

Hasta última hora, las matemáticas fueron para él una fuente placentera de problemas interesantes y, finalmente, creo que un bálsamo y un consuelo; como lo demuestra el trabajo que ha realizado durante sus meses de enfermedad, en el que aborda de nuevo uno de nuestros temas favoritos: la convergencia de las sumas esféricas de las series y las integrales de Fourier.

En cierta manera fue un hombre afortunado. Gozó de unas oportunidades excepcionales y tuvo las cualidades necesarias para sacar partido de ellas. Tenía también un carácter afable y generoso, muy adecuado para la colaboración científica. El resultado fue una productividad extrema durante sus años en la Universidad Autónoma. Era un expositor claro y elegante; un buen profesor. Quienes hayan asistido a sus seminarios o a sus cursos de doctorado y licenciatura, no podrán evitar la tristeza de saber que ya nunca volveremos a verle junto a una pizarra, ni escucharemos de nuevo su interpretación segura y sosegada de la belleza lejana de las Matemáticas.

Conocí a José Luis en la Universidad de Princeton, en el mes de septiembre de 1974. Recién leída su tesis doctoral en la Universidad de Zaragoza, y recién casado, obtuvo una beca que le permitió, en compañía de su esposa Amelia, disfrutar de una estancia de dos años en uno de los centros de mayor actividad matemática del mundo. Este período fue decisivo para su carrera y supuso un cambio importante en los temas y un giro en la calidad de su investigación. En la Universidad de Princeton desempeñaba yo mi primer trabajo como profesor, después de doctorarme en la Universidad de Chicago. Recuerdo aquel tiempo como unos años fascinantes. Había una actividad enorme en diversas áreas y era difícil escoger entre varios seminarios interesantes. Me parece que José Luis siguió con especial ahínco el dirigido por Stein, y se sintió en gran parte fascinado por las exposiciones de Eli acerca de integrales singulares

asociadas a curvas homogéneas, funciones maximales de medias esféricas y teoría de Littlewood-Paley que, entre otros temas, constituyeron el objeto de los cursos de Stein en aquellos años, y que también son temas constantes en los trabajos postreros de José Luis. Stein es un gran expositor. Sus clases tienen a menudo la estructura de una sinfonía romántica. Personalmente he sido testigo de como un público adulto, matemáticamente hablando, ha seguido con embeleso su cálculo detallado de la constante de una fórmula, cuyo valor exacto era, por lo demás, irrelevante en la discusión general del tema. No es extraño que José Luis, cuya formación consistía entonces en Análisis Armónico abstracto, en grupos abelianos localmente compactos, con métodos suaves de topología general, se sintiese fascinado por el «análisis duro» y decidiera endurecer su repertorio.

Durante aquel año hablamos muy a menudo de matemáticas. También seguíamos con interés la evolución política española, junto con los otros miembros del grupo hispano de Princeton. La conversación de las reuniones de fin de semana estaba, con frecuencia, llena de proyectos e intenciones de mejora de nuestra universidad.

Recuerdo una tarde en la que me contó los detalles de los resultados de su tesis doctoral: una delicada extensión de los teoremas de distribución uniforme al caso de grupos abelianos localmente compactos, en los que existe una sucesión de subgrupos encajados. Tesis que publicó más tarde en los Anales de la Sociedad Matemática de Polonia. Conversábamos en mi despacho que estaba situado en la planta undécima de la torre de Fine Hall, y desde donde se divisa perfectamente el bello colorido que presentan, en el otoño, los numerosos árboles del campus de Princeton. Compartimos mis proyectos de entonces para derrotar a los elusivos operadores de la sumación esférica; una de cuyas piezas la habíamos preparado en colaboración, C. Fefferman y yo, durante el anterior verano. Este trabajo, que contiene una desigualdad con pesos para la transformada de Hilbert y que fue publicado en Studia, es uno de los artículos más citados por José Luis en su bibliografía posterior.

En el verano de 1975 volvió a España y en el mes de diciembre obtuvo el primer puesto en la oposición convocada para cubrir dos plazas de Análisis Matemático. Entre la Autónoma y la Complutense escogió a esta última, por razones que creo evidentes en aquellos años. Volvió a Princeton en el mes de enero de 1976, en donde residió el resto del curso académico. 1976-77 fue su año en el Departamento de Teoría de Funciones de la Universidad Complutense. Pronto pidió el traslado a Zaragoza en donde desempeñó la cátedra (agregación) hasta el año 1981, y desarrolló una actividad enorme; tanto en la consolidación de su nueva línea de investigación (fundamentalmente la extensión al caso vectorial de la Teoría de Calderón-Zygmund), como en la dirección de tesis doctorales, contando con cuatro y hasta cinco alumnos al mismo tiempo. Sus esfuerzos por elevar el nivel científico del departamento de

Matemáticas de la Universidad de Zaragoza, deben tener difícil parangón en la historia de esa universidad.

Durante el curso 1977-78 nos volvimos a encontrar en un congreso fundacional del grupo español de Análisis Armónico que tuvo lugar en Segovia. Debo decir que yo gozaba de un permiso de la Universidad de Princeton y que, después de una oposición y un traslado, había ocupado la plaza que dejó vacante José Luis en la Universidad Complutense. Aquella fue una época especialmente estimulante para mí, había logrado demostrar una vieja conjetura de A. Zygmund y tuve la oportunidad de exponerla en el congreso organizado en el verano de 1978 en la localidad de Williamstown. Por primera vez en la historia del Análisis Armónico, un congreso de la categoría del de Williamstown contaba con una nutrida representación española; con ponencias interesantes y una conferencia plenaria. Es esta una realidad que se ha ido consolidando desde entonces y a la que José Luis ha contribuido de manera importante.

La conferencia dada por John Gilbert, en la que expuso la teoría de factorización de Nikishin, fue una revelación del congreso que impresionó especialmente a José Luis quién, de vuelta a Zaragoza, logró entender, simplificar y potenciar las ideas de Nikishin y Maurey y conectarlas con sus trabajos sobre integrales singulares. El resultado lo hemos visto crecer desde entonces, hasta convertirse en un instrumento al que José Luis ha sacado registros exquisitos que han vencido a varios problemas clásicos del Análisis Armónico.

Durante las tres semanas que duró el congreso, tuvimos tiempo de conversar acerca del empeño de crear en España un departamento moderno de Matemáticas; reuniendo a la gente capaz de dirigir proyectos de investigación y con una docencia racionalizada donde, por ejemplo, cambiar de asignatura no provocase traumas de posesión territorial. Nuestra común experiencia en la universidad española nos había sensibilizado al respecto. Para mí, además, era el momento de tomar una decisión importante: la de volver a España o permanecer en los Estados Unidos.

En el año 1980, ya en la Universidad Autónoma de Madrid, comenzamos a llevar a la práctica algunas de estas ideas foquistas. Un famoso decreto de plantillas mínimas nos permitió dotarnos de algunas cátedras. De común acuerdo con otros compañeros del departamento le llamé a Zaragoza para ofrecerle la posibilidad de trasladarse a nuestra universidad y convencerle de que así lo hiciera. José Luis mostró un gran interés, pero por motivos diversos (acababa de adquirir un apartamento en Zaragoza, por ejemplo) no podía asegurarme su decisión.

El lunes 23 de febrero de 1981 fue nuestro orador invitado en el seminario del departamento. Para festejar su visita, mi esposa y yo habíamos organizado una pequeña fiesta en casa con los participantes en el seminario. Cuando José Luis se presentó, algo despistado, a la hora convenida de las siete de la

tarde, estábamos escuchando por la radio las lamentables noticias del asalto al Congreso de los Diputados.

En contra de muchos pronósticos, José Luis estaba en nuestra universidad al comienzo del curso académico 1981-82. A pesar de que la vida académica no es siempre todo lo agradable que pudiera, creo que él se ha encontrado a gusto en nuestro departamento, siendo estos años uno de nuestros valores más sólidos. Pero también hay que decir que este período ha sido fundamental para su obra, y que el matemático en ciernes que ganó la cátedra de la Complutense con solo tres artículos, ha publicado, en las mejores revistas, más de veinte entre 1982 y 1988. Muchos de ellos largos y llenos de análisis duro. Ha colaborado con otros miembros del departamento; especialmente con J. García-Cuerva, con quién ha publicado una monografía que es ya una referencia obligada del Análisis Armónico contemporáneo y ha colaborado también con otros conocidos analistas, tales como R. Coifman y P. Jones, de la Universidad de Yale, y M. Christ de UCLA.

José Luis está generalmente considerado como uno de los mejores matemáticos españoles de esta década; la primera en la que puede hablarse de la existencia de buenos matemáticos españoles, como regla, y no como excepción. Es también un figura del Análisis Armónico. En los últimos años ha sido invitado como uno de los conferenciantes principales en los congresos más importantes del área. En el verano de 1986 tuvo lugar una reunión de la élite del Análisis de Fourier, en ese maravilloso palacio de las matemáticas que los alemanes tienen en la Selva Negra. Allí fui testigo de como José Luis se convirtió en la estrella del congreso, tanto por el precioso resultado que expuso, como por las citas de sus trabajos, hechas en las conferencias de otros participantes. Para sus amigos es un consuelo saber que, al final de su corta vida, alcanzó el reconocimiento que se merecía por parte de la comunidad matemática mundial.

Para un analista armónico el objeto fundamental de estudio lo constituye la transformación de Fourier en el espacio $\mathbb{R}^n$

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} e^{-2\pi i x\xi} f(x)\, dx, \qquad \xi \in \mathbb{R}^n,$$

y su análoga periódica en el caso del toro T^n:

$$\hat{g}(\nu) = \int_{T^n} e^{-2\pi i x\nu} g(x)\, dx, \qquad \nu \in Z^n.$$

La transformación de Fourier tiene un buen comportamiento con respecto a las simetrías del espacio euclídeo subyacente y, entre otras, tiene la propiedad de convertir la operación de derivación en multiplicación y, recíprocamente, la integración en división. Esta propiedad es responsable de la enorme importancia del análisis armónico en la teoría de las ecuaciones diferenciales lineales. La transformada de Fourier resulta también un instrumento adecua-

do para modelar la mecánica cuántica y la teoría de la difracción y desempeña un papel importante en la teoría de la comunicación.

¿En qué sentido son ciertas las fórmulas

$$f(x) = \int_{\mathbb{R}^n} e^{2\pi i x\xi} \hat{f}(\xi)\, d\xi, \qquad g(x) = \sum_{\nu} \hat{g}(\nu) e^{2\pi i \nu x} ?$$

¿De qué manera afectan el tamaño y la suavidad de una función a su transformada de Fourier?

Muchos trabajos de análisis matemático están motivados por estas preguntas. Para contestarlas se han creado teorías y operadores vicarios y auxiliares.

Entre los operadores más importantes que pueden ser analizados con la transformada de Fourier, destacan los *Operadores Integrales Singulares* que aparecen al invertir los operadores diferenciales elípticos:

$$Tf(x) = \lim_{\epsilon \to 0} \int_{|y| \geq \epsilon} \frac{\Omega(y)}{|y|^n} f(x - y)\, dy, \qquad f \in C_0^\infty(\mathbb{R}^n)$$

donde Ω es una función homogénea de grado cero, cuya restricción a la esfera unidad es una función lisa de promedio nulo.

El ejemplo fundamental es la transformada de Hilbert:

$$Hf(x) = \lim_{\epsilon \to 0} \frac{f(x - y)}{y}\, dy, \qquad f \in C_0^\infty(\mathbb{R}).$$

Entre los operadores auxiliares hay que mencionar a las funciones maximales y las funciones cuadrado o de Littlewood-Paley. La función maximal de Hardy-Littlewood

$$Mf(x) = \sup_{x \ni Q} \frac{1}{m(Q)} \int_Q |f(y)|\, dy,$$

donde Q designa a un cubo de $\mathbb{R}^n$, define a un operador sublineal acotado en los espacios de Lebesgue $L^p(\mathbb{R}^n)$, $1 < p \leqslant \infty$, y satisface una desigualdad débil en el espacio $L^1(\mathbb{R}^n)$.

La función cuadrado más genuina fue introducida por Littlewood y Paley: dada la serie de Fourier

$$f \sim \sum \hat{f}(\nu) e^{2\pi i \nu x}$$

se define

$$g(f)(x) = \left[\sum_n \left| \sum_{k \in I_n} \hat{f}(k) e^{2\pi i k x} \right|^2 \right]^{1/2}, \qquad I_n = \{k : 2^n \leqslant |k| < 2^{n+1}\}$$

Las normas L^p, $1 < p < \infty$, de $g(f)$ y f son comparables. Este resultado es, en realidad, una relación de L^p-ortogonalidad entre funciones cuyo espectro

de Fourier está localizado en distintas regiones «diádicas» del espacio de frecuencias, y ha sido extendido de diversas maneras tanto al caso no-periódico como a dimensiones mayores. Sin embargo, el «carácter diádico» de la descomposición del espacio de frecuencias es fundamental para que resulte cierta la equivalencia de normas $\|g(f)\|_p \sim \|f\|_p$.

El paradigma fundamental de la Teoría de Calderón-Zygmund puede expresarse de la siguiente manera: la función maximal controla a los operadores integrales singulares con la ayuda de la función cuadrado.

Una formulación de este principio que resulta apropiada para analizar los operadores de la sumación esférica, viene dada por la desigualdad siguiente:

$$(*) \qquad \int |Tf(x)|^p \omega(x)\,dx \leqslant C_{p,s}(T) \int |f(x)|^p [M\omega^s(x)]^{1/s}\,dx, \qquad 1 < p, s < \infty$$

donde T designa a un operador integral singular y la constante finita $C_{p,s}(T)$ es independiente de las funciones f y ω.

Esta estimación es un caso particular de las llamadas desigualdades con peso para las integrales singulares: para todo p, $1 < p < \infty$, existe una clase de pesos A_p tales que las integrales singulares están acotadas en los espacios $L^p(\omega(x)\,dx)$, $\omega \in A_p$.

Uno de los mejores resultados obtenidos por José Luis Rubio consiste en una demostración constructiva, muy ingeniosa y notablemente sencilla, denominada por varios autores el algoritmo de Rubio de Francia, del teorema previamente probado por P. Jones acerca de la estructura de las mencionadas clases de pesos A_p.

Teorema (P. Jones-J. L. Rubio). *$\omega \in A_p$ si y sólo si ω es comparable a una función de la forma $(M\omega_0)^\delta \cdot (M\omega_1)^{\delta(1-p)}$ donde ω_0, ω_1 son funciones localmente integrables y δ es un número positivo menor que* 1.

Este resultado y su método de demostración dado por José Luis, han originado diversos teoremas llamados de extrapolación y otra perspectiva de la dualidad entre los espacios H^1 y BMO. La desigualdad (*) produce automáticamente estimaciones vectoriales para familias de integrales singulares. José Luis cambió el sentido de la implicación y observó como las estimaciones vectoriales dan lugar a desigualdades con peso. La monografía publicada en North-Holland contiene una detallada exposición de esta teoría.

La desigualdad (*) resulta también especialmente útil cuando han de considerarse en $\mathbb{R}^n$, conjuntamente, varias transformadas de Hilbert unidimensionales asociadas a distintas direcciones del espacio. Esto es precisamente lo que ocurre en el análisis de los operadores de Bochner-Riesz ($\widehat{T_\alpha f}(\xi) = (1 - |\xi|^2)_+^\alpha \hat{f}(\xi)$, $\alpha \geqslant 0$) donde (*) permite reducir la acotación de T_α a la estimación precisa del crecimiento de la norma, respecto al número de direcciones involucradas, de

la denominada función maximal de Kakaya. En este empeño se observó la importancia de tener desigualdades unilaterales $\|g(f)\|_p \leqslant C_p \|f\|_p$ para ciertas funciones cuadrado «no diádicas».

En el año 1980 tuvo lugar un congreso de análisis armónico en la Escuela Normal Superior de Pisa, al que asistió José Luis. En mi conferencia presenté una demostración de la acotación:

$$\int |g(f)(x)|^2 \omega(x)\, dx \leqslant C_s \int |f(x)|^2 [M\omega^s(x)]^{1/s}\, dx, \qquad 1 < s < \infty,$$

válida para cualquier par de funciones f, ω con una constante finita C_s independiente de ellas.

Aquí

$$g(f)(x) = \left[\sum |P_\alpha f(x)|^2\right]^{1/2}, \qquad \widehat{P_\alpha f}(\xi) = \chi_{Q_\alpha}(\xi)\hat{f}(\xi)$$

y los $\{Q_\alpha\}$ constituyen una familia disjunta de cubos de $\mathbb{R}^n$ de igual tamaño y lados paralelos a los ejes coordenados. En particular, la desigualdad anterior implica la siguiente familia de estimaciones:

$$(**) \qquad \|g(f)\|_p \leqslant C_p \|f\|_p, \qquad 2 \leqslant p < \infty.$$

Resulta, pues, natural preguntarse si (**) seguirá siendo válida al sustituir cubos por rectángulos (paralelepípedos) de lados paralelos a los ejes, pero de dimensiones arbitrarias.

Entre los asistentes al seminario se encontraba mi gran amigo y excelente analista Yves Meyer, a quién no logré convencer de la bondad de la conjetura. Como al término del congreso no habíamos logrado encontrar ora la demostración, ora el contraejemplo, apostamos una botella de buen caldo. En el verano de 1983 José Luis obtuvo una preciosa demostración del caso unidimensional. Posteriormente Jean L. Journé, un alumno de Yves, logró demostrarla para dimensión arbitraria. El resultado es un bello teorema.

Teorema (Journé-Rubio de Francia). *Sea $\{R_k\}$ una familia disjunta de «rectángulos» de $\mathbb{R}^n$ de lados paralelos a los ejes coordenados. Sea el operador P_k definido por la fórmula*

$$\widehat{P_k f}(\xi) = \chi_{R_k}(\xi)\hat{f}(\xi)$$

y la función cuadrado

$$g(f)(x) = \left[\sum |P_k f(x)|^2\right]^{1/2}.$$

Entonces, para todo p, $2 \leqslant p < \infty$ existe una constante finita C_p tal que $\|g(f)\|_p \leqslant C_p \|f\|_p$, para toda función $f \in L^p(\mathbb{R}^n)$.

Debo añadir que cuando finalmente recibí mi premio, un excelente chateau, pude decirle a mi amigo Yves que nunca dudé de mi triunfo, por cuanto contaba con un poderoso paladín, José Luis Rubio, y, además, uno de sus mejores hombres se había cambiado a mi bando.

La teoría de las integrales singulares ha tenido diversas generalizaciones. Una de ellas es el operador

$$Hf(x) = \text{v.p.} \int_{-\infty}^{+\infty} f(x - \Gamma(t)) \frac{dt}{t}$$

donde $\Gamma\colon \mathbb{R} \to \mathbb{R}^n$, $n \geqslant 2$, es una curva que pasa por el origen de coordenadas: $\Gamma(0) = 0$. Este objeto, denominado transformada de Hilbert a lo largo de la curva Γ, aparece en el estudio de los operadores diferenciales parabólicos y en el análisis armónico de ciertos espacios de tipo homogéneo. Su teoría, acotaciones L^p, funciones maximales asociadas etcétera, ha sido desarrollada con notable éxito por A. Nagel, E. Stein y S. Wainger, entre otros. En el año 1986 colaboré con José Luis, y con el equipo de Madison, en demostrar que una cierta condición de curvatura era necesaria y suficiente para que el operador H estuviera acotado en todos los espacios L^p, $1 < p < \infty$. En el número de la Revista Matemática Iberoamericana dedicado a Alberto P. Calderón publicamos un trabajo que contiene el resultado siguiente: sea

$$\Gamma = \{(t, \gamma(t))\colon -\infty < t < +\infty\}, \qquad \gamma(0) = \gamma'(0) = 0$$

una curva de clase C^1 en el plano con las propiedades:

(*a*) γ es biconvexa, es decir $|\gamma'(t)|$ es decreciente en $(-\infty, 0)$ y creciente en $(0, +\infty)$.

(*b*) Existe una constante $C > 1$ tal que $|\gamma'(Ct)| \geqslant 2|\gamma'(t)|$.

(*c*) γ es equilibrada, es decir: existe una constante $k > 1$ tal que $|\gamma(k^{-1}t)| \leqslant |\gamma(-t)| \leqslant |\gamma(kt)|$ para todo $t > 0$.

Teorema. *Bajo las hipótesis (a), (b) y (c) acerca de la curva, la integral singular H y las funciones maximales M y H* están acotadas en* $L^p(\mathbb{R}^2)$, $1 < p < \infty$.

Siendo

$$Mf(x) = \sup_{h>0} \frac{1}{2h} \int_{-h}^{+h} |f(x - \Gamma(t))|\, dt,$$

$$H^*f(x) = \sup_{\epsilon>0} \left| \int_{|t| \geq \epsilon} f(x - \Gamma(t)) \frac{dt}{t} \right|.$$

Los detalles de la demostración habían sido elaborados en varias sesiones de trabajo celebradas en mi despacho, pero había que hacer la redacción final

que fuimos dejando de un día para otro. Se acercaba la fecha de su partida para Yale, donde había de permanecer durante todo un semestre, así es que decidimos pasar un domingo juntos, con las familias, en mi casa de Soto del Real, y dedicar unas horas a redactar el artículo. Carecíamos entonces de la condición (c) ya que creíamos que el resultado era cierto sin ella; de manera que transcurrían las horas y las cuentas no cuadraban del todo; nos acercábamos peligrosamente a la hora del almuerzo y la familia daba muestras de ansiedad creciente; sin embargo, el teorema no acababa de salir. Finalmente, cuando la presión era insostenible, decidimos cortar por lo sano y nos inventamos la condición que llamamos de equilibrio, que nos permitió acabar la escritura y gozar de una tarde apacible. A los pocos días descubrimos la hilarante realidad de que (c) es también una propiedad necesaria.

He recogido tan solo una muestra de sus contribuciones al análisis, pero su producción es mucho más amplia y contiene bastantes gamas. Más que un luthier o un director de orquesta, José Luis fue un virtuoso que obtuvo registros nuevos e insospechados de todos sus instrumentos. Algunos de sus teoremas son muy difíciles, de una elaborada orfebrería, otros son comparativamente fáciles, una vez que uno está en posesión de la idea adecuada, pero ninguno de ellos es banal y en todos sus escritos encontramos una presentación elegante y exquisita. Sus amigos echaremos de menos sobre todo su amistad, su carácter afable y positivo, más interesado en crear que en lamentarse, sin complejos, y cuyo mayor defecto era su notoria incapacidad para decir que no.

Si exceptuamos algunos casos aislados del pasado, puede decirse que las Matemáticas existen en España desde mediados de los años setenta. A ello han contribuido fundamentalmente dos factores: la cantidad notable de españoles que han realizado estudios de doctorado en los mejores centros del mundo, y las posibilidades que ha habido estos años para recuperarles y ofrecerles puestos en algunas universidades.

Aunque aún nos queda un largo trecho por recorrer, la región de las matemáticas ya presenta entre nosotros algunos frondosos bosques y caudalosos ríos, incluso contamos con varios picos de montaña. La pérdida de José Luis entristece nuestro paisaje, pero estamos convencidos de que su semilla lo hará reverdecer durante mucho tiempo.

Antonio Córdoba
Marzo, 1988

REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.° 1, 1988

Non-Negative Solutions to Fast Diffusions

Bjorn E. J. Dahlberg and Carlos E. Kenig

Dedicated to the memory of our friend José Luis Rubio de Francia

Introduction

The purpose of this work is to study the class of non-negative continuous weak solutions of the non-linear evolution equation

$$\frac{\partial u}{\partial t} = \Delta\varphi(u), \qquad x \in \mathbb{R}^n, 0 < t < T \leqslant +\infty. \tag{1.1}$$

Here the non-linearity φ is assumed to be continuous, increasing, with $\varphi(0) = 0$. In the case when $\varphi(u)$ is super-linear when u tends to $+\infty$, the corresponding study was carried out in [7] (and in the pure power case in [2], [6] and [4]). Here, we treat the case when $\varphi(u)$ is sub-linear when u tends to $+\infty$ (but not *too* sub-linear, as we shall see below). In doing so, we complement and extend (and make extended use of) some results of M. A. Herrero and M. Pierre [12], who have studied the Cauchy problem for the equation

$$\frac{\partial u}{\partial t} = \Delta u^m, \qquad 0 < m < 1, \text{ in } \mathbb{R}^n \times (0, \infty),$$

with any initial data f in $L^1_{\text{loc}}(\mathbb{R}^n)$. They prove that, given any such f, there exists a unique (in an appropriate class) solution u to the Cauchy problem. Moreover, in the case when $m > (n-2)/n$, they establish a regularizing effect from L^1_{loc} to L^∞_{loc}, for the solution, exploiting the well-known Aronson-Bénilan [1] differential inequality. It is known (see [5]) that for $0 < m \leqslant (n-2)/n$, such a regularizing effect cannot hold.

Note that existence of solutions for arbitrary $L^1_{\text{loc}}(\mathbb{R}^n)$ is in sharp contrast with the super-linear case ([2], [4], [6], [7]) where, even to get existence for finite time, growth conditions at infinity on the initial data are necessary.

The quantitative assumptions on our non-linearity φ are

$$(1.2)\quad \begin{cases} \text{(i)} & a \leqslant \dfrac{u\varphi'(u)}{\varphi(u)} \leqslant \dfrac{1}{a}, \qquad u > 0 \\ \text{(ii)} & \dfrac{n-2}{n} + a \leqslant \dfrac{u\varphi'(u)}{\varphi(u)} \leqslant 1 - a, \qquad u > u_0 \end{cases}$$

Assumption (i) is a global, scale invariant, polynomial growth for φ, while (ii) reflects sub-linearity for u large, but also reflects a power behavior larger than $(n-2)/n$ for u large.

Under (1.2), we prove (Theorem 3.11) that any non-negative continuous weak solution of (1.1) in $\mathbb{R}^n \times (0, T)$, $0 < T \leqslant +\infty$ has an initial trace μ, as $t \downarrow 0$, which is a locally finite Borel measure (in fact, only (1.2)(i) is needed for this). We also establish the L^1_{loc} into L^∞_{loc} regularizing effect (Corollary 3.20). Here, the full force of (1.2) is needed, since, without the left-hand side of (1.2)(i)), this fails. At this point, there is a crucial methodological difference with the work of Herrero and Pierre ([12]), since the Aronson-Bénilan inequality [1] is unavailable to us, given our assumptions on φ. We then prove (Theorem 3.22), that given any locally finite, positive Borel measure μ there exists a non-negative continuous weak solution of (1.1) in $\mathbb{R}^n \times (0, \infty)$, with initial trace μ. Again, the left-hand inequality of (1.2)(ii) is needed here, because of the work of Brézis and Friedman [5], which shows that for $\varphi(u) = u^m$, $0 < m \leqslant (n-2)/n$, there is no solution with initial data the delta mass. Moreover, the right-hand inequality in (1.2)(ii) is also needed, since there are no growth restrictions on the measure μ at ∞. Finally, (Theorem 3.28), we show that non-negative, continuous weak solutions of (1.2) in $\mathbb{R}^n \times (0, T)$, $0 < T \leqslant +\infty$, are uniquely determined by their initial trace. To show this, we adopt techniques in [16], [12], [6], and [7]. This result may be somewhat surprising, since such solutions verify no growth restrictions at infinity. A corollary of these results (Theorem 3.33) is that any non-negative continuous weak solution of (1.1) in $\mathbb{R}^n \times (0, T)$, $0 < T \leqslant +\infty$, with φ verifying (1.2) has a unique extension to a solution in $\mathbb{R}^n \times (0, \infty)$. This again is in sharp contrast with the super-linear case, where blow-up in finite time can occur.

Several interesting questions along the lines of this work remain. For instance, one would like to be able to classify, if one does not assume the left-hand inequality in (1.2)(ii), the class of «admissible» initial traces. Also, it would be of interest to develop, assuming (1.2), the theory for the initial-Dirichlet problem in finite cylinders, along the lines of [8]. Another important question is whether, under (1.2), non-negative continuous weak solutions of (1.1), and weak solutions of (1.1) are the same, as in the super-linear case ([9], [10]). We hope to return to some of these questions at a later time.

All the solutions u considered in this work are non-negative. We will sometimes omit specific mention of this fact.

Acknowledgement. Part of this research was carried out while the second author was a visiting member of the M.S.R.I., Berkeley, California. He gratefully acknowledges their hospitality and support.

2. Preliminary Results

In this section we collect some preliminary results that will be used throughout the paper. Most of the proofs are omitted, since they are essentially contained in [7]. We will work with continuous, strictly increasing φ on $0 \leqslant u < +\infty$, that are positive on $0 < u < +\infty$, with $\varphi(0) = 0$. In this section we only impose the growth condition

$$0 < a \leqslant \frac{u\varphi'(u)}{\varphi(u)} \leqslant \frac{1}{a}, \qquad 0 < u < \infty \tag{2.1}$$

and the normalization

$$\varphi(1) = 1. \tag{2.2}$$

When φ verifies (2.1) and (2.2), we say that $\varphi \in \mathfrak{A}_a$.

We say that u is a continuous non-negative weak solution of

$$\frac{\partial u}{\partial t} = \Delta\varphi(u) \quad \text{in} \quad \Omega \subset \mathbb{R}^{n+1} \tag{2.3}$$

if u is continuous in Ω, $u \geqslant 0$, and (2.3) holds in Ω in the distribution sense.

Set $B = \{x \in \mathbb{R}^n : |x - x_0| < r\}$, $\tau_1 < \tau_2$, and let $Q = B \times (\tau_1, \tau_2)$. Denote by $\partial_p Q$ the parabolic boundary of Q, *i.e.*, $\partial_p Q = \bar{Q} \setminus (B \times \{\tau_2\})$.

Let $\varphi \in \mathfrak{A}_a$, $0 < a < 1$, and let $g \in C(\partial_p Q)$ be a given non-negative function. Consider the boundary value problem

$$\begin{cases} \dfrac{\partial v}{\partial t} = \Delta\varphi(v) & \text{in} \quad Q \\ v = g & \text{in} \quad \partial_p Q \end{cases} \tag{2.4}$$

A function $v(x, t)$ is said to be a weak solution of (2.4) in Q, if $v \in C([\tau_1, \tau_2]; L^1(B)) \cap L^\infty(Q)$, $v \geqslant 0$, and v satisfies the integral identity

$$\int_Q \int \left[\varphi(v)\,\Delta\eta + v\frac{\partial \eta}{\partial t} \right] dx\,dt = \int_{\tau_1}^{\tau_2} \int_{\partial B} \varphi(g) \frac{\partial \eta}{\partial N}\, d\sigma\, dt + \int_B v(x, \tau_2)\eta(x, \tau_2)\, dx - \int_B g(x, \tau_1)\eta(x, \tau_1)\, dx \tag{2.5}$$

for all smooth functions η on $\bar{Q}$, which vanish on $\partial B \times [\tau_1, \tau_2]$. Here $\partial/\partial N$ denotes the exterior normal derivative on ∂B, and σ denotes the surface measure on ∂B.

Lemma 2.6. *Let u be a continuous non-negative weak solution of* (2.3) *in $\Omega \subset \mathbb{R}^n \times \mathbb{R}$. Suppose that $\bar{Q} \subset \Omega$, and let $g = u|_{\partial_p Q}$. Then, u is a weak solution of* (2.4) *in Q.*

The proof given in [2], Theorem 3.1 applies without changes.

We next need some maximum principles and approximation results.

Lemma 2.7. *Let $g_1, g_2 \in C(\partial_p Q)$, and let v_1, v_2 be weak solutions of* (2.4) *with boundary values g_1 and g_2 respectively. If $0 \leqslant g_1 \leqslant g_2$, then $v_1 \leqslant v_2$ in Q.*

The proof given in [7], Lemma 2.3 applies verbatim.

Corollary 2.8. *Let $g \in C(\partial_p Q)$, $g \geqslant 0$, and assume that v is a weak solution of* (2.4), *and that v is continuous in $\bar{Q}$. Choose $G_k \in C^\infty(\mathbb{R}^n \times \mathbb{R})$ such that $g_k = G_k|_{\partial_p Q}$ are strictly positive, $g \leqslant g_{k+1} \leqslant g_k$, $g_k \to g$ uniformly. Let $\varphi_k \in C^\infty(0, \infty)$, $\varphi_k \in \mathfrak{A}_a$, $\varphi_k \to \varphi$ uniformly on compact subsets of $[0, \infty)$. Let v_k solve $\partial v_k/\partial t = \Delta \varphi_k(v_k)$ in Q, $v_k = g_k$ on $\partial_p Q$. Then, each $v_k \in C^\infty(\bar{Q})$, and $v_k \to v$ uniformly on compact subsets of Q.*

The proof is the same as the one of Corollary 2.1 of [7]. The next preliminary result that we need is an extension of M. Pierre's uniqueness theorem [16], to our class of non-linearities $\mathfrak{A}_a$. Again, the proof is given in [7], Lemma 2.12.

Lemma 2.9. *Suppose that $\varphi \in \mathfrak{A}_a$, and that u and v are continuous, non-negative weak solutions of* (2.3) *in $\mathbb{R}^n \times (0, \infty)$, $n \geqslant 3$. Suppose that*

$$\sup_{t>0} \int_{\mathbb{R}^n} [u(x, t) + v(x, t)]\, dx < \infty,$$

and that $u, v \in L^\infty(\mathbb{R}^n \times [\tau_1, \tau_2])$ for each $0 < \tau_1 < \tau_2 < +\infty$. Suppose also that

$$\lim_{t \downarrow 0} \int_{\mathbb{R}^n} [u(x, t) - v(x, t)]\eta(x)\, dx = 0$$

for each $\eta \in C_0^\infty(\mathbb{R}^n)$. Then, $u \equiv v$ in $\mathbb{R}^n \times (0, T)$.

We now need some *a priori* estimates. Let $1/2 < r < \rho < 2$, $S = B_r \times (-r^2, 0]$, $R = B_\rho \times (-\rho^2, 0]$, where $B_r = \{x \in \mathbb{R}^n : |x| < r\}$.

Lemma 2.10. *Let u be a smooth, non-negative solution of the equation $\partial u/\partial t = \Delta\varphi(u)$ in R, where $\varphi \in \mathfrak{A}_a \cap C^\infty[0, \infty)$. Then,*

$$\|u\|_{L^\infty(S)} \leqslant C\left\{1 + \frac{1}{(\rho - r)^N}\int_R \int u^p \, dx\, dt\right\}^\theta, \tag{2.11}$$

where C, p, θ, N are positive constants which depend only on n and a.

The proof of Lemma 2.10 follows from a variant of the Moser ([4]) iteration technique, which is essentially given in [7], Lemma 3.1.

Corollary 2.12. *Let u be a continuous non-negative weak solution of* (2.3), *with φ in $\mathfrak{A}_a$, in R. Then,* (2.11) *holds.*

PROOF. Corollary 2.12 follows from 2.6, 2.8 and 2.10.

3. Nonnegative solutions

In this section we establish the main properties of nonnegative continuous weak solutions u of (2.3), in the strips $S_T = \mathbb{R}^n \times (0, T)$, $0 < T \leqslant +\infty$, under the assumptions (2.1) and (2.2) on the non-linearity φ, together with the additional assumption

$$\frac{n-2}{n} + a \leqslant \frac{u\varphi'(u)}{\varphi(u)} \leqslant a, \qquad u \geqslant u_0 \tag{3.1}$$

where $0 < a < 1$. If we also impose the normalization $u_0 = 1$, we say that $\varphi \in \mathfrak{B}_a$. It is easy to see that (2.1) and (3.1) imply the existence of numbers μ, ν, γ, $0 < \mu < 1$, $0 < \nu < +\infty$, $\gamma > (n-2)/n$ which depend only on a, so that, if $\varphi \in \mathfrak{B}_a$,

$$\begin{aligned} &\text{(i)} && \varphi(u) \leqslant u^\mu, && u \geqslant 1 \\ &\text{(ii)} && \varphi(u) \leqslant u^\nu, && 0 < u < 1 \\ &\text{(iii)} && \varphi(u) \geqslant u^\gamma, && u \geqslant 1 \end{aligned} \tag{3.2}$$

We will also need the following elementary estimate

$$\int_b^a \frac{du}{Au^\beta + Bu^\mu} \approx \begin{cases} \dfrac{1}{B}[a^{1-\mu} - b^{1-\mu}], & Ca > 2 \\ \dfrac{1}{A}[a^{1-\beta} - b^{1-\beta}], & Ca \leqslant 2, \end{cases} \tag{3.3}$$

where $0 < \beta < \mu < 1$, $A > 0$, $B > 0$, $C = (B/A)^{1/(\mu-\beta)}$, and as usual $\simeq$ means that the ratio is bounded above and below. (3.3) remains valid if $\beta = \mu$, provided $A = B$. Let now μ, ν be as in (3.2), and define $\beta = \min\{\mu, \nu\}$. Then, $0 < \beta \leqslant \mu < 1$, and,

$$\varphi(u) \leqslant u^{\beta}, \qquad 0 < u < 1. \tag{3.4}$$

Following Herrero and Pierre ([12]), for any number θ, $0 < \theta < 1$, and $\psi \in \mathcal{S}(\mathbb{R}^n)$, $\psi \geqslant 0$, we set

$$C_\theta(\psi) = \left[\int_{\mathbb{R}^n} |\Delta\psi|^{1/(1-\theta)} \psi^{-\theta/(1-\theta)}\right]^{1-\theta}. \tag{3.5}$$

Clearly $C_\theta(\psi)$ could be infinite, but, if $C_\theta(\psi) < \infty$, the same is true for $C_\theta(\psi_{x_0,R})$, where, for $x_0 \in \mathbb{R}^n$, $R > 0$,

$$\psi_{x_0,R}(x) = \psi\left(\frac{x - x_0}{R}\right) \Big/ R^n. \tag{3.6}$$

Moreover, once β, μ are fixed as above, we can find $0 \leqslant \bar{\psi} \leqslant 1$,

$$\operatorname{supp} \bar{\psi} \subset \{|x| < 2\},$$

$\bar{\psi} \equiv 1$ on $\{|x| < 1\}$, so that $C_\beta(\bar{\psi}) < +\infty$, $C_\mu(\bar{\psi}) < +\infty$. (See the remarks after (3.6) in [12].) We will fix such a $\bar{\psi}$ for the remainder of the paper.

Our first lemma is merely a version of Lemma 3.2 in [12], in our context.

Lemma 3.6. *Let u be a continuous non-negative weak solution of* (2.3) *in* $\mathbb{R}^n \times (0, T)$, *with* $\varphi \in \mathcal{B}_a$. *Let* β, μ *be as in* (3.3), (3.2) *respectively, and let* $\psi \in C_0^\infty(\mathbb{R}^n)$ *be such that* $A = C_\beta(\psi) < +\infty$, $B = C_\mu(\psi) < +\infty$. *Let*

$$G(b) = \int_0^b \frac{du}{Au^\beta + Bu^\mu}.$$

Let

$$f(t) = \int_{\mathbb{R}^n} u(x, t)\psi(x)\, dx.$$

Then

$$|G(f(t)) - G(f(s))| \leqslant |t - s| \quad \textit{for} \quad 0 < s, t < T. \tag{3.7}$$

PROOF. By Corollary 2.8, the following formal calculation is justified.

$$f'(t) = \int_{\mathbb{R}^n} \varphi(u(x, t))\, \Delta\psi(x)\, dx,$$

and hence,

$$\begin{aligned} |f'(t)| &\leqslant \int_{u\leqslant 1} \varphi(u(x,t))|\Delta\psi|\,dx + \int_{u>1} \varphi(u(x,t))|\Delta\psi|\,dx \\ &\leqslant \int_{u\leqslant 1} u^{\beta}(x,t)|\Delta\psi|\,dx + \int_{u>1} u^{\mu}(x,t)|\Delta\psi|\,dx \\ &\leqslant C_{\beta}(\psi) f(t)^{\beta} + C_{\mu}(\psi) f(t)^{\mu}, \end{aligned}$$

by Hölder's inequality. The lemma now follows by integrating the differential inequality.

Corollary 3.7. *Let u be as in Lemma* 3.6, *$\bar{\psi}$ as above. Then, there exists $C_1 > 0$, $C_2 > 0$, which depend only on $\beta, \mu, \bar{\psi}$ and n such that, if $0 < s, t < T$, and*

$$\max\left\{\int_{\mathbb{R}^n} u(x,t)\bar{\psi}_{x_0,R}(x)\,dx, \int_{\mathbb{R}^n} u(x,s)\bar{\psi}_{x_0,R}(x)\,dx\right\} \leqslant C_1, \tag{3.8}$$

then

$$\left|\left[\int_{\mathbb{R}^n} u(x,t)\bar{\psi}_{x_0,R}(x)\,dx\right]^{1-\beta} - \left[\int_{\mathbb{R}^n} u(x,s)\bar{\psi}_{x_0,R}(x)\,dx\right]^{1-\beta}\right| \leqslant C_2|t-s|/R^2, \tag{3.9}$$

while if the max *in* (3.8) *is bigger than C_1,*

$$\left|\left[\int_{\mathbb{R}^n} u(x,t)\bar{\psi}_{x_0,R}(x)\,dx\right]^{1-\mu} - \left[\int_{\mathbb{R}^n} u(x,s)\bar{\psi}_{x_0,R}(x)\,dx\right]^{1-\mu}\right| \leqslant C_2|t-s|/R^2. \tag{3.10}$$

PROOF. Note that $C_{\theta}(\bar{\psi}_{x_0,R}) = R^{-2}C_{\theta}(\bar{\psi})$. The corollary now follows from (3.17) and (3.3).

We are now ready to establish the existence of a trace as $t \downarrow 0$

Theorem 3.11. *Let u be a continuous non-negative weak solution of* (2.3) *in $\mathbb{R}^n \times (0,T)$, with $\varphi \in \mathcal{B}_a$. Then there exists a unique locally finite Borel measure μ on $\mathbb{R}^n$, such that*

$$\lim_{t\downarrow 0} \int_{\mathbb{R}^n} u(x,t)\psi(x)\,dx = \int_{\mathbb{R}^n} \psi(x)\,d\mu(x),$$

for all $\psi \in C_0^{\infty}(\mathbb{R}^n)$.

PROOF. Let $s = T/2$, and consider $0 < t < T/2$. The continuity of u in $\mathbb{R}^n \times (0,T)$, together with Corollary 3.7, shows that $\{u(x,t)\}$ has locally uniformly bounded mass. Hence, given $\{t_j\} \to 0$, we can find a subsequence $\{t_{r_j}\} \to 0$, and a locally finite Borel measure μ such that

$$\lim_{j\to\infty}\int_{\mathbb{R}^n} u(x,t_{r_j})\psi(x)\,dx = \int_{\mathbb{R}^n}\psi(x)\,d\mu(x)$$

for all $\psi\in C_0^\infty(\mathbb{R}^n)$.

Suppose that $\{s_j\}\to 0$, and let us consider the corresponding $\{s_{\theta_j}\}$ and $\tilde{\mu}$. We want to show that $\mu=\tilde{\mu}$. However, once again, Corollary 3.7, applied to $t=t_{r_j}$, $s=s_{\theta_j}$, shows that, for all x_0, R,

$$\int \bar{\psi}_{x_0,R}\,d\mu = \int \bar{\psi}_{x_0,R}\,d\tilde{\mu}.$$

The non-trivality of $\bar{\psi}$ now implies that $d\mu=d\tilde{\mu}$, and the theorem follows.

We now turn our attention to local L^∞ bounds for continuous weak solutions. This is the heart of the matter, and where we are forced to depart from the techniques in [12], due to the general class of non-linearities considered, since the Aronson-Bénilan inequality ([1]) is not available. Also, this is the only point where the lower bound in (3.1) enters (and of course, it necessarily enters here (see [5] and [12])).

Lemma 3.12. *Let u be a continuous, non-negative weak solution of* (2.3) *in* $Q^*=\{(x,t)\colon |x|<2,\ -4<t<0\}$, *with* $\varphi\in\mathfrak{B}_a$. *Let*

$$Q=\{(x,t)\colon |x|<1,\ -1<t<0\},$$

and define $H_\varphi(s)$ *to be* 1 *for* $0<s<1$, *and* $s[\varphi(s)/s]^{n/2}$ *for* $s\geqslant 1$. *Then, there is a constant* $C=C(0,n)>0$ *such that*

$$\|H_\varphi(u)\|_{L^\infty(Q)}\leqslant C\{\|u\|_{L^1(Q^*)}+1\}.$$

PROOF. First note that the lower bound in (3.1) implies that H_φ is increasing for $s>1$. Next, let p be as in Lemma 2.10 (and we can also assume that $p\geqslant 1$). For such a fixed p, we define now $H_{p,\varphi}(s)$ to be 1 for $0<s\leqslant 1$, and $s^p[\varphi(s)/s]^{n/2}$ for $s\geqslant 1$. Again, $H_{p,\varphi}$ is increasing for $s\geqslant 1$. Let $F_{p,\varphi}(s)$ be the inverse function to $H_{p,\varphi}$, defined again for $s\geqslant 1$. We first claim that

$$u(0,0)\leqslant CF_{p,\varphi}\left\{\iint_{Q^*} u^p(x,t)\,dx\,dt+1\right\} \tag{3.13}$$

In order to establish (3.13) we first note that if u is a continuous weak solution of $\partial u/\partial t=\Delta\varphi(u)$ in Ω, and $\alpha>0$, $\beta>0$, $\gamma>0$ are given, and we define $v(x,t)=u(\alpha x,\beta t)/\gamma$, then v is a continuous weak solution of $\partial v/\partial t=\Delta\psi(v)$ in the appropriate Ω', where $\psi(s)=\beta\varphi(\gamma s)\alpha^{-2}\gamma^{-1}$. Suppose now that $\beta/\alpha^2\geqslant 1$, and $\varphi\in\mathfrak{B}_a$. Let $G_\varphi(\gamma)=\gamma/\varphi(\gamma)$, which by the right-hand side inequality in (3.1), is increasing for $\gamma\geqslant 1$. Let $\Gamma_\varphi(s)$ be its inverse function defined for $s\geqslant 1$, and choose $\gamma=\Gamma_\varphi(\beta/\alpha^2)$. Then, $\psi(1)=\beta\varphi(\gamma)\alpha^2\gamma^{-1}=1$, and it is easy to see that $\psi\in\mathfrak{B}_a$.

Assume now that (3.13) fails. We can then find $\varphi_k \in \mathfrak{B}_a$, u_k continuous weak solutions such that

$$u_k(0,0) \geqslant kF_{p,\varphi_k}\left\{\iint_{Q^*} u_k^p(x,t)\,dx\,dt + 1\right\}.$$

First note that $F_{p,\varphi_k}(1) = 1$, and that $F_{p,\varphi_k}(s)$ is increasing for $s \geqslant 1$, so that $u_k(0,0) \geqslant k$. Because of (2.11), this forces that

$$I_k = \iint_{Q^*} u_k^p(x,t)\,dx\,dt \to +\infty \quad \text{as} \quad k \to \infty.$$

For α_k small (to be chosen momentarily), let

$$v_k(x,t) = \frac{u_k(\alpha_k x, t)}{\gamma_k}, \qquad \gamma_k = \Gamma_{\varphi_k}\left(\frac{1}{\alpha_k^2}\right).$$

The v_k are continuous weak solutions of $\partial v/\partial t = \Delta\psi_k(v)$, $\psi_k \in \mathfrak{B}_a$ in Q^*, by our previous discussion.

$$\int_{|x|<2/\alpha_k}\int_{-4}^{0} v_k^p(x,t)\,dx\,dt = \frac{1}{\Gamma_{\varphi_k}(1/\alpha_k^2)^p}\,\alpha_k^{-n} I_k.$$

Choose α_k so that

$$\frac{1}{\Gamma_{\varphi_k}(1/a_k^2)^p}\,\alpha_k^{-n} I_k = 1,$$

or equivalently

$$I_k = \frac{\Gamma_{\varphi_k}(1/\alpha_k^2)^p}{1/\alpha_k^n}.$$

This is possible because $\Gamma_{\varphi_k}^p(s)/s^{n/2}$ is increasing for $s \geqslant 1$, by the left-hand side inequality in (3.1), and $I_k \to +\infty$. Moreover, $\alpha_k \to 0$ as $k \to \infty$. But then,

$$\begin{aligned}
v_k(0,0) &= \frac{1}{\varphi_{\varphi_k}(1/\alpha_k^2)}\,u_k(0,0) \\
&\geqslant \frac{k}{\Gamma_{\varphi_k}(1/\alpha_k^2)}\,F_{p,\varphi_k}\{I_k + 1\} \\
&\geqslant \frac{k}{\Gamma_{\varphi_k}(1/\alpha_k)}\,F_{p,\varphi_k}(I_k) \\
&= kF_{p,\varphi_k}\left\{\frac{\Gamma_{\varphi_k}(1/\alpha_k^2)^p}{\alpha_k^n}\right\}\frac{1}{\Gamma_{\varphi_k}(1/\alpha_k^2)} \\
&= k,
\end{aligned}$$

by the definition of F_{p,φ_k}. However, this contradicts (2.11), by our choice of α_k, and thus (3.13) is established.

We next note that by using translation and dilations $(x, t) \to (\alpha x, \alpha^2 t)$, (3.13) has, as a consequence,

$$\|u\|_{L^\infty(B_{r/2}\times(-r^2/4,0])} \leqslant CF_{p,\varphi}\left\{\frac{1}{r^{n+2}}\iint_{B_r\times(-r^2,0)} u^p(x,t)\,dx\,dt + 1\right\} \tag{3.14}$$

for $0 < r < 2$. Finally, using (3.14), translations, the dilations $(x, t) \to (\alpha x, \alpha^2 t)$, and a simple covering argument, one can show

$$\|u\|_{L^\infty(S)} \leqslant CF_{p,\varphi}\left\{\frac{1}{(\rho-r)^n + (\rho-r)^2}\iint_R u^p(x,t)\,dx\,dt + 1\right\}, \tag{3.15}$$

where $S = B_r \times (-r^2, 0]$, $R = B_\rho \times (-\rho^2, 0]$, $1/2 < r < \rho < 2$.

To conclude the proof of Lemma 3.12, we use an argument which originates in the work of Hardy and Littlewood (see [11]), and which was first used in the context of the porous medium equation in [6].

We will first show that

$$\|u\|_{L^\infty(Q)} \leqslant C\{\|u\|_{L^1(Q^*)} + 1\}^\sigma, \tag{3.16}$$

where $\sigma = \sigma(a, n)$. Once (3.16) is established, the lemma will follow by repeating the argument that we used to establish (3.15), with $p = 1$, and using (3.16) instead of (2.11).

In order to establish (3.16) we need to point out to two properties of $F_{p,\varphi}$. First, $F_{p,\varphi}(As) \leqslant CA^\sigma F_{p,\varphi}(s)$, for $A \geqslant 1$, $s \geqslant 1$. This is an easy consequence of (3.1). Another easy consequence of (3.1) is that $F_{p,\varphi}(s^{p-1}) \leqslant Cs^{1-\epsilon}$, where $\epsilon = \epsilon(a, n)$.

Now, for $1/2 < r < 1$, let

$$\begin{aligned} S_r &= B_{2r} \times (-4r^2, 0], \\ m(r) &= \|u\|_{L^\infty(S_r)}, \\ I &= \iint_{Q^*} u(x,t)\,dx\,dt, \\ J &= \max\{I, 1\}. \end{aligned}$$

We want to show that

$$m(1/2) \leqslant C\{I + 1\}^\sigma. \tag{3.17}$$

If there exists r, $1/2 < r < 1$ such that $m(r) \leqslant 1$, we are done, and hence, we can assume that $m(r) > 1$ for all r, $1/2 < r < 1$. Pick now $1/2 < r < \rho < 1$. Then, (3.15) implies that

$$m(r) \leqslant CF_{p,\varphi}\left\{\frac{1}{(r-\rho)^{n+2}}\iint_{S_\rho} u^p(x,t)\,dx\,dt + 1\right\}$$

$$\leqslant CF_{p,\varphi}\left\{\frac{4J}{(r-\rho)^{n+2}}\, m(\rho)^{p-1}\right\}$$

$$\leqslant C\left\{\frac{J}{(r-\rho)^{n+2}}\right\}^{\sigma} F_{p,\varphi}(m(\rho)^{p-1})$$

$$\leqslant C\left\{\frac{J}{(r-\rho)^{n+2}}\right\}^{\sigma} m(\rho)^{1-\epsilon}.$$

Choose now γ, $0 < \gamma < 1$ so that $\theta = (1-\epsilon)/\gamma < 1$, and let $\rho = r^\gamma$. Taking logarithms, we see that

$$\log m(r) \leqslant C\log J + C\log\frac{1}{(r-r^\gamma)} + (1-\epsilon)\log m(r^\gamma).$$

Integrating with respect to the measure dr/r, between $1/2$ and 1, we obtain

$$\int_{1/2}^{1} \log m(r)\frac{dr}{r} \leqslant C\log J + C + \theta\int_{1/2^\gamma}^{1} \log m(r)\frac{dr}{r}$$

$$\leqslant C\log J + C + \theta\int_{1/2}^{1} \log m(r)\,\frac{dr}{r}$$

(3.17) immediately follows from this, and our lemma is established.

Remark 3.18. In the case when $\varphi(u) = u^m$, $0 < m < 1$, the technique of proof of Lemma 3.12 allows one to show

$$\|u\|_{L^\infty(Q)} \leqslant C_q\left\{\left(\iint_{Q^*} u^q\right)^{\frac{1}{q-(1-m)n/2}} + 1\right\}, \tag{3.19}$$

for each q such that $q - (1-m)n/2 > 0$. Note that $q = 1$ is allowed precisely when $m > (n-2)/n$, giving another explanation to the result in [5]. Moreover, at least in the range $m > (n-2)/n$, an inequality of the form

$$\|u\|_{L^\infty(Q)} \leqslant C\left\{\iint_{Q^*} u^q + 1\right\}^{\sigma}$$

can only hold if $q - (1-m)n/2 > 0$, as can be seen by considering the Barenblatt solutions ([3], [1], [12], [5])

$$U_\alpha(x,t) = t^{-\beta}\{\alpha + 2m\gamma^{-1}|x|^2 t^{-2\beta/n}\}^{-s},$$

where $\alpha > 0$, $s = \dfrac{1}{1-m}$, $\gamma = \dfrac{2}{1-m} - n$, $\beta^{-1} = m - 1 + \dfrac{2}{n}$.

Corollary 3.20. *Let u be a continuous weak solution of* (2.3) *in $B_{4R} \times (0, T)$, continuous in $\bar{B}_{4R} \times [0, T]$, with $\varphi \in \mathfrak{B}_a$. Suppose that $T/R^2 > 1$. Then,*

(3.21)
$$\sup_{|x|<R} H_\varphi(u(x, T)) \leqslant C\left\{\frac{1}{T^{n/2}} \int_{|x|<4R} u(x, 0)\, dx + \Gamma_\varphi(T/R^2) \cdot (T/R^2)^{n/2}\right\}.$$

PROOF. Apply Lemma 3.12 to

$$v(x, t) = \frac{u(\alpha x, \beta t)}{\gamma},$$

where $\alpha = R$, $\beta = T$, and $\gamma = \Gamma_\varphi(\beta/\alpha^2)$, and then apply Corollary 3.7 to the same v.

Note that in the case when $\varphi(u) = u^m$, $\dfrac{n-2}{n} < m < 1$, Theorems 3.20 is exactly Theorem 2.2 in [12].

We now turn our attention to the existence of solutions in $\mathbb{R}^n \times (0, \infty)$.

Theorem 3.22. *Let μ be a locally finite Borel measure on $\mathbb{R}^n$. Then, for any $\varphi \in \mathfrak{B}_a$, there exists a continuous weak solution u to* (2.3) *in $\mathbb{R}^n \times (0, \infty)$, such that*

$$\lim_{t \downarrow 0} \int_{\mathbb{R}^n} u(x, t)\eta(x)\, dx = \int_{\mathbb{R}^n} \eta(x)\, d\mu(x), \tag{3.23}$$

for all $\eta \in C_0^\infty(\mathbb{R}^n)$.

PROOF. Our starting point is the following classical result (see [15]): for any $f \in C_0^\infty(\mathbb{R}^n)$, there exists $u \in C([0, \infty); L^1(\mathbb{R}^n)) \cap L^\infty(\mathbb{R}^n \times [0, \infty))$, which is a weak solution of (2.3) in $\mathbb{R}^n \times (0, \infty)$, and such tht $u(x, 0) = f(x)$. Moreover, the results in [17] show that, if φ verifies (2.1), then u is continuous in $\mathbb{R}^n \times (0, \infty)$.

Our next claim is that if $R \geqslant 1$, $\int_{B_{4R}} f\, dx \leqslant M$, then the corresponding u are equicontinuous in compact subsets of $B_R \times (0, \infty)$. To establish this, note that Corollary 3.20 establishes the local uniform boundedness in $B_R \times [R^2, +\infty)$, and hence [17] gives the equicontinuity there. If $t < R^2$, choose $r^2 = t < R^2$, and apply the same argument to $B_r \times [r^2, +\infty)$ and its translates, to obtain the desired equicontinuity.

Finally, note that if $\eta \in C_0^\infty(\mathbb{R}^n)$, $\operatorname{supp} \eta \subset B_R$, $\int_{B_{4R}} f\, dx \leqslant M$, and u is as above, then

$$\left| \int \eta(x)[u(x, t) - f(x)]\, dx \right| \leqslant tC_{M,\eta}. \tag{3.24}$$

In fact,

$$\int \eta(x)[u(x,t) - f(x)] = \int_0^t \int_{\mathbb{R}^n} \Delta\eta\,(x)\varphi(u(x,s))\,dx\,ds,$$

and so, by (3.2)(i), the left-hand side of (3.24) is bounded by

$$\int_0^t \int_{\mathbb{R}^n} |\Delta\eta\,(x)|\,dx\,ds + \int_0^t \int_{\mathbb{R}^n} |\Delta\eta\,(x)| u^\mu(x,s)\,dx\,ds, \qquad 0 < \mu < 1.$$

An application of Hölder's inequality and Corollary 3.7 establishes (3.24).

Fix now the measure μ, and pick $f_k \in C_0^\infty(\mathbb{R}^n)$ such that

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} f_k(x)\eta(x)\,dx = \int \eta(x)\,d\mu(x).$$

We can choose f_k so that, in addition

$$\int_{B_R} f_k(x)\,dx \leqslant M_R,$$

where M_R is independent of k. Let u_R be the corresponding solutions, constructed at the beginning of the proof. By equicontinuity (after possibly passing to a subsequence, which we still denote $\{u_k\}$), there exists a continuous weak solution u of (2.3), such that the u_k converge to u, uniformly on compact subsets of $\mathbb{R}^n \times (0, \infty)$. (3.24) now establishes (3.23).

Remark 3.25. The results in [5] show that, unless the left-hand inequality of (3.1) is verified (in the pure power case $\varphi(u) = u^m$), Theorem 3.22 fails when μ is the delta mass at the origin. As our proof shows, this is because of the lack of an $L^\infty - L^1$ «regularity» effect, as Lemma 3.12. This clarifies Remark 3.18.

Finally, we turn to the uniqueness of the solution constructed in Theorem 3.22. The general strategy is the one developed in [6]. As in [6] and [7], we start out with a version of the maximum principle. Its proof follows closely that of Theorem 2.3 in [12].

Lemma 3.26. *Let u, v be continuous weak solutions of* (2.3) *in*

$$S_T = \mathbb{R}^n \times (0, T),$$

with $\varphi \in \mathfrak{B}_a$. Assume that

$$\lim_{t\downarrow 0} \int_{|x|<R} [v(x,t) - u(x,t)]^+\,dx = 0 \tag{3.27}$$

for all $R > 0$, where $A^+ = \max\{A, 0\}$. Then, $v \leqslant u$ in S_T.

PROOF. Let $w = v - u$, and let q denote the characteristic function of the set where $u < v$. If u, v and φ are smooth, Kato's inequality ([13]) shows that

$$\Delta[\varphi(v) - \varphi(u)]^+ \geqslant q\Delta[\varphi(v) - \varphi(u)],$$

where

$$\frac{\partial w^+}{\partial t} = q\frac{\partial w}{\partial t}.$$

By (3.1), (3.2) and (3.4) we see that

$$[\varphi(v) - \varphi(u)]^+ \leqslant C\{([v-u]^+)^\mu + ([v-u]^+)^\beta\}.$$

Hence, (still under the assumption that u, v, φ are smooth), we have, for $\psi \in C_0^\infty(\mathbb{R}^n)$, $\psi \geqslant 0$.

$$\begin{aligned}
\frac{\partial}{\partial t}\int \psi(x)w^+(x,t)\,dx &\leqslant \int \psi(x)\,\Delta[\varphi(v) - \varphi(u)]^+\,dx \\
&\leqslant \int |\Delta\psi(x)|[\varphi(v) - \varphi(u)]^+\,dx \\
&\leqslant C\int |\Delta\psi(x)|([v-u]^+)^\mu + C\int |\Delta\psi(x)|([v-u]^+)^\beta \\
&\leqslant CC_\mu(\psi)\left(\int \psi(x)w^+(x,t)\,dx\right)^\mu \\
&\quad + CC_\beta(\psi)\left(\int \psi(x)w^+(x,t)\,dx\right)^\beta.
\end{aligned}$$

Integrating this (one-sided) differential inequality in a manner similar to the proof of Lemma 3.6, using Corollary 2.8 to justify the formal argument above, and using (3.27), we see that, if $C_\mu(\psi) < \infty$, $C_\beta(\psi) < \infty$, there are constants C_1, C_2 such that, if

$$\frac{1}{R^n}\int_{|x-\xi|<R}[v(x,t) - u(x,t)]^+\,dx \leqslant C_1,$$

then

$$\left(\frac{1}{R^n}\int_{|x-\xi|<R}[v(x,t) - u(x,t)]^+\,dx\right)^{1-\beta} \leqslant C_2 tR^{-2},$$

while if the above quantity exceeds C_1, then,

$$\left(\frac{1}{R^n}\int_{|x-\xi|<R}[v(x,t) - u(x,t)]^+\,dx\right)^{1-\mu} \leqslant C_2 tR^{-2},$$

for any $R > 0$, $\xi \in \mathbb{R}^n$.

Let now $h(x,t) = \int_0^t [\varphi(v) - \varphi(u)]^+ \, dx$. It is easy to see that, for each $t > 0$, h is a subharmonic function of x. Hence, for $\xi \in \mathbb{R}^n$,

$$
\begin{aligned}
h(\xi,t) &\leqslant \frac{1}{\omega_n R^n} \int_{B_R(\xi)} h(x,t)\,dx \\
&\leqslant \frac{1}{\omega_n R^n} \int_0^t \int_{B_R(\xi)} [\varphi(v) - \varphi(u)]^+ \\
&\leqslant \frac{C}{R^n} \int_0^t \int_{B_R(\xi)} \{([v-u]^+)^\mu + ([v-u]^+)^\beta\} \\
&\leqslant \frac{C}{R^n} \int_0^t R^{n(1-\mu)} \left(\int_{B_R(\xi)} [v-u]^+ \right)^\mu + R^{n(1-\beta)} \left(\int_{B_R(\xi)} [v-u]^+ \right)^\beta \\
&\leqslant \frac{C}{R^n} \int_0^t R^{n(1-\mu)} \{ s^{\mu/(1-\mu)} R^{-2\mu/(1-\mu)} + s^{\mu/(1-\beta)} R^{-2\mu/(1-\beta)} \} \\
&\quad + \frac{C}{R^n} \int_0^t R^{n(1-\beta)} \{ s^{\beta/(1-\beta)} R^{-2\beta/(1-\beta)} + s^{\beta/(1-\mu)} R^{-2\beta/(1-\mu)} \}
\end{aligned}
$$

For fixed t, this tends to 0 as $R \to \infty$, and hence $h(x,t) \equiv 0$, which establishes the lemma.

Our general uniqueness result now follows from the approximation procedure in [6], using Lemma 2.9 and Lemma 3.26.

Theorem 3.28. *Let u, v be continuous weak solutions of* (2.3) *in S_T, with $\varphi \in \mathfrak{B}_a$. Assume that*

$$\lim_{t \downarrow 0} \int u(x,t)\eta(x)\,dx = \lim_{t \downarrow 0} \int v(x,t)\eta(x)\,dx$$

for all $\eta \in C_0^\infty(\mathbb{R}^n)$. Then $u \equiv v$ in S_T.

PROOF. Let μ be the locally finite Borel measure on $\mathbb{R}^n$, attached to u by Theorem 3.11. Pick $h \in C_0^\infty(\mathbb{R}^n)$, $0 \leqslant h \leqslant 1$, and let $w(x,t,h)$ be a solution in $\mathbb{R}^n \times (0,\infty)$, with initial trace $h\mu$. By Theorem 3.22, at least one such w exists. Moreover, Corollary 3.20 shows that any such w belongs to $L^\infty(\mathbb{R}^n \times [\tau_1, \tau_2])$, for each $0 < \tau_1 < \tau_2 < +\infty$. We now want to show that, for any such w,

$$\sup_{t>0} \int_{\mathbb{R}^n} w(x,t,h)\,dx < +\infty. \tag{3.29}$$

In fact, Corollary 3.7 easily implies that for each $t > 0$, $R > 0$,

$$(3.30)\quad \int_{R/2<|x|<R} w(x,t,h)\frac{dx}{|x|^{n-2}}$$

$$\leqslant \int_{R/2<|x|<2R} h(x)\frac{d\mu(x)}{|x|^{n-2}} + Ct^{1/(1-\beta)}R^{2-[2/(1-\beta)]} + Ct^{1/(1-\mu)}R^{2-[2/(1-\mu)]},$$

and hence

$$(3.31)\quad \int_{|x|>1} w(x,t,h)\frac{dx}{|x|^{n-2}} < +\infty.$$

Pick now $\psi(x) \in C^\infty(\mathbb{R}^n)$, $\psi(x) > 0$, $\psi(0) = 1$, ψ bounded, $\Delta\psi \leqslant 0$, $\psi(x) \leqslant C/|x|^n$ for $|x| > 1$, $|\nabla\psi(x)| \leqslant C/|x|^{n-1}$ for $|x| > 1$, and let $\psi_R(x) = \psi(x/R)$. Let $\theta_N(x)$ be a $C_0^\infty(\mathbb{R}^n)$ function, $0 \leqslant \theta_N \leqslant 1$, $\theta_N \equiv 1$ for $|x| < N$, $\operatorname{supp}\theta_N \subset \{|x| < 2N\}$, $|\nabla\theta_N| \leqslant C/N$, $|\Delta\theta_N| \leqslant C/N^2$. Then, for $0 < s < t < +\infty$, (with $w(x,t) = w(x,t,h)$),

$$\int w(x,t)\psi_R(x)\,dx - \int w(x,s)\psi_R(x)\,dx = \lim_{N\to\infty}\int [w(x,t) - w(x,s)]\theta_N(x)\psi_R(x)\,dx$$

$$= \lim_{N\to\infty}\int_s^t\int_{\mathbb{R}^n} \varphi(w(x,r))\Delta[\theta_N(x)\psi_R(x)]\,dx$$

$$\leqslant \overline{\lim_{N\to\infty}}\int_s^t\int_{\mathbb{R}^n} \varphi(w(x,r))[\Delta\theta_N(x)\psi_R(x)$$

$$+ 2\nabla\theta_N(x)\nabla\psi_R(x)]\,dx\,dr.$$

If we now use the pointwise estimates for ψ_R, $\nabla\psi_R$, $\Delta\theta_N$, $\nabla\theta_N$, the support properties of the last two functions, (3.2), (3.30), and the boundedness of w in $\mathbb{R}^n \times [s,t]$, we see that the above $\overline{\lim}_{N\to\infty}$ is non-positive. Hence

$$\int w(x,t)\psi_R(x)\,dx \leqslant \int w(x,s)\psi_R(x)\,dx.$$

Moreover, (3.30), and our pointwise bounds on ψ show that

$$\lim_{s\to 0}\int w(x,s)\psi_R(x)\,dx = \int h(x)\psi_R(x)\,d\mu(x).$$

The last expression is bounded, independently of R, and so

$$\int w(x,t)\psi_R(x)\,dx \leqslant C$$

and Fatou's lemma implies (3.29), letting $R \to \infty$. Hence, Lemma 2.9 now shows there is only one such $w(x,t,h)$. We next claim that $w(x,t,h) \leqslant u(x,t)$.

Let $U_\epsilon(x, t)$ be a solution with initial data $h(x)u(x, \epsilon)$. Theorem 3.11 and the above argument show that there exists exactly one such U_ϵ, and that

$$\sup_{t>0} \int_{\mathbb{R}^n} U_\epsilon(x, t)\, dx \leqslant C \sup_{0<\epsilon<T/2} \int h(x)u(x, \epsilon)\, dx \leqslant C_h. \tag{3.32}$$

where the last inequality is a consequence of Corollary 3.7. Moreover, for any smooth η, we have that

$$\lim_{\epsilon \to 0} \int \eta(x)h(x)u(x, \epsilon)\, dx = \int \eta(x)h(x)\, d\mu\,(x).$$

Note now that $\lim_{\epsilon \to 0} U_\epsilon(x, t) = w(x, t)$. In fact, (3.32), Corollary 3.20 and the results in [17] show that $\{U_\epsilon(x, t)\}$ is equicontinuous on compact subsets of $\mathbb{R}^n \times (0, \infty)$. Let $\{U_{\epsilon_j}(x, t)\}$ be a subsequence, which converges uniformly on compact subsets of $\mathbb{R}^n \times (0, \infty)$ to $\hat{w}(x, t)$. Clearly, $\hat{w}(x, t)$ is a continuous weak solution of (2.3) in $\mathbb{R}^n \times (0, \infty)$. We claim that for any $\eta \in C_0^\infty(\mathbb{R}^n)$,

$$\lim_{t \downarrow 0} \int \hat{w}(x, t)\eta(x)\, dx = \int h(x)\eta(x)\, d\mu\,(x).$$

In fact, the argument leading to (3.24), together with the second inequality in (3.32) show that

$$\left| \int \eta(x)[U_\epsilon(x, t + \epsilon) - h(x)u(x, \epsilon)]\, dx \right| \leqslant tC_h$$

which proves our claim. But the uniqueness of $w(x, t, h)$ then shows that $\hat{w}(x, t) = w(x, t, h)$, and our claim follows. The construction of $U_\epsilon(x, t)$ in [15] gives that $U_\epsilon(x, t) \in C([0, \infty); L^1(\mathbb{R}^n))$, so that

$$\lim_{t \downarrow 0} \int_{\mathbb{R}^n} |U_\epsilon(x, t) - h(x)u(x, \epsilon)|\, dx = 0.$$

By the continuity of $u(x, t)$,

$$\lim_{t \downarrow 0} \int_K |u(x, t + \epsilon) - u(x, \epsilon)| dx = 0,$$

for each $K \subset\subset \mathbb{R}^n$. Hence, by Lemma 3.26, $U_\epsilon(x, t) \leqslant u(x, t + \epsilon)$, and thus, $w(x, t, h) \leqslant u(x, t)$. Pick now $0 \leqslant h_j \leqslant h_{j+1} \leqslant 1$, $h_j \in C_0^\infty(\mathbb{R}^n)$, $\lim_{j \to \infty} h_j(x) \equiv 1$. By the construction of $w(x, t, h)$ given in Theorem 3.11,

$$w(x, t, h_j) \leqslant w(x, t, h_{j+1}).$$

Moreover, by the first part of the proof, $w(x, t, h_j) \leqslant u(x, t)$. Let $w_\infty(x, t)$ be a limit of some subsequence of $w(x, t, h_j)$, which exists and is a continuous

weak solution, by equicontinuity. $w_\infty(x,t)$ has a trace, as $t \downarrow 0$ by Theorem 3.22. The trace is between $h_j\, d\mu$ and $d\mu$ for each j, and hence it equals $d\mu$. Since $w_\infty \leqslant u$,

$$\lim_{t\downarrow 0} \int_{|x|<R} |u(x,t) - w_\infty(x,t)| = 0$$

for all $R > 0$. By Lemma 3.26, $u(x,t) = w(x,t)$. Similarly, $v(x,t) = w_\infty(x,t)$, and hence $u(x,t) = v(x,t)$.

Theorem 3.33. *Let u be a continuous weak solution of* (2.3) *in S_T, with $\varphi \in \mathfrak{B}_a$. Then there exists a unique $\hat{u}$ in $S_\infty = \mathbb{R}^n \times (0,\infty)$, which is a continuous weak solution of* (2.3) *in S_∞, with $u = \hat{u}$ in S_T.*

PROOF. Let μ be the trace of u given by Theorem 3.11, and let $\hat{u}$ be the corresponding solution in S_∞, constructed in Theorem 3.22. By Theorem 3.28, $u = \hat{u}$ in S_T.

References

[1] Aronson, D. G. and Bénilan, Ph. Régularité des solutions de l'équation des milieux poreux dans $\mathbb{R}^N$, *C. R. Acad. Sci. Paris*, Série A, **288**(1979), 103-105.

[2] Aronson, D. G. and Caffarelli, L. A. The initial trace of a solution of the porous medium equation, *Trans. Amer. Math. Soc.* **280**(1983), 351-366.

[3] Barenblatt, G. I. On some unsteady motions of a liquid or a gas in a porous medium, *Prikl. Mat. Mekh.* **16**(1952), 67-78.

[4] Bénilan, Ph., Crandall, M. and Pierre, M. Solutions of the porous medium equation in $\mathbb{R}^N$ under optimal conditions on initial values, *Indiana U. Math. J.* **33**(1984), 51-87.

[5] Brézis, H., and Friedman, A. Nonlinear parabolic equations involving measures as initial conditions, *J. Math. Pures et Appl.* **62**(1983), 73-97.

[6] Dahlberg, B. E. J. and Kening, C. E., Non-negative solutions of the porous medium equation, *Comm. P. D. E.* **9**(1984), 409-437.

[7] Dahlberg, B. E. J. and Kenig, C. E. Non-negative solutions of generalized porous medium equations, *Revista Matemática Iberoamericana,* **2**(1986), 207-305.

[8] Dahlberg, B. E. J. and Kenig, C. E. Non-negative solutions of the initial-Dirichlet problem for generalized porous medium equations in cylinders, *Journal of the Amer. Math. Soc.*, **1**(1988), 401-412.

[9] Dahlberg, B. E. J. and Kenig, C. E. Weak solutions of the porous medium equation, preprint.

[10] Dahlberg, B. E. J. and Kenig, C. E. Weak solutions of the porous medium equation in a cyclinder, preprint.

[11] Fefferman, C. and Stein, E. M. H^p spaces of several variables, *Acta Math.* **129**(1972), 137-193.

[12] Herrero, M. and Pierre, M. The Cauchy problem for $u_t = \Delta u^m$ when $0 < m < 1$, *Trans. Amer. Math. Soc.* **291**(1985), 145-158.
[13] Kato, T. Schrödinger operators with singular potentials, *Israel J. Math.* **13**(1973), 135-148.
[14] Moser, J. A Harnack inequality for parabolic differential equations, *Comm. Pure Appl. Math.* **17**(1964), 101-134.
[15] Peletier, L. A. The porous medium equation, *Applications of Nonlinear Analysis in the Physical Sciences*, Bielefeld, 1979.
[16] Pierre, M. Uniqueness of the solutions of $u_t - \Delta\varphi(u) = 0$ with initial datum a measure, *Nonlinear Anal. T.M.A.* **6**(1982), 175-187.
[17] Sacks, P. E. Continuity of solutions of a singular parabolic equation, *Nonlinear Anal. T.M.A.* **7**(1983), 387-409.

Bjorn E. J. Dahlberg*
Department of Mathematics
Washington University
St. Louis, Mo. 63130, USA

Carlos E. Kenig*
Department of Mathematics
University of Chicago
Chicago, Il 60637, USA

* Both authors were partially supported by the N. S. F.

REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.° 1, 1988

Constructions de Bases Orthonormées D'Ondelettes

Yves Meyer

A la mémoire de José-Luis Rubio de Francia

1. Introduction

Bien des bases orthonormées ont été découvertes, puis utilisées par les mathématiciens, les physiciens et les ingénieurs. Ces bases étaient, le plus souvent construites par l'un ou l'autre des deux procédés suivants. Soit elles se composaient des fonctions propres d'un opérateur auto-adjoint compact, soit elles étaient obtenues en appliquant la procédure standard de Gram-Schmidt à une suite convenable de fonctions. Dans les deux cas, les problèmes que l'on rencontre sont les suivants:

(1.1) la base orthonormée $u_k(x)$ n'est pas numériquement explicite
(1.2) les coefficients a_k d'une série $\sum a_k u_k(x)$ ne donnent qu'une information en moyenne et de nature globale sur les propriétés de la somme $f(x)$.

Par exemple, il est impossible d'estimer la norme L^p, $p \neq 2$, de la somme $f(x)$ d'une série trigonométrique en ne connaissant que les modules des coefficients de Fourier de $f(x)$. De même, les coefficients de Fourier ne fournissent pas d'information directe et accessible sur des propriétés locales, comme la dérivabilité en un point donné.

Au début des années 80, les ondelettes sont apparues comme la seule méthode permettant d'éviter toutes les difficultés que nous venons de décrire. Par exemple, en traitement du signal, on a cherché à enregistrer numériquement des phénomènes transitoires, relatifs à des échelles de temps qu'on ne connait pas à l'avance, ce qui est impossible lorsqu'on se limite à des séries ou des intégrales de Fourier traditionnelles. Les ondelettes ont une description numéri-

que essentiellement triviale, puisqu'elles sont engendrées à partir d'une fonction de base $w(x)$, la «mère des ondelettes», par des translations entières, suivies de dilatations dyadiques. En revanche, la base orthonormée des ondelettes n'est pas l'ensemble des fonctions propres d'un opérateur différentiel auto-adjoint. C'est évident car un tel opérateur devrait être invariant par translations et dilatations, c'est à dire être un opérateur à coefficients constants. Or seules les exponentielles sont les fonctions propres de tels opérateurs. Tout cela signifie que nous ne disposerons d'aucune «recette bon marché» pour construire les bases orthonormées d'ondelettes et explique pourquoi elles ont été découvertes si tardivement.

En fait, le système de Haar a été inventé par Alfred Haar en 1909 et il a fallu attendre 1981 pour que les chaînons suivants apparaissent.

Rappelons que le système de Haar, dans sa version convenant à $L^2(\mathbb{R})$, se compose des fonctions $2^{j/2}h(2^jx - k)$, $j \in \mathbb{Z}$, $k \in \mathbb{Z}$, où $h(x)$ est nulle hors de $[0,1)$, est égale à 1 sur $[0, 1/2)$ et à -1 sur $[1/2, 1)$.

Nous dirons qu'une fonction $w(x)$, de la variable réelle x, est une *ondelette* d'ordre $m \geqslant 1$ si $w(x)$ appartient à l'espace de Sobolev $H^m(\mathbb{R})$, si $w(x)$, ainsi que ses dérivées d'ordre $\leqslant m$, est bornée et a une décroissance rapide à l'infini et si $2^{j/2}w(2^jx - k)$, $j \in \mathbb{Z}$, $k \in \mathbb{Z}$, est une base orthonormée de $L^2(\mathbb{R})$. Ces conditions impliquent que tous les moments de $w(x)$, d'ordre $\leqslant m$ soient nuls. Une ondelette est donc une fonction réguliére, bien localisée et oscillante, c'est à dire une petite onde.

Dans cet article nous nous proposons d'abord de rendre justice à un travail de Stromberg, paru en 1981, où la première base orthonormée d'ondelettes, était construite, puis de comparer les ondelettes de Stromberg à celles que l'on obtient aujourd'hui par des procédés plus généraux. Nous terminerons en examinant un problème ouvert: existe-*t*-il une méthode universelle fournissant toutes les bases orthonormées d'ondelettes?

2. Les Ondelettes de Stromberg

Nous débutons par la description d'une suite emboîtée V_j, $j \in \mathbb{Z}$, de sous espaces de $L^2(\mathbb{R})$. Désignons, tout d'abord, par V_0 le sous espace de $L^2(\mathbb{R})$ qui se compose de toutes les fonctions splines d'ordre $m \geqslant 1$ dont les noeuds appartiennent à $\mathbb{Z}$. En d'autres termes, $f(x)$ appartient à V_0 si $f(x)$ appartient à l'espace de Sobolev H^m et si la restriction de $f(x)$ à tout intervalle $[k, k+1]$, $k \in \mathbb{Z}$, est un polynome de degré au plus égal à m. A partir de V_0 on construit les V_j, $j \in \mathbb{Z}$, en imposant la condition que $f(x)$ appartient à V_j si et seulement si $f(2x)$ appartient à V_{j+1}. Alors V_j est inclus dans V_{j+1} et l'on désigne par W_j le complément orthogonal de V_j dans V_{j+1}. Tous les espaces W_j se déduisent de W_0 par la remarque évidente que $f(x)$ appartient à W_0 si et seulement

si $f(2^j x)$ appartient à W_j. Le lemme qui suit nous apprend qu'une fonction $w(x)$, appartenant à W_0, ayant une décroissance rapide à l'infini et telle que la suite $w(x-k)$, $k \in \mathbb{Z}$, soit une base ortonormée de W_0 est automatiquement une ondelette.

En fait, il suffit de demander que $w(x-k)$, $k \in \mathbb{Z}$, soit une suite orthonormée et que $w(x)$ ait une décroissance rapide à l'infini.

Lemme 1. *Sous les hypothéses précédentes, la suite* $2^{j/2} w(2^j x - k)$ $j \in \mathbb{Z}$, $k \in \mathbb{Z}$, *est automatiquement une base orthonormée de l'espace* L^2.

En effet L^2 est la somme orthogonale des W_j, $j \in \mathbb{Z}$, et, pour chaque j, les fonctions $2^{j/2} w(2^j x - k)$, $k \in \mathbb{Z}$, forment une base orthonormée de W_j, comme on le voit par un simple changement d'échelle.

Pour construire $w(x)$, Stromberg emploie la méthode suivante. Il introduit un espace à mi-chemin entre V_0 et V_1, à savoir l'espace $V_{1/2}$ des fonctions appartenant à V_1 et dont la restriction à $[0, \infty)$ coincide avec celle d'une fonction de V_0. En d'autres termes, une fonction de $V_{1/2}$ est une fonction spline, de carré sommable, dont les noeuds sont les entiers positifs, négatifs ou nuls et les demi-entiers négatifs. On désigne par T l'opérateur de translation par 1, agissant sur $L^2(\mathbb{R})$. Il est immédiat que $V_{1/2}$ est inclus dans $T(V_{1/2})$ mais il est moins évident que la codimension de $E = V_{1/2}$ dans $F = T(V_{1/2})$ soit égale à 1. Pour le voir, on suit Stromberg et l'on considère la forme linéaire L sur F définie par

$$L(f) = (d/dx)^m f(1/2 + 0) - (d/dx)^m f(1/2 - 0).$$

Le noyau de L est $V_{1/2}$ puisque deux polynômes $P(x)$ et $Q(x)$ de degré $\leqslant m$ qui se raccordent en $1/2$ à l'ordre m sont nécessairement identiques. Il en résulte que la codimension de E dans F est égale à 1 et qu'il existe une fonction (unique à la multiplication près par une constante de module 1) $w(x) = w_m(x)$, appartenant à F, vérifiant $\|w\|_2 = 1$ et orthogonale à E (au sens du produit scalaire de l'espace de Hilbert de référence $L^2(\mathbb{R})$). *C'est cette fonction* $w(x)$ *qui est l'ondelette de Stromberg.* Par construction $w(x)$ appartient à $H^m(\mathbb{R})$ et, plus précisément, les dérivées de $w(x)$ d'ordre $\leqslant (m-1)$ sont continues, la dérivée m-ième est une fonction en escalier, mais il est beaucoup moins évident que $w(x)$ soit à décroissance exponentielle. Nous nous proposons d'établir cette dernière propriété en démontrant que la construction de $w(x)$ coincide aves les algorithmes généraux de construction d'ondelettes orthogonales développés par S. Mallat et l'auteur. Nous allons décrire ces algorithmes et les appliquer au cas qui nous occupe en privilégiant des solutions de «type causal» au lieu de rechercher, comme nous le faisions jusque là les solutions symétriques. Nous verrons alors apparaître les ondelettes $w_m(x)$ de Stromberg. Bien que les particularités géométriques de la construction de Stromberg ne

soient pas généralisables à d'autres cadres que celui des suites emboîtées d'espaces de fonctions splines, la forme analytique obtenue s'insère donc dans un contexte d'une plus grande portée.

3. Analyses Multirésolutions et Ondelettes

Une analyse multirésolution de l'espace de référence $H = L^2(\mathbb{R})$ est, par définition, une suite emboîtée $V_j, j \in \mathbb{Z}$, de sous-espaces fermés de H ayant les propriétés suivantes.

(3.1) $f(x) \in V_j$ si et seulement si $f(2x) \in V_{j+1}$
(3.2) pour tout $k \in \mathbb{Z}$, $f(x) \in V_0$ si et seulement si $f(x-k) \in V_0$
(3.3) l'intersection des V_j est réduite à $\{0\}$
(3.4) la réunion des V_j est dense dans $H = L^2$
(3.5) il existe une fonction $g(x)$ dans V_0 telle que $g(x-k)$, $k \in \mathbb{Z}$, soit une base de Riesz de H.

Une base de Riesz d'un espace de Hilbert H est, par définition, l'image d'une base orthonormée par un isomorphisme, non nécessairement isométrique. En d'autres termes, une base de Riesz, est une suite e_j, $j \in J$, ayant les deux propriétés suivantes

(3.6) il existe deux constantes C_1 et C_2, strictemente positives, telles que, pour toute suite (finie) α_j, $j \in J$, de coefficients scalaires, on ait

$$C_1(\Sigma|\alpha_j|^2)^{1/2} \leqslant \|\Sigma\alpha_j e_j\| \leqslant C_2(\Sigma|\alpha_j|^2)^{1/2}$$

(3.7) les combinaisons linéaires finies $\Sigma\alpha_j e_j$ sont denses dans H.

On dira qu'une analyse multirésolution de $L^2(\mathbb{R})$ est m-régulière si, en outre, on peut choisir $g(x)$ bornée, à décroissance rapide à l'infini, ainsi que toutes ses dérivées d'ordre $\leqslant m$.

La première opération que l'on effectue sur une analyse multirésolution m-réguliére est de remplace $g(x)$ par $q(x)$, ayant les mêmes propriétés de régularité et de décroissance à l'infini et de sorte que (3.5) devienne la condition que $q(x-k)$, $k \in \mathbb{Z}$, soit une base orthonormée de V_0. Pour ce faire, on ne peut employer le procédé d'orthonormalisation de Gram-Schmidt car il n'est pas compatible avec la structure particulière des bases que l'on cherche à construire. On le remplace par un algorithme différent que nous décrivons maintenant en situation générale. Si e_j, $j \in J$, est une base de Riesz d'un espace de Hilbert H, on considère la matrice de Gram associée qui est, par définition,

$$G = (\langle e_j, e_k \rangle)_{(j,k) \in J \times J}.$$

Cette matrice est définie-positive, précisément parce que e_j, $j \in J$, est une base de Riesz de H. On forme alors

$$G^{-1/2} = (\mu_{j,k})_{(j,k)\in J\times J}$$

et $\Sigma \mu_{j,k} e_k = f_j$ est la base orthonormée cherchée.

En revenant au cas particulier qui nous occupe, ce procédé d'orthonormalisation produit une suite de la forme $q(x-k)$, $k \in \mathbb{Z}$, lorsqu'on l'applique à une suite $g(x-k)$. Qui plus est, la transformée de Fourier

$$Q(u) = \int e^{-iux} q(x)\, dx$$

de $q(x)$ et celle, notée $G(u)$ de $g(x)$ sont reliées par

$$Q(u) = G(u)\left(\sum_{-\infty}^{\infty} |G(u+2k\pi)|^2\right)^{-1/2}$$

ce qui permet le calcul effectif de $q(x)$. En outre, on a

$$C_1 \leqslant \left(\sum_{-\infty}^{\infty} |G(u+2k\pi)|^2\right)^{1/2} \leqslant C_2$$

où C_1 et C_2 sont deux constantes positives. On montre sans difficulté que $\sum_{-\infty}^{\infty} |G(u+2k\pi)|^2$ est indéfiniment dérivable et que $q(x)$ a la même régularité et la même décroissance à l'infini que $g(x)$.

Se pose alors le problème de l'unicité de la fonction $q(x)$ telle que la suite $q(x-k)$ soit une base orthonormée de l'espace V_0. Si $s(x)$ est une autre solution, les transformées de Fourier de $q(x)$ et de $s(x)$ sont reliées par $Q(u) = S(u)\pi(u)$ où $\pi(u)$ est une fonction 2π-périodique et unimodulaire. Réciproquement cette relation fournit toujours une base orthonormée et enfin, si $q(x)$ et $s(x)$ sont toutes deux bornées et à décroissance rapide à l'infini, la fonction $\pi(u)$ est nécessairement C^∞. La réciproque est vraie, ce qui permet d'expliciter tous les choix possibles de $q(x)$. La fonction, $q(x)$ est «le père des ondelettes». Nous pouvons alors construire la fonction $w(x)$ telle que $w(x-k)$, $k \in \mathbb{Z}$, soit une base orthonormée de W_0. On se sert de l'inclusion $V_0 \subset V_1$ des analyses multirésolutions et l'on définit la fonction $m_0(u)$ par

$$Q(2u) = m_0(u)Q(u)$$

et la condition que $m_0(u)$ soit 2π-périodique. Alors $m_0(u)$ est indéfiniment dérivable et, pour construire $w(x)$, on procède comme suit. On forme

$$m_1(u) = e^{-iu}\overline{m_0(u+\pi)}$$

puis

$$W(u) = m_1(u/2)Q(u/2)$$

et enfin $w(x)$ est la transformée de Fourier inverse de $W(u)$. Dans ces conditions $w(x-k)$, $k \in \mathbb{Z}$, est une base orthonormée de W_0. Une fois de plus se pose le problème de l'unicité. Comme ci-dessus, les seuls autres choix sont donnés par la possibilité de multiplier $W(u)$ par une fonction indéfiniment dérivable et 2π-périodique. Les vérifications de ces assertions sont faciles ([5]). Nous venons de décrire la construction générale de toutes les ondelettes. Il reste à indiquer quels choix particuliers conduisent aux ondelettes de Stromberg, ce que nous ferons dans la section suivante.

4. Les Ondelettes de Stromberg

Pour produire les ondelettes de Stromberg, on fixe un entier $m \geqslant 1$ et l'on considère la suite emboîtee V_j, $j \in \mathbb{Z}$, des espaces de fonctions splines que nous avons déjà décrite dans la Section 2. On commence par désigner par $g(x) = g_m(x)$ le produit de convolution $c * c * \cdots * c$ ($m+1$ termes) où $c(x)$ est la fonction caractéristique de $[-1, 0]$. La transformée de Fourier de $g(x)$ est

$$G(u) = (e^{iu} - 1)^{m+1}/(iu)^{m+1}.$$

Ensuite, on cherche $q(x)$, portée par $]-\infty, 0]$, telle que $q(x-k)$, $k \in \mathbb{Z}$, soit une base orthonormée de V_0. A cet effet, on utilise l'identité

$$\sum_{-\infty}^{\infty} |G(u + 2k\pi)|^2 = P_m(\cos u)$$

où P_m est un polynôme de degré m. En outre P_m est positif sur $[-1, 1]$ et l'on a donc, grâce à un lemme dû à F. Riesz,

$$P_m(\cos u) = c_m |(1 + z_1 e^{iu})(1 + z_2 e^{iu}) \cdots (1 + z_m e^{iu})|^2 \tag{4.1}$$

où $|z_1| < 1, \ldots, |z_m| < 1$ et où c_m est une constante positive. En fait, un calcul plus explicite, dû à I. J. Schoenberg, montre que les z_j sont des réels positifs deux à deux distinctis que l'on notera $s_m < s_{m-1} < \cdots < s_1$.

On pose

$$A_m(u) = \sqrt{c_m}(1 + s_1 e^{iu}) \cdots (1 + s_m e^{iu})$$

et l'on définit la transformée de Fourier $Q(u)$ de $q(x)$ par

$$Q(u) = G(u)/A_m(u). \tag{4.2}$$

Il est immédiat de vérifier que $q(x)$ est portée par $]-\infty, 0]$ et a les propriétés annoncées. On suit ensuite la procédure générale du paragraphe précédent ce qui conduit à

$$W(u) = e^{-iu/2}\overline{m_0(u/2+\pi)}\,Q(u/2) \tag{4.3}$$

où $Q(2u) = m_0(u)Q(u)$. Nous nous proposons de démontrer le résultat suivant.

Théorème 1. *La fonction $w(x)$ dont la transformée de Fourier est donnée par* (4.3) *est, à la multiplication près par une constante unimodulaire, l'ondelette de Stromberg.*

La preuve de cet énoncé se ramène à trois vérifications. Par construction, $w(x-k)$, $k \in \mathbb{Z}$, est une base orthonormée de W_0. Nous voulons, en outre, montrer que $w(x)$ et orthogonale à l'espace X des fonctions de V_1 portées par $]-\infty, 0]$ et enfin vérifier que $w(x)$ appartient à $V_{1/2}$. La décroissance exponentielle de $w(x)$ sera essentiellement évidente.

Pour faire la première vérification, il convient d'utiliser le lemme suivant.

Lemme 2. *La suite $\sqrt{2}\,q(2x+k)$, $k \in \mathbb{N}$, est une base orthonormée de X.*

Il est évident que cette suite est orthonormée et il faut démontrer qu'elle est complète. On commence par faire un changement d'échelle évident et se ramener au cas où V_1 est remplacé par V_0 et l'on désigne par Y l'espace des fonctions de V_0 portées par $]-\infty, 0]$. Ensuite, il faut observer que $g(x+k)$, $k \in \mathbb{N}$, est une base de Riesz de Y.

Les transformées de Fourier des fonctions de Y sont donc de la forme $\mu(u)G(u)$ où

$$\mu(u) = a_0 + a_1 e^{iu} + \cdots + a_k e^{ikx} + \cdots$$

appartient à l'espace de Hardy H^2.

Or les transformées de Fourier des séries $\sum_0^\infty a_k q(x+k)$ sont les produits $\mu(u)G(u)/A_m(u)$ et la preuve se conclut par la remarque que $A_m(u)$ est un polynôme en e^{iu} et que $1/A_m(u)$ est holomorphe et bornée dans le disque unité.

Revenons à $w(x)$. On a successivement

$$Q(u) = c_m^{-1/2}(e^{iu}-1)^{m+1}(iu)^{-m-1}(1+s_1e^{iu})^{-1}\cdots(1+s_me^{iu})^{-1}$$

puis

$$m_0(u) = 2^{-m-1}(1+e^{iu})^{m+1}(1+s_1e^{iu})(1+s_1e^{2iu})^{-1} \cdots (1+s_me^{iu})(1+s_me^{2iu})^{-1}$$

et enfin

$$m_1(u) = \frac{2^{-m-1}e^{-iu}(1-e^{-iu})^{m+1}\prod_1^m (1-s_je^{-iu})}{\prod_1^m (1+s_je^{-2iu})}. \tag{4.4}$$

Il vient donc

$$w(x) = a_1 q(2x - 1) + \cdots + a_k q(2x - k) + \cdots \tag{4.5}$$

où les a_k sont à décroissance exponentielle. Cela entraîne évidemment l'orthogonalité avec X.

La vérification de l'appartenance à $T(V_{1/2})$ est un peu plus subtile. Il faut vérifier que $w(x + 1) = u(x) + v(x)$ où $u \in V_0$ et $v \in X$. Par transformée de Fourier, cela se ramène à

$$\begin{aligned} e^{iu} W(u) = (e^{iu/2} - 1)^{m+1} (iu/2)^{-m-1} A(u/2) \\ + (e^{iu} - 1)^{m+1} (iu)^{-m-1} B(u) \end{aligned} \tag{4.6}$$

où $A(u)$ est 2π-périodique et appartient à l'espace de Hardy H^2 et $B(u)$ est également 2π-périodique et de carré sommable. Après simplification, on est amené, on posant

$$f(u) = e^{iu}(1 - e^{-iu})^{m+1} \prod_1^m (1 - s_j e^{-iu})(1 + s_j e^{iu})^{-1}(1 + s_j e^{-2iu})^{-1} \tag{4.7}$$

à savoir si

$$f(u) = g(u) + (1 + e^{iu})^{m+1} h(2u) \tag{4.8}$$

où $g(u)$ est 2π-périodique et appartient à l'espace de Hardy H^2 et $h(u)$ est aussi 2π-périodique et de carré sommable. On utilise alors le lemme suivant

Lemme 3. *Pour une fonction 2π-périodique et de carré sommable $f(u)$,* (4.8) *est équivalent à*

$$r(u) = f(u)(1 - e^{iu})^{m+1} - f(u + \pi)(1 + e^{iu})^{m+1} \in H^2. \tag{4.9}$$

Dans un sens, c'est immédiat. Dans l'autre, on commence par observer que toute fonction $r(u)$ de H^2 s'écrit

$$r(u) = s(u)(1 - e^{iu})^{m+1} + t(u)(1 + e^{iu})^{m+1}$$

où $s(u)$ et $t(u)$ appartiennent toutes deux à H^2. Cette observation est une conséquence immédiate du théorème de Bezout. Mais on a $r(u + \pi) = -r(u)$ et il en résulte que $t(u) = -s(u + \pi)$. Finalement on obtient

$$\begin{aligned} [f(u) - s(u)](1 - e^{iu})^{m+1} &= [f(u + \pi) - s(u + \pi)](1 + e^{iu})^{m+1} \\ &= h(2u) \end{aligned}$$

où $h(u)$ a les propriétés annoncées.

Revenant au cas particulier qui nous occupe, on conclut grâce à l'identité remarquable

$$(4.10) \quad P_m(\cos 2u) = [\cos(u/2)]^{2m+2} P_m(\cos u) + [\sin(u/2)]^{2m+2} P_m(-\cos u).$$

L'identité (4.10) découle immédiatement de la définition des polynômes P_m.

5. Le Problème Général de la Construction des Ondelettes

Nous venons de démontrer que la construction faite par J. O. Stromberg rentre dans le cadre général des analyses multirésolutions et des formules explicites qui fournissent alors l'ondelette $w(x)$ à partir de la fonction $q(x)$ correspondante. Mais nous ne savons pas, à l'heure actuelle, s'il s'agit d'un fait général. Voici un énoncé plus précis. Soit $w(x)$ une ondelette d'ordre $m \geqslant 1$. Existe-t-il une analyse multirésolution m-régulière V_j, $j \in \mathbb{Z}$, telle que $w(x-k)$, $k \in \mathbb{Z}$, soit une base orthonormée du complément orthogonal W_0 de V_0 dans V_1? Nous ne savons pas répondre à cette question. Dans tous les exemples dont nous disposons, l'ondelette $w(x)$ provient d'une analyse multirésolution m-régulière.

Références

[1] Daubechies, I. The wavelet transform, time-frequency localization and signal analysis, AT&T Bell Labs., Murray Hill, USA.
[2] Daubechies, I. Orthonormal basis of compactly supported wavelets, AT&T Bell Labs., Murray Hill, USA.
[3] Mallat, S. Multiresolution approximation and wavelets orthonormal bases of $L^2(\mathbb{R})$, à paraître dans *Trans. Amer. Math. Soc.*
[4] Meyer, Y. Wavelets and Operators, *Proceedings of the special year in modern analysis*, Urbana 1986/87, to be published by Cambridge University Press, (1989).
[5] Meyer, Y. Ondelettes, fonctions splines et analyses graduées, Cahiers du CEREMADE, n.° 8703.
[6] Stromberg, J. O. A modified Fanklin system and higher order spline systems on $\mathbb{R}^n$ as unconditional bases for Hardy spaces. *Conference in Harmonic Analysis in Honor of Antoni Zygmund,* Vol. II, 475-493, edited by W. Beckner and al., Wadsworth Math. Series.

Yves Meyer
CEREMADE
Université Paris IX-Dauphine
Place de Lattre de Tassigny
75775 Paris Cedex 16, France

REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.º 1, 1988

The Boundedness of Calderón-Zygmund Operators on the Spaces $\dot{F}_p^{\alpha,q}$

M. Frazier, R. Torres and G. Weiss

To our dear friend José-Luis Rubio de Francia

Introduction

Calderón-Zygmund operators are generalizations of the singular integral operators introduced by Calderón and Zygmund in the fifties [CZ]. These singular integrals are principal value convolutions of the form

$$Tf(x) = \lim_{\epsilon \to 0} \int_{|x-y|>\epsilon} K(x-y) f(y)\, dy = \text{p.v.}\, K * f(x),$$

where f belongs to some class of test functions. The kernel K is a function defined in $\mathbb{R}^n \setminus \{0\}$ and has the form:

$$K(x) = \Omega(x/|x|)|x|^{-n}$$

where Ω is usually assumed to satisfy some «smoothness» property (see [S] or [GC-RF]) on the unit sphere S^{n-1} as well as the important cancellation property

$$\int_{S^{n-1}} \Omega(x')\, d\sigma = 0.$$

This cancellation property may be rewritten, at least formally, as $T1 = 0$; *i.e.*, the action of T on the function identically equal to 1 (properly defined) is identically zero.

The L^2 boundedness of these operators can be proved by using Fourier transform arguments and, from this, the L^p boundedness for $1 < p < \infty$ is obtained by what are now considered to be standard real variable techniques. Although Fourier transform arguments are not applicable, the L^p boundedness results are still true for some more general principal value integrals

$$Tf(x) = \lim_{\epsilon \to 0} \int_{|x-y|>\epsilon} K(x,y) f(y)\, dy$$

that are not convolutions.

The general operators in question can be described as follows. Let $T: \mathcal{D} \to \mathcal{D}'$ be a continuous linear operator, where $\mathcal{D} = \mathcal{D}(\mathbb{R}^n)$ is the space of Schwartz test functions and $\mathcal{D}' = \mathcal{D}'(\mathbb{R}^n)$ is its dual. Let K denote the distributional kernel of T. That is, $K \in \mathcal{D}'(\mathbb{R}^n \times \mathbb{R}^n)$ and satisfies:

$$\langle T\varphi, \psi \rangle = \langle K, \psi \otimes \varphi \rangle$$

for all φ and ψ in $\mathcal{D}$. We assume that the restriction of K to the set $\{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : x \neq y\}$ is a continuous function. Hence, if φ and ψ are in $\mathcal{D}$ and $\operatorname{supp} \varphi \cap \operatorname{supp} \psi = \varnothing$, then

$$\langle T\varphi, \psi \rangle = \iint K(x,y)\varphi(y)\psi(x)\, dy\, dx.$$

Moreover, we suppose that K satisfies the «size and smoothness» conditions:

$$|K(x,y)| \leqslant C|x-y|^{-n}, \quad \text{and} \tag{0.1}$$

$$|K(x,y) - K(x',y)| + |K(y,x) - K(y,x')| \leqslant C|x-x'|^{\epsilon}|x-y|^{-n-\epsilon} \tag{0.2}$$

whenever $2|x-x'| \leqslant |x-y|$, for some constant $C > 0$ and some ϵ in $(0, 1]$.

These operators were introduced by Coifman and Meyer in [CM] where they showed that if T is bounded on L^2 then it is also bounded on L^p for $1 < p < \infty$.

The problem of charaterizing which operators with this type of kernel are bounded on L^2 was solved by David and Journé in 1985 with the «$T1$-Theorem». They found two necessary and sufficient conditions that T must satisfy. The first of these conditions is that both $T1$ and T^*1 (where T^* denotes the formal transpose of T) must be in BMO. In fact, David and Journé proved in [DJ] that the above condition can be reduced to the case $T1 = T^*1 = 0$; thus, a cancellation property appears again.

The operators studied by Calderón and Zygmund are invariant under the action of the group of translation and dilations of $\mathbb{R}^n$ (the «$ax + b$ group»). However, this is not always the case for this more general class of operators. It is then necessary and natural to impose some constraints on the behaviour of T under the action of this group. The second condition in the David and

Journé result is the so called Weak Boundedness Property (WBP) and it is intimately related to these constraints. We will describe it in detail in the next section.

Several authors have obtained versions of the $T1$-Theorem in other spaces of functions or distributions (see for example [L], [MM], [M2]). In general, the results obtained are based on the same kind of conditions mentioned before. In addition, a higher degree of smoothness on K and a higher degree of cancellation on T (namely that T vanishes on polynomials up to certain order) is usually imposed. In [FHJW] Calderón-Zygmund operators acting on the Triebel-Lizorkin spaces $\dot{F}_p^{\alpha,q}$ were studied for $\alpha \in [0, 1)$ and $1 \leqslant p, q \leqslant \infty$. In this paper we extend the study to the general case $\alpha \in \mathbb{R}$ and $0 < p, q \leqslant \infty$. Our main results are Theorems 3.1, 3.7 and 3.13.

One of the reasons for studying Calderón-Zygmund operators in the context of Triebel-Lizorkin spaces in that the results obtained can be translated to some other classical spaces. For example, we recover the result of David and Journé for L^2, and, using interpolation arguments, a criterion for the boundedness of Calderón-Zygmund operators on the Besov-Lipschitz spaces $\dot{B}_p^{\alpha,q}$. Also, for $\alpha > 0$ and $1 < p < \infty$, the inhomogeneous Triebel-Lizorkin space $F_p^{\alpha,2}$ coincides with the Sobolev space L_p^α and a criterion for this space can be easily obtained from the homogeneous case (see Corollary 3.33 below).

The main results in this paper were proved independently in [FW] and [T]. They were motivated by problems posed by B. Jawerth and E. Stein. Also, we would like to thank G. David for some helpful discussions.

We have included some technicals facts that can, essentially, be found elsewhere in the literature. However, we introduce some different notation that can be used to simplify the presentation. Hence, we restate some of these known facts here. The reader is referred to the expository article [FW] and to [T] for further details.

1. Some Preliminary Definitions

Generalized Calderón-Zygmund operators

Frazier and Jawerth have shown in [FJ1] and [FJ2] that the spaces $\dot{F}_p^{\alpha,q}$ can be decomposed in terms of some building blocks called «smooth atoms», and similarly into some more general building blocks, the «smooth molecules». Using this, it follows, in the tradition of [CW], [TW] and others, that to prove that an operator is bounded on these spaces, it is sufficient to prove that it maps «smooth atoms» into «smooth molecules». This was the strategy used in [FHJW] and it is also the one that we want to apply here.

We start by recalling the definitions of these atoms and molecules. Fix $\alpha \in \mathbb{R}$, and $0 < p, q \leqslant \infty$. Let $J = n/\min\{1, p, q\}$ and $L = \max\{[J - n - \alpha], [J - n]\}$

where, for $x \in \mathbb{R}$, $x_+ = \max\{0, x\}$, $[x]$ denotes the standard «greatest integer function» at x, and $x^* = x - [x]$.

Let Q be a dyadic cube of $\mathbb{R}^n$ with «lower left corner» at the point x_Q and with side length $\ell(Q)$. A smooth atom for $\dot{F}_p^{\alpha,q}$ associated to the cube Q, is a function $a_Q \in \mathcal{D}$ satisfying

$$\text{Supp } a_Q \subset 3Q, \tag{1.1}$$

$$\int x^\gamma a_Q(x)\,dx = 0 \quad \text{if} \quad |\gamma| \leqslant L, \tag{1.2}$$

$$|D^\gamma a_Q(x)| \leqslant |Q|^{-1/2-|\gamma|/n} \quad \text{if} \quad |\gamma| \leqslant [\alpha]_+ + 1, \tag{1.3}$$

(where as usual, for a multi-index of non-negative integers $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n)$, we let $D^\gamma = (\partial^{\gamma_1}/\partial x_1^{\gamma_1})(\partial^{\gamma_2}/\partial x_2^{\gamma_2}) \cdots (\partial^{\gamma_n}/\partial x_n^{\gamma_n})$, $x^\gamma = x_1^{\gamma_1} x_2^{\gamma_2} \cdots x_n^{\gamma_n}$ and $|\gamma| = \gamma_1 + \gamma_2 + \cdots + \gamma_n$).

Let $M > J$ and $1 > \delta > \alpha^*$. A smooth (δ, M)-molecule is a function m_Q which satisfies

$$|m_Q(x)| \leqslant C|Q|^{-1/2}(1 + \ell(Q)^{-1}|x - x_Q|)^{-\max\{M, M-\alpha\}}, \tag{1.4}$$

$$\int x^\gamma m_Q(x)\,dx = 0 \quad \text{if} \quad |\gamma| \leqslant [J - n - \alpha], \tag{1.5}$$

$$|D^\gamma m_Q(x)| \leqslant C|Q|^{-1/2-|\gamma|/n}(1 + \ell(Q)^{-1}|x - x_Q|)^{-M} \quad \text{for} \quad |\gamma| \leqslant [\alpha], \tag{1.6}$$

and, for $|\gamma| = [\alpha]$,

$$|D^\gamma m_Q(x) - D^\gamma m_Q(x')| \leqslant C|Q|^{-1/2-(|\gamma|+\delta)/n}|x - x'|^\delta \sup_{|z| \leqslant |x-x'|} (1 + \ell(Q)^{-1}|z - (x - x_Q)|)^{-M}, \tag{1.7}$$

where C is an absolute constant. (Our convention is that conditions (1.5), (1.6) or (1.7) are void if the term bounding $|\gamma|$ is negative.)

The degree of smoothness of the molecules increases with the parameter α; thus, in order to carry out our project, we will need to consider derivatives of the image of an atom by the Calderón-Zygmund operators that we want to study. This, in turn, will cause us to consider operators with sufficiently smooth kernels.

The usual «$|x - y|^{-n}$ size condition» of the kernels of Calderón-Zygmund operators, in general, is not satisfied by the kernels of the derivatives of these operators. In fact, the behaviour of those near the «diagonal» is worse, and, consequently, we are forced to consider a more general class of operators that we will now define. Let $T: \mathcal{D} \to \mathcal{D}'$ be a continuous linear operator and let K be its distributional kernel. For $m, l \in \mathbb{N}_0$ and $\epsilon \in (0, 1)$, we say that T is a (generalized) Calderón-Zygmund operator of «size m» and «smoothness $l + \epsilon$» if the restriction of K to $\Omega = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : x \neq y\}$ is a function with

continuous partial derivatives in the variables x up to order l which satisfy, for some constant $C > 0$:

$$|D_1^\gamma K(x,y)| \leqslant C|x-y|^{-n-m-|\gamma|} \quad \text{for} \quad |\gamma| \leqslant l \tag{1.8}$$

$$|D_1^\gamma K(x,y) - D_1^\gamma K(x',y)| \leqslant C|x-y|^{-n-m-l-\epsilon}|x-x'|^\epsilon \tag{1.9}$$

for $|\gamma| = l$ and $2|x-x'| < |x-y|$(where the subindex 1 stands for derivatives in the first variable, x). In such a case we will write $T \in \mathrm{CZO}(m, l+\epsilon)$ or $T \in \mathrm{CZO}(l+\epsilon)$ if $m = 0$. If β and γ are two multi-indices, we denote by $T_{\beta,\gamma}$ the continuous linear operator from $\mathfrak{D}$ to $\mathfrak{D}'$ whose distributional kernel is:

$$K_{\beta,\gamma} = (y-x)^\beta D_1^\gamma K, \tag{1.10}$$

where the derivatives are taken in the distributional sense. It is easy to check that if $T \in \mathrm{CZO}(m, l+\epsilon)$, $|\gamma| \leqslant l$ and $|\beta| \leqslant m$, then $T_{\beta,\gamma} \in \mathrm{CZO}(m - |\beta| + |\gamma|, l - |\gamma| + \epsilon)$ and that the restriction of $K_{\beta,\gamma}$ to Ω is the function:

$$^{\circ}K_{\beta,\gamma}(x,y) = (y-x)^\beta D_1^\gamma K(x,y).$$

We will write T_β if $|\gamma| = 0$ and $D^\gamma T$ if $|\beta| = 0$.

In contrast with (0.2), our definition of Calderón-Zygmund operators does not impose any smoothness in the $y-$variables. Recall that the formal transpose of T, $T^*: \mathfrak{D} \to \mathfrak{D}'$, is defined by

$$\langle T^*\psi, \varphi\rangle = \langle T\varphi, \psi\rangle, \quad \text{for} \quad \varphi, \psi \in \mathfrak{D}, \tag{1.11}$$

and, hence, the kernel of T^* is given by $K^*(x,y) = K(y,x)$. We will require later that the kernel of T be smooth in both x and y; thus, both T and T^* will be generalized Calderón-Zygmund operators for appropriate numbers m and l.

The action of generalized Calderón-Zygmund operators on functions of polynomial growth

Another characteristic of the atoms and molecules for the spaces $\dot{F}_p^{\alpha,q}$ is that they have a certain number of vanishing moments (properties (1.2) and (1.5)). Since we want that the Calderón-Zygmund operators preserve these vanishing moments, some cancellation property should be imposed on them. As we mention before, this is usually achieved by requiring that T (and T^*) vanishes on polynomials. However, since T is originally defined only on $\mathfrak{D}$, its action on polynomials has to be first defined. More generally, assume that $T \in \mathrm{CZO}(m, \epsilon)$. We want to extend T to the space:

$$\mathcal{O}^m = \{f \in C^\infty(\mathbb{R}^n): f(x) = O(|x|^m) \text{ as } x \to \infty\}.$$

Let $\{\varphi_j\}_{j=1}^\infty$ be a sequence of functions in $\mathfrak{D}$ satisfying the following properties:

$$\operatorname{supp} \varphi_j \subseteq B_{j+1} = \{x \in \mathbb{R}^n : |x| < j+1\}, \tag{1.12}$$

$$\varphi_j(x) = 1 \quad \text{for every} \quad x \in B_j, \quad \text{and} \tag{1.13}$$

$$\|\varphi_j\|_\infty \leqslant C \quad \text{for some} \quad C > 0, \quad \text{for every} \quad j. \tag{1.14}$$

Let $f \in \mathcal{O}^m$. By analogy with the definition of the action of convolution singular integrals on L^∞ (see [GC-RF, p. 202]), we define Tf by the equality

$$Tf = \lim_{j \to \infty} \left\{ T(\varphi_j f) - \int_{1 < |y| < j} K(0, y) f(y)\, dy \right\}. \tag{1.15}$$

This is justified by the following

Lemma 1.16. *Suppose $T \in \mathrm{CZO}(m, \epsilon)$, $f \in \mathcal{O}^m$ and $\{\varphi_j\}_j$ is as before. The limit in* (1.15) *exists in the weak*- topology of $\mathcal{D}'$ and defines Tf as a distribution. In particular, if $g \in \mathcal{D}_0$ where $\mathcal{D}_0 = \left\{ \varphi \in \mathcal{D} : \int \varphi\, dx = 0 \right\}$, then*

$$\langle Tf, g \rangle = \lim_{j \to \infty} \langle T\varphi_j f, g \rangle.$$

PROOF. Let

$$F_j = T\varphi_j f - \int_{1 < |y| < j} K(0, y) f(y)\, dy.$$

We want to show that

$$\lim_{j \to \infty} \langle F_j, g \rangle \quad \text{exists for any} \quad g \in \mathcal{D}.$$

Fix $g \in \mathcal{D}$ and let j_0 be large enough so that $\operatorname{supp} g \subseteq B_{j_0/2}$. For $j > j_0$, we can write

$$F_j = F_{j_0} + T(\varphi_j - \varphi_{j_0}) f - \int_{j_0 \leq |y| < j} K(0, y) f(y)\, dy.$$

Then, since $(\varphi_j - \varphi_{j_0}) f$ and g have disjoint supports, we have

$$\begin{aligned}
\langle F_j, g \rangle &= \langle F_{j_0}, g \rangle + \iint K(x, y)[(\varphi_j - \varphi_{j_0}) f](y) g(x)\, dy\, dx \\
&\quad - \iint_{j_0 < |y| < j} K(0, y) f(y) g(x)\, dy\, dx \\
&= \langle F_{j_0}, g \rangle + \iint_{j_0 < |y| \leq j_0 + 1} K(x, y)[1 - \varphi_{j_0} f](y) g(x)\, dy\, dx \\
&\quad - \iint_{j_0 < |y| \leq j_0 + 1} K(0, y) f(y) g(x)\, dy\, dx \\
&\quad + \iint_{j_0 + 1 < |y| < j} [K(x, y) - K(0, y)] f(y) g(x)\, dy\, dx \\
&\quad + \iint_{j \leq |y| < j+1} K(x, y) \varphi_j f(y) g(x)\, dy\, dx,
\end{aligned} \tag{1.17}$$

where all integrals are absolutely convergent.

Now, the first three terms in (1.17) do not depend on j, and, using (1.8), they are easily seen to be bounded by $P(g)$ where P is some continuous seminorm in $\mathcal{S}(\mathbb{R}^n)$ which may depend on the support of g.

By (1.9) and the dominated convergence theorem, the fourth term tends to

$$\iint_{j_0+1<|y|} [K(x,y)-K(0,y)]f(y)g(x)\,dy\,dx \quad \text{as} \quad j\to\infty,$$

and it is bounded by $C\|g\|_{L^1}$ where C is a constant depending on f, the support of g and the constant in (1.8) and (1.9).

Finally, the fifth term is bounded by

$$C\int_{\operatorname{supp} g}\int_{j-j_0/2<|y|<j+1+j_0/2} \frac{|y|^m}{|y|^{n+m}}|g(x)|\,dy\,dx$$

$$\leqslant C\|g\|_{L^1}\log\frac{j+1+j_0/2}{j-j_0/2}\to 0 \quad \text{as} \quad j\to\infty. \quad \square$$

Observe that if $f\in\mathcal{D}$, the above definition of Tf and the original action of T on f differ by a constant. Hence, Tf should be viewed as a distribution on $\mathcal{D}_0$.

It is easy to see that definition of Tf does not depend on the choice of the sequence $\{\varphi_j\}_{j=1}^{\infty}$ and that, for $m=0$, it is equivalent to the one given in [DJ].

If $T\in \mathrm{CZO}(m,\epsilon)$ and $g\in\mathcal{D}_0$, we can use Lemma 1.16 to extend the action of T^*g to $\mathcal{O}^m$ by

$$\langle T^*g,f\rangle \equiv \lim_{j\to\infty}\langle T^*g,\varphi_j f\rangle = \lim_{j\to\infty}\langle T\varphi_j f,g\rangle = \langle Tf,g\rangle,$$

which is consistent with (1.11).

Note also that for $w\in\operatorname{supp} g$ and

$$y\in G=\{z\in\mathbb{R}^n\colon \operatorname{dist}(z,\operatorname{supp} g)>2\operatorname{diam}(\operatorname{supp} g)\},$$

$$\begin{aligned}\left|\int K(x,y)g(x)\,dx\right| &= \left|\int [K(x,y)-K(w,y)]g(x)\,dx\right| \\ &\leqslant \int_{\operatorname{supp} g} C|x-w|^{\epsilon}|y-w|^{-n-m-\epsilon}\|g\|_{\infty}\,dx \\ &\leqslant C[\operatorname{diam}(\operatorname{supp} g)]^{n+\epsilon}\|g\|_{\infty}|y-w|^{-n-m-\epsilon},\end{aligned} \tag{1.18}$$

where we used that g has integral zero. But, if $\operatorname{supp} f\subseteq G$, the function $f(y)\int K(x,y)g(x)\,dx$ is an integrable function of y, since $f\in\mathcal{O}^m$, and, therefore,

$$\begin{aligned}\langle T^*g,f\rangle &= \lim_{j\to\infty}\iint K(x,y)\varphi_j f(y)g(x)\,dy\,dx \\ &= \lim_{j\to\infty}\int \varphi_j f(y)\int_{\operatorname{supp} g} K(x,y)g(x)\,dx\,dy \\ &= \int f(y)\int K(x,y)g(x)\,dx\,dy.\end{aligned}$$

Hence, we see that «far away» from the support of g, the distribution $T*g$ agrees with the continuous function $\int K(x,y)g(x)\,dx$.

We have used the «size» of T to define its action on $\mathcal{O}^m$. A similar argument can be given which uses the «smoothness» of T. More precisely, let

$$\mathcal{D}_l = \left\{\varphi \in \mathcal{D}: \int x^\gamma \varphi(x)\,dx = 0 \quad \forall |\gamma| \leqslant l\right\}$$

and let $\mathcal{D}_l'$ be the dual space of $\mathcal{D}_l$ with respect to the topology inherited from $\mathcal{D}$. Then we have

Lemma 1.19. *Let $T \in \mathrm{CZO}(l+\epsilon)$, $f \in \mathcal{O}^l$ and let $\{\varphi_j\}_{j=1}^\infty$ be as before. Then the limit*

$$Tf = \lim_{j\to\infty}\left\{T\varphi_j f - \sum_{|\gamma|\leq l}\int_{1<|y|<j}\frac{D_1^\gamma K(0,y)}{\gamma!}f(y)\,dy\,x^\gamma\right\}$$

exists in the weak-topology of $\mathcal{D}_{l-1}'$. In particular, if $g \in \mathcal{D}_l$, then*

$$\langle Tf, g\rangle = \lim_{j\to\infty}\langle T\varphi_j f, g\rangle.$$

The proof of the lemma is similar to that of Lemma 1.16 and we leave it for the reader. Again, to avoid ambiguities, Tf should be regarded as an element of $\mathcal{D}_l'$.

If $T: \mathcal{D} \to \mathcal{D}'$ is linear and continuous, then it is easy to check that

$$(D^\gamma T)\varphi = D^\gamma(T\varphi) \tag{1.20}$$

in the sense of distributions. Namely, let $T \in \mathrm{CZO}(l+\epsilon)$ and $f \in \mathcal{O}^l$. If $g \in \mathcal{D}$, then $D^\gamma g \in \mathcal{D}_{l-1}$ for $|\gamma| = l$ and hence we may consider

$$\langle D^\gamma(Tf), g\rangle \equiv (-1)^{|\gamma|}\langle Tf, D^\gamma g\rangle$$

$$= (-1)^{|\gamma|}\lim_{j\to\infty}\left\{\langle T\varphi_j f, D^\gamma g\rangle - \iint_{1<|y|<j} D_1^\gamma K(0,y)f(y)\,dy\,\frac{x^\gamma}{\gamma!}D^\gamma g(x)\,dx\right\}$$

$$= \lim_{j\to\infty}\left\{\langle D^\gamma T\varphi_j f, g\rangle - \iint_{1<|y|<j} D_1^\gamma K(0,y)f(y)\,dy\,g(x)\,dx\right\},$$

as an easy integration by parts shows. However, since $T \in \mathrm{CZO}(l+\epsilon)$, $D^\gamma T \in \mathrm{CZO}(l,\epsilon)$ and, from the last expression and Lemma 1.19, we have

$$\langle D^\gamma(Tf), g\rangle = \langle (D^\gamma T)f, g\rangle$$

which is consistent with (1.20). As in the remarks following Lemma 1.16, we may define the action of $T*g$ on $\mathcal{O}^l$ for $g \in \mathcal{D}_l$ and verify that

$$\langle T^*g, f\rangle = \langle Tf, g\rangle.$$

Also, as in (1.18), this time using that $g \in \mathcal{D}_l$, we have for $w \in \operatorname{supp} g$ and y «far away» from $\operatorname{supp} g$

$$\begin{aligned}(1.21)\quad \left|\int K(x,y)g(x)\,dx\right| &= \left|\int\left[K(x,y) - \sum_{|\gamma|\le l}\frac{D_1^\gamma K(w,y)}{\gamma!}(x-w)^\gamma\right]g(x)\,dx\right| \\ &\le \int_{\operatorname{supp} g} C|x-w|^{l+\epsilon}|y-w|^{-n-l-\epsilon}\|g\|_\infty\,dx \\ &\le C[\operatorname{diam}(\operatorname{supp} g)]^{n+l+\epsilon}\|g\|_\infty|y-w|^{-n-l-\epsilon},\end{aligned}$$

and, therefore, as before, away from the support of g, the distribution T^*g agrees with the continuous fuction $\int K(x,y)g(x)\,dx$.

We want to point out that if $g \in \mathcal{D}_l$, more general sequences than $\{\varphi_j\}_{j=1}^\infty$ can be used to compute $\langle Tf, g\rangle$. In fact, let $\{\psi_j\}_{j=1}^\infty \subseteq \mathcal{D}$ satisfy the condition (extending (1.13))

(1.13)′ for each compact set $B \subset \mathbb{R}^n$ there exists $N > 0$ such that $\psi_j(x) = 1$ for all $x \in B$ and $j > N$,

as well as

$$(1.14)'\qquad \|\psi_j\|_\infty \le C \quad \text{for all} \quad j.$$

Then it is easy to show, using (1.16) (see [T]) that, as in (1.16),

$$\langle Tf, g\rangle = \lim_{j\to\infty}\langle T\psi_j f, g\rangle.$$

Moreover, the following result due to Y. Meyer holds ([M2], *cf.* also [FW]).

Lemma 1.22. *Let $T \in \mathrm{CZO}(l+\epsilon)$ and let $f \in \mathcal{D}'$. For $\varphi \in \mathcal{D}$ let $\varphi^t(x) = \varphi(x/t)$, $t > 0$. Then for all $g \in \mathcal{D}_l$*

$$\lim_{t\to\infty}\langle T\varphi^t f, g\rangle = \varphi(0)\langle Tf, g\rangle.$$

The Weak Boundedness Property

In [DJ] the following concept of Weak Bondedness Property was introduced.

For $\eta \in \mathcal{D}$, $z \in \mathbb{R}^n$ and $t > 0$, let

$$\eta^{z,t}(x) = \eta\left(\frac{x-z}{t}\right).$$

A linear and continuous operator $T\colon \mathcal{D} \to \mathcal{D}'$ is said to satisfy the Weak Boundedness Property if for every bounded subset $\mathcal{B}$ of $\mathcal{D}$ there exists a positive

constant $C = C(\mathcal{B})$ such that for all $\varphi, \psi \in \mathcal{B}$, all $z \in \mathbb{R}^n$ and $t > 0$,

$$|\langle T\varphi^{z,t}, \psi^{z,t}\rangle| \leqslant Ct^n. \tag{1.23}$$

If K is the kernel of T, then (1.23) may be rewritten as

$$|\langle K, \psi^{z,t} \otimes \varphi^{z,t}\rangle| \leqslant Ct^n. \tag{1.24}$$

Moreover, since the linear span of $\mathcal{D}(\mathbb{R}^n) \otimes \mathcal{D}(\mathbb{R}^n)$ is dense in $\mathcal{D}(\mathbb{R}^n \times \mathbb{R}^n)$, it follows that the above is equivalent to the condition

$$|\langle K, f^{z,t}\rangle| \leqslant Ct^n,$$

for every function f in a bounded subset $\tilde{\mathcal{B}}$ of $\mathcal{D}(\mathbb{R}^n \times \mathbb{R}^n)$, all of whose elements
have support in, say, a cube centered on the diagonal (where, $f^{z,t}(x, y)$ is of course $f((x - z)/t, (y - z)/t)$).

If we denote by $T^{z,t}$ the operator from $\mathcal{D}$ to $\mathcal{D}'$ whose kernel, $K^{z,t}$, is defined by

$$\langle K^{z,t}, f\rangle = \frac{1}{t^n}\langle K, f^{z,t}\rangle,$$

then the Weak Boudedness Property says that for all $f \in \tilde{\mathcal{B}}$

$$|\langle K^{z,t}, f\rangle| \leqslant C.$$

That is, the family of operators $\{T^{z,t}\}_{z\in\mathbb{R}^n, t>0}$ is in some sense uniformly bounded. These are the constraints imposed on the behavior of T under the action of the «$ax + b$ group» that we mentioned in the introduction.

In order to include in our discussion operators of «size m» we now extend the definition of the Weak Bondedness Property in the following way.

A linear and continuous operator $T: \mathcal{D} \to \mathcal{D}'$ is said to satisfy a Weak Boundedness Property of order m if for every bounded subset $\mathcal{B}$ of $\mathcal{D}(\mathbb{R}^n \times \mathbb{R}^n)$ whose elements have supports in a fixed cube centered on the diagonal, there exists a positive constant $C = C(\mathcal{B})$ such that for all $f \in \mathcal{B}$, all $z \in \mathbb{R}^n$, and $t > 0$,

$$|\langle K, f^{z,t}\rangle| \leqslant Ct^{n-m}. \tag{1.25}$$

In such a case we will write $T \in \mathrm{WBP}(m)$ or, simply, $T \in \mathrm{WBP}$ if $m = 0$.

We conclude this section with some remarks that are not hard to check. The reader is referred to [T] for the details.

(*a*) Using translation, the set $\mathcal{B}$ may be replaced by more general ones whose elements have all their derivatives uniformly bounded and are supported in cubes centered on the diagonal with fixed side length (but not with fixed center).

(*b*) If, in addition, the kernel of T satifies the size condition

$$|K(x,y)| \leqslant C|x-y|^{-n-m} \quad \text{for} \quad (x,y) \in \Omega,$$

then the cubes mentioned in (*a*) do not need to be centered on the diagonal.

(*c*) If $T \in \mathrm{WBP}(m)$, then $T_{\beta,\gamma} \in \mathrm{WBP}(m - |\beta| + |\gamma|)$ for all $|\beta| \leqslant m$ and for all $|\gamma| \geqslant 0$ (compare with Lemma 2.9 in [FW]).

2. L^∞-Estimates and Pointwise Definitions

In this section we collect some technical lemmas that will be used later. Some of the results that we are going to present are already known for the case $m = 0$ and $l = 0, 1$. However, even in these cases, the proofs that we give here are somewhat different from the ones found in the references that we shall give.

The following two lemmas are generalizations of what is sometimes known as «Meyer's lemma». For $m = 0$, Y. Meyer obtained these results in [M1]; see also [MM].

Lemma 2.1. *Let $T \in \mathrm{WBP}(m)$, and assume that its kernel satisfies*

$$|K(x,y)| \leqslant C|x-y|^{-n-m}$$

for (x,y) in Ω. If $f \in \mathfrak{D}(\mathbb{R}^n \times \mathbb{R}^n)$ and $D_1^\beta D_2^\gamma f(x,x) = 0$ for all $|\beta| + |\gamma| \leqslant m$, then

$$\langle K, f \rangle = \iint_{x \neq y} K(x,y) f(x,y)\, dx\, dy,$$

where the integral on the right is absolutely convergent.

PROOF. Let $\varphi \in \mathfrak{D}(\mathbb{R}^n)$ be such that $\operatorname{supp} \varphi \subseteq B_2(0)$ and $\varphi = 1$ on $B_1(0)$. Then, for $\epsilon > 0$ we may write

$$\langle K, f \rangle = \left\langle K, \left[1 - \varphi\left(\frac{x-y}{\epsilon}\right)\right] f \right\rangle + \left\langle K, \varphi\left(\frac{x-y}{\epsilon}\right) f \right\rangle. \tag{2.2}$$

Since K is locally integrable on Ω, the first term on the right hand side of (2.2) is

$$\iint K(x,y)\left(1 - \varphi\left(\frac{x-y}{\epsilon}\right)\right) f(x,y)\, dx\, dy,$$

and, by the Lebesgue dominated convergence theorem, tends to

$$\iint K(x,y)f(x,y)\,dx\,dy$$

as $\epsilon \to 0$. Hence, we must show that the second term tends to zero.

Choose a function $\psi \in \mathcal{D}(\mathbb{R}^n)$ with $\operatorname{supp}\psi \subseteq B_2(0)$ and so that

$$\sum_{k\in\mathbb{Z}^n} \psi(x-k) = 1.$$

Then, for $1 \geqslant \epsilon > 0$

$$f(x,y)\varphi\left(\frac{x-y}{\epsilon}\right) = f(x,y)\varphi\left(\frac{x-y}{\epsilon}\right)\sum_{k\in\mathbb{Z}^n}\psi\left(\frac{x}{\epsilon}-k\right).$$

Let $r>0$ be such that $|x|<r$ for $(x,y)\in\operatorname{supp} f$. Then

$$f(x,y)\varphi\left(\frac{x-y}{\epsilon}\right) = f(x,y)\varphi\left(\frac{x-y}{\epsilon}\right)\sum_{|k|<R}\psi\left(\frac{x}{\epsilon}-k\right),$$

where

$$R = 2 + \frac{r}{\epsilon}.$$

If we let

$$F_{k,\epsilon}(x,y) = \epsilon^{-1-m} f(\epsilon x + k, \epsilon y + k)\varphi(x-y)\psi\left(x + \frac{k}{\epsilon} - k\right),$$

then the family $\{F_{k,\epsilon}: k\in\mathbb{Z}^n \text{ and } 1\geqslant\epsilon>0\}$ is one of the sets described in remark (a) at the end of Section 1, and since $T\in \mathrm{WBP}(m)$, there exists a constant C such that

$$|\langle K, F^{z,t}_{k,\epsilon}\rangle| \leqslant Ct^{n-m}$$

for all $k\in\mathbb{Z}^n$, $1\geqslant\epsilon>0$, $z\in\mathbb{R}^n$ and $t>0$. We have used the hypothesis $D_1^\beta D_2^\gamma f(x,x)=0$ to obtain the appropriate cancellation of the powers of ϵ. In particular,

$$|\langle K, F^{k,\epsilon}_{k,\epsilon}\rangle| \leqslant C\epsilon^{n-m},$$

and, hence,

$$\begin{aligned}\left|\left\langle K, f(x,y)\varphi\left(\frac{x-y}{\epsilon}\right)\right\rangle\right| &\leqslant \sum_{|k|<R}\left|\left\langle K, f(x,y)\varphi\left(\frac{x-y}{\epsilon}\right)\psi\left(\frac{x}{\epsilon}-k\right)\right\rangle\right| \\ &\leqslant \epsilon^{1+m}\sum_{|k|<R}|\langle K, F^{k,\epsilon}_{k,\epsilon}\rangle| \\ &\leqslant C\epsilon^{1+m}R^n\epsilon^{n-m} \to 0 \quad \text{as} \quad \epsilon\to 0. \quad \square\end{aligned}$$

The next result gives a crucial L^∞-estimate, showing that T behaves well on smooth bump functions.

Lemma 2.3. *Let $T \in \mathrm{CZO}(m, \epsilon) \cap \mathrm{WBP}(m)$ and assume that $T_\beta 1 = 0$ for all $|\beta| \leqslant m$. Then T maps $\mathcal{D}(\mathbb{R}^n)$ into L^∞. Moreover, there exists a constant C, depending only on T, such that for any $w \in \mathbb{R}^n$, $t > 0$ and $\varphi \in \mathcal{D}$ with* $\operatorname{supp}\varphi \subseteq B_t(w)$ *we have*

$$\|T\varphi\|_\infty \leqslant C \sum_{|\gamma| \leqslant 1+m} t^{|\gamma|-m} \|D^\gamma \varphi\|_\infty. \tag{2.4}$$

PROOF. We will proceed by induction on m. Assume first $m = 0$. Let $\varphi \in \mathcal{D}$, $\operatorname{supp}\varphi \subseteq B_t(w)$.

(*a*) For $x \notin B_{2t}(w)$,

$$T\varphi(x) = \int_{B_t(w)} K(x, y)\varphi(y)\, dy,$$

and, hence,

$$|T\varphi(x)| \leqslant \int_{B_t(w)} C|x-y|^{-n} \|\varphi\|_\infty dy \leqslant C\|\varphi\|_\infty. \tag{2.5}$$

(*b*) Let $\chi \in \mathcal{D}$, $\chi(x) = 1$ on $B_4(0)$ and $\chi(x) = 0$ for $x \notin B_6(0)$, and write

$$\varphi(y) = [\varphi(y) - \varphi(x)]\chi^{w,t}(y) + \varphi(x)\chi^{w,t}(y) = f_1(x, y) + f_2(x, y).$$

Then, for $\psi \in \mathcal{D}$ satisfying $\operatorname{supp}\psi \subseteq B_{3t}(w)$, we have

$$\begin{aligned} |\langle T\varphi, \psi\rangle| &= |\langle K, \psi \otimes \varphi\rangle| \\ &\leqslant |\langle K, \psi(x) f_1(x, y)\rangle| + |\langle K, \psi(x) f_2(x, y)\rangle| \\ &= I + II. \end{aligned}$$

By Lemma 2.1,

$$I = \left| \iint K(x, y)\psi(x) f_1(x, y)\, dx\, dy \right| \tag{2.6}$$

$$\begin{aligned} &\leqslant \int_{B_{3t}(w)} \int_{B_{6t}(w)} C|x-y|^{-n} \sum_{|\gamma|=1} \|D^\gamma\varphi\|_\infty |x-y|\, dy\, |\psi(x)|\, dx \\ &\leqslant Ct \sum_{|\gamma|=1} \|D^\gamma\varphi\|_\infty \|\psi\|_1. \end{aligned}$$

In order to estimate II, it is sufficient to estimate

$$|\langle K, \psi \otimes \chi^{w,t}\rangle| = |\langle T\chi^{w,t}, \psi\rangle|.$$

Assume, first, that $\int \psi(x)\,dx = 0$. Then, since $T(1) = 0$, we have

$$0 = \langle T1, \psi\rangle = \langle T\chi^{w,t}, \psi\rangle + \langle T^*\psi, (1 - \chi^{w,t})\rangle$$

and, hence

$$\begin{aligned}(2.7)\quad |\langle T\chi^{w,t}, \psi\rangle| &= |\langle T^*\psi, (1 - \chi^{w,t})\rangle| \\ &= \left|\iint K(x,y)(1 - \chi^{w,t})(y)\psi(x)\,dx\,dy\right| \\ &\leqslant \int_{B_{3t}(w)} \left|(1 - \chi^{w,t})(y)\int_{cB_{6t}(w)} [K(x,y) - K(w,y)]\psi(x)\,dx\right| dy \\ &\leqslant C\int_{B_{3t}(w)}\int_{cB_{6t}(w)} t^{\epsilon}|w - y|^{-n-\epsilon}|\psi(x)|\,dy\,dx \\ &\leqslant Ct^{\epsilon}\|\psi\|_1 \int_{cB_{6t}(0)} |y|^{-n-\epsilon}\,dy \\ &\leqslant C\|\psi\|_1.\end{aligned}$$

Now consider a general $\psi \in \mathfrak{D}$, $\operatorname{supp}\psi \subseteq B_{3t}(w)$ and $\int \psi(x)\,dx = b$. Fix $\eta \in \mathfrak{D}$, $\operatorname{supp}\eta \subseteq B_1(0)$, $\eta \geqslant 0$ and $\int \eta(x)\,dx = 1$. Then $\psi - bt^{-n}\eta^{w,t}$ has support in $B_{3t}(w)$ and has integral zero. Hence, using (2.7) and that $T \in \mathrm{WBP}(0)$

$$\begin{aligned}(2.8)\quad |\langle T\chi^{w,t}, \psi\rangle| &\leqslant |\langle T\chi^{w,t}, \psi - bt^{-n}\eta^{w,t}\rangle| + |b|t^{-n}|\langle T\chi^{w,t}, \eta^{w,t}\rangle| \\ &\leqslant C\|\psi - bt^{-n}\eta^{w,t}\|_1 + \|\psi\|_1 t^{-n}|\langle T\chi^{w,t}, \eta^{w,t}\rangle| \\ &\leqslant C\|\psi\|_1 + C\|\psi\|_1 t^{-n}t^n \\ &\leqslant C\|\psi\|_1.\end{aligned}$$

From (2.6) and (2.8) we see that

$$\begin{aligned}(2.9)\quad |\langle T\varphi, \psi\rangle| &\leqslant Ct\sum_{|\gamma|=1} \|D^{\gamma}\varphi\|_{\infty}\|\psi\|_1 + C\|\varphi\|_{\infty}\|\psi\|_1 \\ &\leqslant C\sum_{|\gamma|\leqslant 1} t^{|\gamma|}\|D^{\gamma}\varphi\|_{\infty}\|\psi\|_1.\end{aligned}$$

Thus, (2.5) and (2.9) show that $T\varphi$ extends as an element of $(L^1(\mathbb{R}^n))^* = L^{\infty}$ with a norm bounded by the right hand side of (2.4). This ends the proof for $m = 0$.

Assume that the result is true whenever $T \in \mathrm{CZO}(m-1, \epsilon) \cap \mathrm{WBP}(m-1)$ and $T_{\beta}1 = 0$ for all $|\beta| \leqslant m - 1$, and suppose $T \in \mathrm{CZO}(m, \epsilon) \cap \mathrm{WBP}(m)$ and $T_{\beta}1 = 0$ for all $|\beta| \leqslant m$. Fix $\varphi \in \mathfrak{D}$, satisfying $\operatorname{supp}\varphi \subseteq B_t(w)$.

$(a)'$ For $x \notin B_{2t}(w)$

$$(2.10)\qquad |T\varphi(x)| \leqslant \int_{B_t(w)} C|x - y|^{-n-m}\|\varphi\|_{\infty}\,dx \leqslant Ct^{-m}\|\varphi\|_{\infty},$$

which is an estimate of the right order.

$(b)'$ Let χ be as in (b) and write

$$\varphi(y) = \left[\varphi(y) - \sum_{|\gamma|\le m} \frac{D^\gamma\varphi(x)}{\gamma!}(y-x)^\gamma\right]\chi^{w,t}(y)$$
$$+ \left[\varphi(x)\chi^{w,t}(y) + \sum_{1\le|\gamma|\le m} \frac{D^\gamma\varphi(x)}{\gamma!}(y-x)^\gamma\right]\chi^{w,t}(y)$$
$$= f_1(x,y) + f_2(x,y) + f_3(x,y).$$

Then for $\psi \in \mathcal{D}$, $\operatorname{supp}\psi \subseteq B_{3t}(w)$,

$$|\langle T\varphi, \psi\rangle| \le \sum_{i=1}^{3} |\langle K, \psi(x) f_i(x,y)\rangle| = I + II + III.$$

As in (b), by Lemma 2.1

$$(2.11)\quad I = \left|\iint K(x,y)\psi(x) f_1(x,y)\,dx\,dy\right|$$
$$\le \int_{B_{3t}(w)}\int_{B_{6t}(w)} C|x-y|^{-n-m} \sum_{|\gamma|=m+1} \|D^\gamma\varphi\|_\infty |x-y|^{m+1}\,dy\,|\psi(x)|\,dx$$
$$\le Ct \sum_{|\gamma|=m+1} \|D^\gamma\varphi\|_\infty \|\psi\|_1.$$

The same argument, this time using that $T \in \mathrm{WBP}(m)$, yields

$$(2.12)\qquad II \le Ct^{-m}\|\varphi\|_\infty\|\psi\|_1.$$

Finally, to estimate III we can apply the induction hypothesis to the operators T_γ for $1 \le |\gamma| \le m$ to obtain

$$(2.13)\quad |\langle K, \psi(x) f_3(x,y)\rangle| \le \sum_{1\le|\gamma|\le m} |\langle D^\gamma\varphi T_\gamma \chi^{w,t}, \psi\rangle| \frac{1}{\gamma!}$$
$$\le C \sum_{1\le|\gamma|\le m} \|D^\gamma\varphi\|_\infty \sum_{0\le|\beta|\le m-|\gamma|+1} t^{-m+|\gamma|+|\beta|}\|D^\beta\chi^{w,t}\|_\infty\|\psi\|_1$$
$$\le C \sum_{1\le|\gamma|\le m} \|D^\gamma\varphi\|_\infty \sum_{0\le|\beta|\le m-|\gamma|+1} t^{-m+|\gamma|}\|D^\beta\chi\|_\infty\|\psi\|_1$$
$$\le C \sum_{1\le|\gamma|\le m} t^{-m+|\gamma|}\|D^\gamma\varphi\|_\infty\|\psi\|_1.$$

Hence, (2.10), (2.11), (2.12) and (2.13) give (2.4). □

Corollary 2.14. *Let $T \in \mathrm{CZO}(l+\epsilon) \cap \mathrm{WBP}$ and suppose that $T(y^\gamma) = 0$ for all $|\gamma| \le l$. Then T maps $\mathcal{D}$ into L^∞. Moreover, there exists a constant C*

depending only on T such that, if $\varphi \in \mathcal{D}$ and $\operatorname{supp}\varphi \subseteq B_t(w)$, *then*

$$\|D^\gamma T\varphi\|_\infty \leqslant C \sum_{|\nu| \leqslant |\gamma|+1} t^{|\nu|-|\gamma|}\|D^\nu\varphi\|_\infty \quad \text{for every} \quad |\gamma| \leqslant l. \tag{2.15}$$

PROOF. Since $T \in \mathrm{CZO}(l+\epsilon) \cap \mathrm{WBP}$ we have

$$D^\gamma T \in \mathrm{CZO}(|\gamma|, l-|\gamma|+\epsilon) \cap \mathrm{WBP}(|\gamma|) \subseteq \mathrm{CZO}(|\gamma|, \epsilon) \cap \mathrm{WBP}(|\gamma|),$$

since $|\gamma| \leqslant l$. Hence in order to apply Lemma 2.3 we only need to check that $T_{\beta,\gamma}1 = 0$ for every $|\beta| \leqslant |\gamma|$.

Choose $\psi \in \mathcal{D}_0$ and let $\{\varphi_j\}_{j=1}^\infty$ be a sequence as in Lemma 1.16. Then,

$$\begin{aligned}
\langle T_{\beta,\gamma}1, \psi\rangle &= \lim_{j\to\infty} \langle T_{\beta,\gamma}\varphi_j, \psi\rangle \\
&= \lim_{j\to\infty} \langle (y-x)^\beta D_1^\gamma K, \psi \otimes \varphi_j\rangle \\
&= (-1)^{|\gamma|} \lim_{j\to\infty} \langle K, D_1^\gamma[(y-x)^\beta \psi(x)]\varphi_j(y)\rangle \\
&= (-1)^{|\gamma|} \lim_{j\to\infty} \sum_{\substack{0\leqslant\nu\leqslant\gamma \\ 0\leqslant\beta+\nu-\gamma}} C_\alpha \langle K, (y-x)^{\beta+\nu-\gamma} D^\nu\psi(x)\varphi_j(y)\rangle
\end{aligned} \tag{2.16}$$

where the inequalities under the above sum refer to each component of the multi-indexes. We claim that each term in the sum is zero. In fact, $\langle K, (y-x)^{\beta+\nu-\gamma}D^\nu\psi(x)\varphi_j(y)\rangle$ can be written as a sum of terms of the form $C_\mu\langle K, y^\mu x^{\beta+\nu-\gamma-\mu}D^\nu\psi(x)\varphi_j(y)\rangle$ with $\beta+\nu-\gamma-\mu \geqslant 0$. Since $\psi \in \mathcal{D}_0$, it follows that $D^\nu\psi \in \mathcal{D}_{|\nu|}$ and, hence, the function $x^{\beta+\nu-\gamma-\mu}D^\nu\psi$ has at least vanishing moments of all orders not exceeding $|\mu|$. Then by Lemma 1.19, we have

$$\begin{aligned}
\lim_{j\to\infty} \langle K, y^\mu x^{\beta+\nu-\gamma-\mu} D_1^\nu\psi(x)\varphi_j(y)\rangle &= \lim_{j\to\infty} \langle T\varphi_j y^\mu, x^{\beta+\nu-\gamma-\mu}D^\nu\psi\rangle \\
&= \langle Ty^\mu, x^{\beta+\nu-\gamma-\mu}D^\nu\psi\rangle = 0,
\end{aligned}$$

since $|\mu| \leqslant |\beta| + |\nu| - |\gamma| \leqslant |\gamma|$. □

Let $T \in \mathrm{CZO}(m, \epsilon) \cap \mathrm{WBP}(m)$ and assume $T_\beta 1 = 0$ for $|\beta| \leqslant m$. We have proved that for every $\varphi \in \mathcal{D}$, $T\varphi$ is an L^∞-function. We will now show, that $T\varphi(x)$ can be well defined at every point. In fact, we will show that $T\varphi$ agrees almost everywhere with a continuous function.

Let $t > 0$, $w \in \mathbb{R}^n$ and suppose $\xi \in \mathcal{D}$ satisfies $\xi = 1$ on $B_{2t}(w)$ and $\xi(x) = 0$ for $x \notin B_{4t}(w)$. Let $\eta = 1 - \xi$. As in the proof of Lemma 2.3, we have that for all $\psi \in \mathcal{D}_0$ with $\operatorname{supp}\psi \subseteq B_t(w)$,

$$0 = \langle T1, \psi\rangle = \langle T\xi, \psi\rangle + \langle T^*\psi, \eta\rangle;$$

thus,

$$\begin{aligned}\langle T\xi, \psi\rangle &= -\iint [K(x,y) - K(w,y)]\eta(y)\psi(x)\,dy\,dx \\ &= \iint [K(w,y) - K(x,y)]\eta(y)\,dy\,\psi(x)\,dx.\end{aligned}$$

Since both $T\xi$ and

$$\int [K(w,y) - K(x,y)]\eta(y)\,dy$$

are bounded functions on $B_t(w)$ and they agree, as distributions, on $\mathfrak{D}_0(B_t(w))$, we must have that for a.e. x in $B_t(w)$,

$$T\xi(x) = C_{w,t} + \int [K(w,y) - K(x,y)]\eta(y)\,dy, \tag{2.17}$$

where $C_{w,t}$ is some constant which may depend on w and t but not on $x \in B_t(w)$.

Since the right hand side of (2.17) is a continuous function, we can take it as the pointwise definition of $T\xi$ on $B_t(w)$. Note also that if $m \geqslant 1$ then the «faster decay of $K(x,y)$ at infinity» allows us to compute $T\xi(x)$ as

$$T\xi(x) = C'_{w,t} - \int K(x,y)\eta(y)\,dy. \tag{2.18}$$

In any case, if $x, x' \in B_t(w)$, then

$$T\xi(x) - T\xi(x') = \int [K(x',y) - K(x,y)]\eta(y)\,dy. \tag{2.19}$$

We now want to define $T\varphi(x)$ for any $\varphi \in \mathfrak{D}$. Assume again $T \in \mathrm{CZO}(m,\epsilon) \cap \mathrm{WBP}(m)$ and $T_\beta 1 = 0$ for $|\beta| \leqslant m$. Let $\varphi \in \mathfrak{D}$ and let $\psi \in \mathfrak{D}(B_t(w))$. For ξ and η as before,

$$\begin{aligned}\langle T\varphi, \psi\rangle &= \langle K, \psi \otimes \varphi\rangle \\ &= \left\langle K, \psi(x)\left[\varphi(y) - \sum_{|\gamma|\leqslant m} D^\gamma\varphi(x)\frac{(y-x)^\gamma}{\gamma!}\right]\xi(y)\right\rangle \\ &\quad + \left\langle K, \psi(x)\sum_{|\gamma|\leqslant m} D^\gamma\varphi(x)\frac{(y-x)^\gamma}{\gamma!}\,\xi(y)\right\rangle + \langle K, \psi(x)\varphi(y)\eta(y)\rangle \\ &= \iint K(x,y)\left[\varphi(y) - \sum_{|\gamma|\leqslant m} D^\gamma\varphi(x)\frac{(y-x)^\gamma}{\gamma!}\right]\xi(y)\psi(x)\,dy\,dx \\ &\quad + \sum_{|\gamma|\leqslant m}\left\langle \frac{D^\gamma\varphi}{\gamma!}\,T_\gamma\xi, \psi\right\rangle + \iint K(x,y)\varphi(y)\eta(y)\psi(x)\,dy\,dx\end{aligned}$$

(the first integral representation follows from Lemma 2.1 and the second one is true because $\operatorname{supp}\eta\varphi \cap \operatorname{supp}\psi = \varnothing$). Hence, for a.e. $x \in B_t(w)$,

$$T\varphi(x) = \int K(x,y)\left[\varphi(y) - \sum_{|\gamma|\le m} D^\gamma\varphi(x)\frac{(y-x)^\gamma}{\gamma!}\right]\xi(y)\,dy \tag{2.20}$$
$$+ \sum_{|\gamma|\le m}\left\{\frac{D^\gamma\varphi}{\gamma!}T_\gamma\xi\right\}(x)$$
$$+ \int K(x,y)\varphi(y)\eta(y)\,dy.$$

Again, since the right hand side of (2.20) is continuous we can take it as the pointwise definition of $T\varphi$ on $B_t(w)$. The continuity also guarantees that this definition is independent of $B_t(w)$ and ξ.

It will be convenient for later applications to have a formula for the difference between the values of $T\varphi$ (as well as $D^\gamma T\varphi$) at two different points. This is obtained in the next lemma. Once again, for $l = 0$ the result is due to Y. Meyer and for $l = 1$ to M. Meyer. The proof follows from a careful use of the previous pointwise definitions and it is left to the reader. The details can be found in [T]. See also [FW] for an alternative formula.

Lemma 2.21. *Assume* $T \in \mathrm{CZO}(l+\epsilon) \cap \mathrm{WBP}$ *and* $T(x^\gamma) = 0$ *for all* $|\gamma| \le l$. *Let* $x \ne x'$ *be two points of* $\mathbb{R}^n$, $t = |x - x'|$, $\xi \in \mathfrak{D}$, $\xi = 1$ *on* $B_{2t}(x')$, $\xi(x) = 0$ *for* $x \notin B_{4t}(x')$ *and let* $\eta = 1 - \xi$. *Then for all* $\varphi \in \mathfrak{D}$ *and for all* $|\gamma| \le l$

$$D^\gamma T\varphi(x) - D^\gamma T\varphi(x')$$
$$= \int D_1^\gamma K(x,y)\left[\varphi(y) - \sum_{|\beta|\le|\gamma|}\frac{D^\beta\varphi(x)}{\beta!}(y-x)^\beta\right]\xi(y)\,dy$$
$$- \int D_1^\gamma K(x',y)\left[\varphi(y) - \sum_{|\beta|\le|\gamma|}\frac{D^\beta\varphi(x')}{\beta!}(y-x')^\beta\right]\xi(y)\,dy$$
$$+ \int [D_1^\gamma K(x,y) - D_1^\gamma K(x',y)]\left[\varphi(y) - \sum_{|\beta|\le|\gamma|}\frac{D^\beta\varphi(x')}{\beta!}(y-x')^\beta\right]\eta(y)\,dy$$
$$+ \sum_{|\beta|\le|\gamma|}\frac{1}{\beta!}\left[D^\beta\varphi(x) - \sum_{|\nu|\le|\gamma|-|\beta|}\frac{D^{\beta+\nu}\varphi(x')}{\nu!}(x-x')^\nu\right]T_{\beta,\gamma}\xi(x),$$

where all integrals are absolutely convergent.

3. Boundedness criteria

The following three theorems extend the results of [FHJW].

Theorem 3.1. *Let* $\alpha < 0$ *and* $0 < p, q \leqslant \infty$. *Assume that* $T: \mathcal{D} \to \mathcal{D}'$ *satisfies*

(3.2) $T \in \mathrm{WBP}$,

(3.3) $T \in \mathrm{CZO}(\delta)$ *where* $1 > \delta > 0$,

(3.4) $T^* \in \mathrm{CZO}([J - n - \alpha] + \rho)$ *where* $1 > \rho > (J - \alpha)^*$,

(3.5) $T1 = 0$, *and*

(3.6) $T^*(x^\gamma) = 0$ *for* $|\gamma| \leqslant [J - n - \alpha]$,

then T *extends to a bounded operator on* $\dot{F}_p^{\alpha,q}$.

Theorem 3.7. *Let* $\alpha \geqslant 0$ *and* $\min\{p, q\} \geqslant 1$. *Assume that* $T: \mathcal{D} \to \mathcal{D}'$ *satisfies*

(3.8) $T \in \mathrm{WBP}$,

(3.9) $T \in \mathrm{CZO}([\alpha] + \delta)$ *where* $1 > \delta > \alpha^*$,

(3.10) $T^* \in \mathrm{CZO}(\rho)$ *where* $1 > \rho > 0$,

(3.11) $T(y^\gamma) = 0$ *for all* $|\gamma| \leqslant [\alpha]$, *and*

(3.12) $T^*1 = 0$ *if* $\alpha = 0$,

then T *extends to a bounded operator on* $\dot{F}_p^{\alpha,q}$.

Theorem 3.13. *Let* $\alpha \geqslant 0$ *and* $\min\{p, q\} < 1$. *Assume that* $T: \mathcal{D} \to \mathcal{D}'$ *satisfies*

(3.14) $T \in \mathrm{WBP}$,

(3.15) $T \in \mathrm{CZO}([\alpha] + \delta)$ *where* $1 > \delta > \alpha^*$ *if* $J - n - \alpha < 0$ *and* $1 > \delta > \max\{\alpha^*, J^*\}$ *if* $J - n - \alpha \geqslant 0$,

(3.16) $T^* \in \mathrm{CZO}([J - n] + \rho)$ *where* $1 > \rho > J^*$,

(3.17) $T(y^\gamma) = 0$ *for all* $|\gamma| \leqslant [\alpha]$,

(3.18) $T^*(x^\gamma) = 0$ *for all* $|\gamma| \leqslant [J - n - \alpha]$ *if* $J - n - \alpha \geqslant 0$,

(3.19) $|D_2^\beta D_1^\gamma K(x, y) - D_2^\beta D_1^\gamma K(x, z)| \leqslant C|y - z|^\rho |x - y|^{-(n + |\beta| + |\gamma| + \rho)}$ *for* $2|y - z| \leqslant |x - y|$ *if* $|\gamma| \leqslant \min\{[J] - n, [\alpha]\}$ *and* $|\beta + \gamma| = [J] - n$, *and*

(3.20) $|D_2^\beta D_1^\gamma K(x, y) - D_2^\beta D_1^\gamma K(z, y)| \leqslant C|x - z|^\rho |x - y|^{-(n + |\beta| + |\gamma| + \delta)}$ *for* $2|x - z| \leqslant |x - y|$ *if* $|\gamma| = [\alpha]$ *and* $|\beta| = ([J] - n - [\alpha])_+$,

then T *extends to a bounded operator on* $\dot{F}_p^{\alpha,q}$.

As we mentioned in Section 1 we will prove these theorems by showing that smooth atoms are mapped by T into smooth molecules. Actually, we will show that if $a \in \mathcal{D}$ is a smooth atom associated with Q_0, the unit cube with lower

left corner at the origin, then Ta is a smooth molecule associated with Q_0. Moreover, we will see that Ta satisfies conditions (1.4) to (1.7) with a constant C depending only on the dimension n, the constants appearing in Lemma 2.3 and Corollary 2.14, and the constants associated to T and T^* by the hypotheses of the theorems and the definitions of generalized Calderón-Zygmund operators and WBP. It is not hard to check that all these constants, as well as the hypotheses of the theorems, remain unchanged if we replace T by any of the operators $T^{z,t}$. Now using a simple translation and dilation argument, together with the particular form of conditions (1.4) to (1.7), one can easily conclude that T maps a smooth atom associated with an arbitrary cube Q to a smooth molecule associated with Q and with the same constant C obtained for Q_0.

PROOF OF THEOREM 3.1. Let $\epsilon = \rho - (J - \alpha)^*$ and $M = J + \epsilon$. Let $a \in \mathfrak{D}$ be a smooth atom for $\dot{F}_p^{\alpha,q}$ associated with Q_0. We will show that Ta is a smooth M-molecule. Observe that since $\alpha < 0$ we only need to check that conditions (1.4) and (1.5) hold (this is also the reason why we use the term «smooth M-molecule» and not «smooth (δ, M)-molecule»).

Assume first $|x| > 4\sqrt{n}$. Then, since $a(x)$ has L-vanishing moments (where $L = [J - n - \alpha]$) and $\|a\|_\infty \leqslant 1$, we have

$$
\begin{aligned}
|Ta(x)| &= \left| \int_{3Q_0} K(x,y)a(y)\,dy \right| \\
&= \left| \int_{3Q_0} \left[K(x,y) - \sum_{|\beta| \leqslant L} D_2^\beta K(x,0) \frac{y^\beta}{\beta!} \right] a(y)\,dy \right| \\
&\leqslant \int_{3Q_0} \left| K(x,y) - \sum_{|\beta| \leqslant L} D_2^\beta K(x,0) \frac{y^\beta}{\beta!} \right| dy \\
&\leqslant \int_{3Q_0} \left| \sum_{|\beta| = L} [D_2^\beta K(x, z(y)) - D_2^\beta K(x,0)] \frac{y^\beta}{\beta!} \right| dy
\end{aligned}
$$

where the point $z(y) \in Q_0$. But, then, $2|z(y)| \leqslant 4\sqrt{n} < |x|$ and since $T^* \in \mathrm{CZO}(L + \rho)$, we obtain from the last inequality that

$$
|Ta(x)| \leqslant \int_{3Q_0} C \frac{|z(y)|^\rho}{|x|^{n+L+\rho}} |y|^L\,dy \leqslant C|x|^{-n-L-\rho} = C|x|^{-M+\alpha}, \tag{3.21}
$$

because $n + L + \rho = n + [J - n - \alpha] + (J - \alpha)^* + \epsilon = J - \alpha + \epsilon = M - \alpha$, where the constant C depends only on T and the dimension n. Clearly (3.21) gives (1.4) for Q_0 and $|x| \geqslant 4\sqrt{n}$.

Now let $|x| \leqslant 4\sqrt{n}$. We can use (3.2), (3.3) and (3.5) to apply Lemma 2.3 and obtain

$$|Ta(x)| \leqslant C(\|a\|_\infty + 2\sqrt{n}\,\|\nabla a\|_\infty) \leqslant C, \tag{3.22}$$

where the constant C, by (1.3), is independent of $a(x)$. Hence, (3.22) gives (1.4) for $|x| \leqslant 4\sqrt{n}$.

Since (1.4) holds, $x^\gamma Ta$ is integrable for $|\gamma| \leqslant L$; hence, by (3.4) and (3.6),

$$0 = \langle T^* x^\gamma, a\rangle = \langle Ta, x^\gamma\rangle = \int Ta(x)x^\gamma\,dx,$$

which is condition (1.5). □

PROOF OF THEOREM 3.7. Let $\epsilon = \min\{\delta - \alpha^*, \rho\}$ and let $M = J + \epsilon = n + \epsilon$. We will show that if a is a smooth atom associated with Q_0, then Ta is smooth (δ, M)-molecule also associated with Q_0.

We use the same arguments of the previous proof, and, when $|x| > 4\sqrt{n}$ we use (3.10), in order to obtain

$$\begin{aligned} |Ta(x)| &= \left|\int_{3Q_0} [K(x,y) - K(x,0)]a(y)\,dy\right| \\ &\leqslant C\int_{3Q_0} |y|^\rho |x|^{-n-\rho}\,dy \\ &\leqslant C|x|^{-n-\rho} \\ &\leqslant C|x|^{-M} \\ &\leqslant C(1+|x|)^{-M}, \end{aligned}$$

which is (1.4).

As in the previous proof, if $|x| \leqslant 4\sqrt{n}$ (3.8), (3.9), (3.11) and Lemma 2.3 yield

$$|Ta(x)| \leqslant C \leqslant C(1+|x|)^{-M}.$$

Condition (1.5) needs to be checked only if $\alpha = 0$. But in that case $T^*1 = 0$ and so

$$0 = \langle T^*1, a\rangle = \langle Ta, 1\rangle = \int Ta(x)\,dx.$$

We turn now to condition (1.6). If $|\gamma| = 0$, (1.6) is just (1.4). Thus, we may assume $\alpha \geqslant |\gamma| > 0$. Again there are two different cases. If $|x| > 4\sqrt{n}$ we have the immediate estimate

$$\begin{aligned} |D^\gamma Ta(x)| &\leqslant \int_{3Q_0} |D_1^\gamma K(x,y)|\,\|a\|_\infty\,dy \\ &\leqslant C|x|^{-n-|\gamma|} \\ &\leqslant C|x|^{-n-\delta} \\ &\leqslant C(1+|x|)^{-M}, \end{aligned}$$

while if $|x| \leqslant 4\sqrt{n}$, we may use (3.8), (3.9) and (3.11) to apply, this time, Corollary 2.14 and obtain

$$|D^\gamma Ta(x)| \leqslant C \sum_{|\beta| \leq |\gamma|+1} (2\sqrt{n})^{|\beta|-|\gamma|} \|D^\beta a\|_\infty \leqslant C \leqslant C(1+|x|)^{-M},$$

where the constant C is, again by (1.3), independent of $a(x)$.

We now show that Ta satisfies condition (1.7). Let $|\gamma| = [\alpha]$ and let x and x' be two different points of $\mathbb{R}^n$. If $|x-x'| \geqslant 1$ then, by (1.6), for $|\gamma| = [\alpha]$,

$$\begin{aligned} |D^\gamma Ta(x) - D^\gamma Ta(x')| &\leqslant |D^\gamma Ta(x)| + |D^\gamma Ta(x')| \\ &\leqslant C\{(1+|x|)^{-M} + (1+|x'|)^{-M}\} \\ &\leqslant C|x-x'|^\delta\{(1+|x|)^{-M} + (1+|x'|)^{-M}\}, \end{aligned}$$

which certainly gives (1.7) for this case.

Assume now $|x-x'| < 1$. We consider the following different positibilies:

(*a*) If $|x|, |x'| > 6\sqrt{n}$, then for $y \in 3Q_0$, $|x-y| \geqslant 4\sqrt{n} \geqslant 2|x-x'|$, and, also, $|x-y| \geqslant |x| - |y| \geqslant |x|/2$; thus,

$$\begin{aligned} |D^\gamma Ta(x) - D^\gamma Ta(x')| &= \left| \int_{3Q_0} D^\gamma K(x,y)a(y)\,dy - \int_{3Q_0} D^\gamma K(x',y)a(y)\,dy \right| \\ &\leqslant \int_{3Q_0} |D^\gamma K(x,y) - D^\gamma K(x',y)|\,\|a\|_\infty\,dy \\ &\leqslant \int_{3Q_0} C|x-x'|^\delta |x-y|^{-n-[\alpha]-\delta}\,dy \\ &\leqslant C|x-x'|^\delta |x|^{-n-[\alpha]-\delta} \\ &\leqslant C|x-x'|^\delta (1+|x|)^{-M}. \end{aligned}$$

(*b*) If $|x| \geqslant 6\sqrt{n}$ and $|x'| < 6\sqrt{n}$ we still have that x' is «far» from Q_0, that is $|x'| \geqslant |x| - |x-x'| \geqslant 6\sqrt{n} - 1 \geqslant 5\sqrt{n}$ and the situation can be handled as in (*a*). The same is true, of course, if $|x'| \geqslant 6\sqrt{n}$ and $|x| < 6\sqrt{n}$.

(*c*) The last case to be considered is when $|x|, |x'| < 6\sqrt{n}$. In this case we use the representation formula of Lemma 2.21. In fact, fix $\xi \in \mathfrak{D}$ satisfying $\operatorname{supp} \xi \subseteq B_4(0)$ and $\xi \equiv 1$ on $B_2(0)$, and let

$$\xi^{x',t}(y) = \xi\left(\frac{y-x'}{t}\right)$$

where $t = |x-x'|$. Then since $T \in \mathrm{CZO}([\alpha]+\delta) \cap \mathrm{WBP}$ and $T^\gamma(x^\gamma) = 0$, for $|\gamma| \leqslant [\alpha]$, we have for $|\gamma| = [\alpha]$

$$\begin{aligned} &|D^\gamma Ta(x) - D^\gamma Ta(x')| \\ &\qquad \leqslant \left| \int D_1^\gamma K(x,y) \left[a(y) - \sum_{|\beta| \leq [\alpha]} \frac{D^\beta a(x)}{\beta!} (y-x)^\beta \right] \xi^{x',t}(y)\,dy \right| \end{aligned}$$

$$+\left|\int D_1^\gamma K(x',y)\left[a(y)-\sum_{|\beta|\le[\alpha]}\frac{D^\beta a(x')}{\beta!}(y-x')^\beta\right]\xi^{x',t}(y)\,dy\right|$$

$$+\left|\int [D_1^\gamma K(x,y)-D_1^\gamma K(x',y)]\right.$$
$$\left.\left[a(y)-\sum_{|\beta|\le[\alpha]}\frac{D^\beta a(x')}{\beta!}(y-x')^\beta\right][1-\xi^{x',t}(y)]\,dy\right|$$

$$+\sum_{|\beta|\le[\alpha]}\frac{1}{\beta!}\left|D^\beta a(x)-\sum_{|\nu|\le[\alpha]-|\beta|}\frac{D^{\beta+\nu}a(x')}{\nu!}(x-x')^\nu\right||T_{\beta,\gamma}\xi^{x',t}(x)|$$

$$=I+II+III+IV.$$

We have

$$I\le C\int_{|x'-y|<4|x-x'|}|x-y|^{-n-[\alpha]}\left|a(y)-\sum_{|\beta|\le[\alpha]}\frac{D^\beta a(x)}{\beta!}(y-x)^\beta\right|\|\xi\|_\infty\,dy$$

$$\le C\int_{|x'-y|<4|x-x'|}|x-y|^{-n-[\alpha]}\sum_{|\beta|=[\alpha]+1}\|D^\beta a\|_\infty|y-x|^{[\alpha]+1}\,dy$$

$$\le C\int_{|x'-y|<4|x-x'|}|x-y|^{-n+1}\,dy$$

$$\le C|x-x'|$$

$$\le C|x-x'|^\delta,$$

since $|x-x'|<1$, where the constant C depends on ξ but not on $a(x)$ since $\|D^\beta a\|_\infty\le 1$ for all $|\beta|\le[\alpha]+1$. The term II can be estimated in the same way.

By (3.9),

$$III\le C\int_{|x'-y|>2|x-x'|}|x-x'|^\delta|x-y|^{-n-[\alpha]-\delta}$$
$$\left|a(y)-\sum_{|\beta|\le[\alpha]}\frac{D^\beta a(x')}{\beta!}(y-x')^\beta\right|dy$$

$$\le C|x-x'|^\delta\left\{\int_{|x'-y|<2}|x'-y|^{-n-[\alpha]-\delta}\sum_{|\beta|=[\alpha]+1}\|D^\beta a\|_\infty|y-x'|^{[\alpha]+1}\,dy\right.$$
$$\left.+\int_{|x'-y|\ge 2}2|x'-y|^{-n-[\alpha]-\delta}\sum_{|\beta|=[\alpha]}\|D^\beta a\|_\infty|y-x'|^{[\alpha]}\,dy\right\}$$

$$\le C|x-x'|^\delta\left\{\int_{|x'-y|<2}|x'-y|^{-n+1-\delta}\,dy+\int_{|x'-y|\ge 2}|x'-y|^{-n-\delta}\,dy\right\}$$

$$\le C|x-x'|^\delta.$$

Finally applying Lemma 2.3 to the operators $T_{\beta,\gamma}$ we obtain

$$
\begin{aligned}
IV &\leqslant \sum_{|\beta| \leq [\alpha]} \frac{1}{\beta!} \left\{ \left| D^\beta a(x) - \sum_{|\nu| \leq [\alpha] - |\beta|} \frac{D^{\beta+\nu} a(x')}{\nu!} (x - x')^\nu \right| \right. \\
&\qquad \left. C \sum_{|\mu| \leq [\alpha] - |\beta| + 1} (2t)^{|\mu| - [\alpha] + [\beta]} \| D^\mu \xi^{x', t} \|_\infty \right\} \\
&\leqslant C \sum_{|\beta| \leq [\alpha]} \sum_{|\nu| = [\alpha] - [\beta] + 1} \| D^\nu a \|_\infty |x - x'|^{[\alpha] - [\beta] + 1} \\
&\qquad \sum_{|\mu| \leq [\alpha] - |\beta| + 1} t^{|\mu| - [\alpha] + [\beta]} t^{-|\mu|} \| D^\mu \xi \|_\infty \\
&\leqslant C \sum_{|\beta| \leq [\alpha]} |x - x'|^{[\alpha] - [\beta] + 1} t^{-[\alpha] + [\beta]} \\
&\leqslant C|x - x'| \\
&\leqslant C|x - x'|^\delta.
\end{aligned}
$$

We have shown that $|D^\gamma Ta(x) - D^\gamma Ta(x')| \leqslant |x - x'|^\delta$, which implies (1.7) for $|x - x'| < 1$ and $|x|, |x'| < 6\sqrt{n}$. This concludes the proof of Theorem 3.7. □

PROOF OF THEOREM 3.13. Let $\epsilon = \min\{\delta - \alpha^*, \rho - J^*\}$ if $J - n - \alpha < 0$ and $\epsilon = \min\{\delta - \max\{\alpha^*, J^*\}, \rho - J^*\}$ if $J - n - \alpha \geqslant 0$, let $M = J + \epsilon$, and let $a(x)$ be a smooth atom for $\dot{F}_p^{\alpha,q}$ associated to Q_0. For these choices of M, the same arguments used in the proof of Theorem 3.1 imply that Ta satisfies conditions (1.4) and (1.5) for a smooth (δ, M)-molecule associated with Q_0, and so they will not be repeated here. In order to prove that Ta also satisfies conditions (1.6) and (1.7) we will use the mixed derivatives hypotheses (3.19) and (3.20).

Let $|x| > 4\sqrt{n}$. Assume, first, that $0 < |\gamma| \leqslant \min\{[J] - n, [\alpha]\}$. Then by (3.19)

$$
\begin{aligned}
(3.23) \quad |D^\gamma Ta(x)| &\leqslant \int_{3Q_0} \left| D_1^\gamma K(x, y) - \sum_{|\beta| \leq [J] - n - |\gamma|} D_2^\beta D_1^\gamma K(x, 0) \frac{y^\beta}{\beta!} \right| \| a \|_\infty dy \\
&\leqslant \int_{3Q_0} \sum_{|\beta| = [J] - n - |\gamma|} |D_2^\beta D_1^\gamma K(x, z(y)) - D_2^\beta D_1^\gamma K(x, 0)| \left| \frac{y^\beta}{\beta!} \right| dy \\
&\leqslant \int_{3Q_0} C|z(y)|^\rho |x|^{-n - ([J] - n - |\gamma|) - |\gamma| - \rho} |y|^{[J] - n - |\gamma|} dy \\
&\leqslant C|x|^{-[J] - \rho} \\
&= C|x|^{-J - (\rho - J^*)} \\
&\leqslant C|x|^{-J - \epsilon} = C|x|^{-M}.
\end{aligned}
$$

While if $[J] - n + 1 \leqslant |\gamma| \leqslant [\alpha]$,

$$
\begin{aligned}
|D^\gamma Ta(x)| &\leqslant \int_{3Q_0} |D_1^\gamma K(x,y)| \, \|a\|_\infty \, dy \\
&\leqslant \int_{3Q_0} C|x-y|^{-n-|\gamma|} \, dy \\
&\leqslant C|x|^{-n-|\gamma|} \\
&\leqslant C|x|^{-n-([J]-n+1)} \\
&\leqslant C|x|^{-M}.
\end{aligned} \tag{3.24}
$$

The inequalities (3.23) and (3.24) imply (1.6) for $|x| > 4\sqrt{n}$. On the other hand, if $|x| \leqslant 4\sqrt{n}$, we use (3.14), (3.15) and (3.17) in order to apply, as in the previous theorems, Corollary 2.14 to obtain

$$|D^\gamma Ta(x)| \leqslant C,$$

which again gives (1.6) for $|x| \leqslant 4\sqrt{n}$.

We now want to check condition (1.7) for $|\gamma| = [\alpha]$. Going back to the proof of Theorem 3.7, we see that the only case that needs a different argument is when $|x - x'| < 1$ and $|x|, |x'| > 6\sqrt{n}$. We consider three different possibilities:

(*a*) If $J - n - \alpha < 0$, then

$$
\begin{aligned}
|D^\gamma Ta(x) - D^\gamma Ta(x')| &\leqslant \int_{3Q_0} |D_1^\gamma K(x,y) - D_1^\gamma K(x',y)| \, \|a\|_\infty \, dy \\
&\leqslant C|x-x'|^\delta |x|^{-n-[\alpha]-\delta} \\
&\leqslant C|x-x'|^\delta |x|^{-M},
\end{aligned}
$$

since

$$n + [\alpha] + \delta = n + \alpha - \alpha^* + \delta \geqslant n + \alpha + \epsilon \geqslant J + \epsilon = M.$$

(*b*) If $J - n - \alpha \geqslant 0$ and $[J] - n - [\alpha] = 0$, then again

$$
\begin{aligned}
|D^\gamma Ta(x) - D^\gamma Ta(x')| &\leqslant \int_{3Q_0} |D_1^\gamma K(x,y) - D_1^\gamma K(x',y)| \, \|a\|_\infty \, dy \\
&\leqslant C|x-x'|^\delta |x|^{-n-[\alpha]-\delta} \\
&\leqslant C|x-x'|^\delta |x|^{-M},
\end{aligned}
$$

because, in this case, $\epsilon = \min\{\delta - \max\{\alpha^*, J^*\}, \rho - J^*\}$ and, thus,

$$n + [\alpha] + \delta = [J] + \delta = J - J^* + \delta \geqslant J + (\delta - \max\{\alpha^*, J^*\}) \geqslant J + \epsilon = M.$$

(*c*) If $J - n - \alpha \geqslant 0$ and $k = [J] - n - [\alpha] > 0$, then, using the vanishing moments of $a(x)$, the integral form of the remainder and (3.20), we get

$$
\begin{aligned}
&|D^\gamma Ta(x) - D^\gamma Ta(x')| \\
&\leqslant \left| \int_{3Q_0} \left[D_1^\gamma K(x,y) - \sum_{|\beta|<k} D_2^\beta D_1^\gamma K(x,0) \frac{y^\beta}{\beta!} \right] a(y)\,dy \right. \\
&\quad \left. - \int_{3Q_0} \left[D_1^\gamma K(x',y) - \sum_{|\beta|<k} D_2^\beta D_1^\gamma K(x',0) \frac{y^\beta}{\beta!} \right] a(y)\,dy \right| \\
&\leqslant \int_{3Q_0} \int_0^1 \frac{(1-s)^{k-1}}{(k-1)!} \left| \sum_{|\beta|=k} [D_2^\beta D_1^\gamma K(x,sy) - D_2^\beta D_1^\gamma K(x',sy)] \frac{y^\beta}{\beta!} \right| ds\,dy \\
&\leqslant C \int_{3Q_0} \int_0^1 |x-x'|^\delta |x-sy|^{-n-|\gamma|-k-\delta} |y|^k \,ds\,dy \\
&\leqslant C|x-x'|^\delta |x|^{-n-|\gamma|-k-\delta} \\
&\leqslant C|x-x'|^\delta |x|^{-M},
\end{aligned}
$$

since

$$
n + |\gamma| + k + \delta = n + [\alpha] + [J] - n - [\alpha] + \delta = [J] + \delta = J + \delta - J^* \geqslant M.
$$

In all of the above three cases we obtain (1.7) because $|x-x'| < 1$ and $|x|, |x'| > 6\sqrt{n}$. This ends the proof of Theorem 3.13. □

We want to make some remarks regarding the above theorems, duality and interpolation for the case $1 \leqslant p, q < \infty$. In this case, $(\dot{F}_p^{\alpha,q})^* = \dot{F}_{p'}^{-\alpha,q'}$ for any $\alpha \in \mathbb{R}$, where $p' = p/(p-1)$ and $q' = q/(q-1)$. Thus, for $\alpha > 0$ and $1 < p, q \leqslant \infty$, Theorem 3.7 can be obtained from Theorem 3.1 under the additional assumption $T^*1 = 0$ (which we know is actually unnecessary). In fact, in such a case, T^* is, by Theorem 3.1, bounded on $\dot{F}_{p'}^{-\alpha,q'}$ and, hence, T is bounded on $\dot{F}_p^{\alpha,q}$. This, together with an interpolation argument, gives an alternative proof of the L^2 boundedness result of David and Journé [DJ] for the case $T1 = T^*1 = 0$. As was already observed in [FHJW], the result of David and Journé can also be obtained from Theorem 3.7 since $L^2 = \dot{F}_2^{0,2}$. Thus, Theorem 3.7 can also be used to prove directly the boundedness of Calderón-Zygmund operators on $L^p = \dot{F}_p^{0,2}$, for $1 < p < \infty$, and $H^1 = \dot{F}_1^{0,2}$, avoiding the usual duality and interpolation argument (see, for example, [CM]).

For $\alpha > 0$, $\dot{F}_2^{\alpha,2} = \dot{L}_2^\alpha$, where $\dot{L}_2^\alpha$ is the homogeneous (L^2) Sobolev or Beppo Levi space of orther α. In [MM] a criterion for the boundedness of generalized Calderón-Zygmund operators on $\dot{L}_2^\alpha$ for $0 < \alpha < 2$ is obtained, without the assumption (3.10) of our Theorem 3.7 *i.e.* without smoothness in the variable y). The reason for our assumption is that we are proving a little more about T, namely that it maps smooth atoms into smooth molecules. We also recover and extend (again under the additional smoothness in y), the result of P. Le-

marié [L], for Besov-Lipschitz spaces. In fact, for $\alpha > 0$ and $\min\{p, q\} \geqslant 1$, we can use the real interpolation result:

$$\dot{B}_p^{\alpha,q} = (\dot{B}_{p_0}^{\alpha_0,p_0}, \dot{B}_{p_1}^{\alpha_1,p_1})_{\theta,q},$$

where $\alpha = (1-\theta)\alpha_0 + \theta\alpha_1$ and $1/p = (1-\theta)(1/p_0) + \theta(1/p_1)$ with $0 < \theta < 1$, and the fact that $\dot{B}_p^{\alpha,p} = \dot{F}_p^{\alpha,p}$, to get from Theorem 3.7 a criterion for these spaces.

Theorem 3.13 can be used, in particular, to obtain a boundedness criterion for real Hardy spaces, since $\dot{F}_p^{0,2} = H^p$ for $p \leqslant 1$. The statement of this theorem is clearly more technical than the previous two due to the presence of the mixed derivative conditions. However, if T is a convolution singular integral operator, and, thus $K(x, y) = K(x - y)$, then the hypotheses of Theorem 3.13 are much easier to state. In fact, T is (essentially) its own transpose, and for $\alpha = 0$, $q = 2$ and $0 < p \leqslant 1$ the hypotheses of the theorem are reduced to:

$$T \in \mathrm{WBP},$$

$$|D^\gamma K(x)| \leqslant C|x|^{-n-|\gamma|} \quad \text{for} \quad |\gamma| \leqslant [n(1/p-1)],$$

$$|D^\gamma K(x) - D^\gamma K(x')| \leqslant C|x-x'|^\epsilon |x|^{-n-|\gamma|-\epsilon} \quad \text{for} \quad |\gamma| = [n(1/p-1)],$$

$2|x - x'| < |x|$ and $1 > \epsilon > (n/p)^*$, and

$$T(y^\gamma) = 0 \quad \text{for} \quad |\gamma| \leqslant [n(1/p-1)].$$

Moreover, the last of these conditions is really superfluous since it is guaranteed by the translation invariance of T, as the reader may check (cf. [T]). Thus we recover the classical result about the boundedness of singular integrals on H^p-spaces (see [FS, p. 191]).

Conditions (3.11) and (3.17) in the above theorems are to some extent necessary as the following theorem shows.

Theorem 3.25. *Suppose that* $\alpha > 0$, $0 < p, q \leqslant \infty$, *and*

$$T \in \mathrm{CZO}([\alpha] + \epsilon) \cap \mathrm{WBP}$$

has the norm boundedness property

$$\|T\varphi\|_{\dot{F}_p^{\alpha,q}} \leqslant C\|\varphi\|_{\dot{F}_p^{\alpha,q}} \tag{3.26}$$

for all $\varphi \in \mathcal{D} \cap \dot{F}_p^{\alpha,q}$. *Then,*

$$T(x^\gamma) = 0 \quad \textit{whenever} \quad |\gamma| < \alpha - \frac{n}{p}. \tag{3.27}$$

Let us delay the proof of this theorem for a moment and make some comments and describe some of its consequences.

The norm boundedness property of the statement of this theorem does not, a priori, allow us to extend T to a continuous operator on $\dot{F}_p^{\alpha,q}$ for the following reason. $\dot{F}_p^{\alpha,q}$ can be regarded as a space of tempered distributions modulo polynomials of degree less than or equal to $[\alpha - n/p]$ (see [FJ2]). Hence, it does not make sense to regard T as defined on $\dot{F}_p^{\alpha,q}$, unless $T(y^\gamma) = 0$ for $|\gamma| \leqslant [\alpha - n/p]$. Nevertheless, it does make sense to consider (3.26) for $\varphi \in \mathcal{D} \cap \dot{F}_p^{\alpha,q}$. Except for the gap in the case where $\alpha - n/p$ is integer, Theorem 3.25 shows that the norm property can not hold unless T annihilates polynomials of appropriate degree, so that T, in fact, can be extended to a bounded operator on $\dot{F}_p^{\alpha,q}$.

Moreover, there is in general, a gap between this theorem and conditions (3.11) or (3.17). However, if $p = \infty$ and $\alpha > 0$ is not an integer, we obtain from Theorems 3.7 and 3.25 the following

Corollary 3.28. *Let $\alpha > 0$, $\alpha \notin \mathbb{Z}$ and $1 < q < \infty$. Assume that $T: \mathcal{D} \to \mathcal{D}'$ satisfies*

$$\begin{aligned} &T \in \mathrm{WBP}, \\ &T \in \mathrm{CZO}([\alpha] + \delta) \qquad 1 > \delta > \alpha^*, \\ &T^* \in \mathrm{CZO}(\rho) \qquad 1 > \rho > 0. \end{aligned}$$

Then T extends to a bounded operator on $\dot{F}_\infty^{\alpha,q}$ if and only if $T(y^\gamma) = 0$ for all $|\gamma| \leqslant [\alpha]$.

If, in addition, $q = \infty$ we have $\dot{F}_\infty^{\alpha,\infty} = \dot{\Lambda}^\alpha$, the (homogeneous) Lipschitz space of order α, and, in this particular case, a careful look at the proof of Theorem 3.7 shows that, if we do not require that T is to map smooth atoms into smooth molecules, we can get rid of the smoothness in the y variable and still obtain the boundedness on $\dot{\Lambda}^\alpha$. Thus, we recover the result of Y. Meyer for the boundednes of generalized Calderón-Zygmund opertors on Lipschitz spaces [M2].

The gap between conditions (3.11) and (3.27) disappears if we consider conditions yielding boundedness for all $\alpha > 0$. In fact, fix p and q with $1 \leqslant p$, $q \leqslant \infty$. From Theorems 3.7 and 3.25 we immediately obtain

Corollary 3.29. *Let $T: \mathcal{D} \to \mathcal{D}'$ be a linear and continuous operator with kernel K whose restriction to Ω is a C^∞-function of x. If, in addition, we have*

$$T \in \mathrm{WBP}, \tag{3.30}$$

$$|D_1^\gamma K(x,y)| \leqslant C_\gamma |x-y|^{-n-|\gamma|} \quad \text{for all} \quad \gamma, \tag{3.31}$$

and

$$|K(x,y) - K(x,z)| \leqslant C|y-z|^{\epsilon}|x-y|^{-n-\epsilon} \tag{3.32}$$

for some $\epsilon > 0$ *whenever* $2|y-z| < |x-y|$, *then* T *extends to a bounded operator on* $\dot{F}_p^{\alpha,q}$ *for all* $\alpha > 0$ *if and only if* $T(P) = 0$ *for all polynomials* P.

Observe that in this case we also have that T is bounded if and only if T maps smooth atoms into smooth molecules for all $\alpha > 0$.

For the inhomogeneous spaces $F_p^{\alpha,q}$ (defined in [Tr], [FJ2]) we have the following

Corollary 3.33. *Suppose* $1 \leqslant p, q < \infty$. *Let* $T: \mathcal{D} \to \mathcal{D}'$ *be a linear and continuous operator satisfying* (3.31) *and* (3.32). *Suppose, also, that* T *is bounded on* L^p *and that* $T(P) = 0$ *for all polynomials* P. *Then* T *extends to a bounded operator on* $F_p^{\alpha,q}$ *for all* $\alpha > 0$.

This result follows from Theorem 3.7 (boundedness in L^p implies WBP) and the standard fact that for $\alpha > 0$ and $0 < p, q \leqslant \infty$, we have

$$\|f\|_{F_p^{\alpha,q}} \cong \|f\|_{\dot{F}_p^{\alpha,q}} + \|f\|_{L^p} \quad \text{for} \quad f \in \mathcal{S}'.$$

Recall that for $\alpha > 0$ and $1 < p < \infty$, $F_p^{\alpha,q} = L_p^{\alpha}$, the ($L^p$) Sobolev space of order α. Thus, the above corollary gives, in particular, conditions for T to be simultaneously bounded on all Sobolev spaces for a fixed p.

PROOF OF THEOREM 3.25. Let $0 < \alpha$ and $0 < \min\{p, q\} < \infty$. We may assume $\alpha > n/p$, otherwise (3.27) is void. We will need to use the following standard facts about Triebel-Lizorkin spaces

Fact 1 (see, for example, [FJ2] or [FW]). Let $\psi \in \mathcal{D}_l$ for some $l \geqslant [\alpha - n/p]$. Then there exists a constant $C_\psi > 0$ such that for all $f \in \dot{F}_p^{\alpha,q}$,

$$|\langle f, \psi \rangle| \leqslant C_\psi \|f\|_{\dot{F}_p^{\alpha,q}}.$$

Fact 2 (see [Tr, p. 239]). Let $f \in \mathcal{D} \cap \dot{F}_p^{\alpha,q}$, then $\|f^t\|_{\dot{F}_p^{\alpha,q}} \leqslant Ct^{(n/p)-\alpha}\|f\|_{\dot{F}_p^{\alpha,q}}$.

Let $|\gamma| = l < \alpha - n/p$. We have to show that $Ty^\gamma = 0$ in $\mathcal{D}_l'$. Now, although Ty^γ is initially only defined on $\mathcal{D}_l'$, using the Hahn-Banach theorem we can extend it to a distribution S in $\mathcal{D}'$ (two such extensions will, of course, differ by a polynomial of degree less than or equal to l). Let β be a multi-index such that $|\beta| > [\alpha - n/p]$. Then, $d^\beta \psi \in \mathcal{D}_{|\beta|-1} \subseteq \mathcal{D}_l$ for all $\psi \in \mathcal{D}$ and we have

$$\langle D^\beta S, \psi \rangle = (-1)^{|\beta|}\langle S, D^\beta \psi \rangle = (-1)^\beta \langle Ty^\gamma, D^\beta \psi \rangle.$$

Choose $\varphi \in \mathcal{D}$ such that $\varphi(0) \neq 0$ and $y^\gamma \varphi(y) \in \mathcal{D}_L$, where as before, $L = [J - n - \alpha]$. It is easy to check that, up to a multiplicative constant, $y^\gamma \varphi$ is a smooth atom for $\dot{F}_p^{\alpha,q}$, and, therefore, $y^\gamma \varphi \in \mathcal{D} \cap \dot{F}_p^{\alpha,q}$. Since

$$T \in \mathrm{CZO}([\alpha] + \epsilon) \subseteq \mathrm{CZO}([\alpha - n/p] + \epsilon),$$

we have, from Lemma 1.22 and the two facts stated at the beginning of this proof, that

$$\begin{aligned}
|\varphi(0)\langle D^\beta S, \psi\rangle| &= |\varphi(0)\langle T(y^\gamma), D^\beta \psi\rangle| \\
&= \lim_{t\to\infty} |\langle T(y^\gamma \varphi^t), D^\beta \psi\rangle| \\
&= \lim_{t\to\infty} t^{|\gamma|} |\langle T[(y^\gamma \varphi)^t], D^\beta \psi\rangle| \\
&\leqslant \lim_{t\to\infty} C_\psi t^{|\gamma|} \| T[(y^\gamma \varphi)^t] \|_{\dot{F}_p^{\alpha,q}} \\
&\leqslant C_{\psi,T} \lim_{t\to\infty} t^{|\gamma|} \| (y^\gamma \varphi)^t \|_{\dot{F}_p^{\alpha,q}} \\
&\leqslant C_{\psi,T} \lim_{t\to\infty} t^{|\gamma| + n/p - \alpha} \| y^\gamma \varphi \|_{\dot{F}_p^{\alpha,q}} \\
&= 0
\end{aligned}$$

since $|\gamma| < \alpha - n/p$. Thus, since $\varphi(0) \neq 0$, we see that $D^\beta S = 0$ in $\mathcal{D}'$ for all $|\beta| > [\alpha - n/p]$. But this implies that S coincides with a polynomial of degree at most $[\alpha - n/p]$ on $\mathcal{D}$, and, a fortiori, Ty^γ coincides on $\mathcal{D}_l$ with a polynomial of degree at most $[\alpha - n/p]$.

Put $P = T(y^\gamma)$. We claim that the weak boundedness property implies that the degree of $P(x)$ is at most l. Since a polynomial of degree at most l acts as the zero linear fuctional on $\mathcal{D}_l$, this will imply that $T(y^\gamma) = 0$ in $\mathcal{D}_l'$.

Suppose that

$$P(x) = \sum_{|\nu| \leq d} c_\nu x^\nu = \sum_{|\nu| < d} c_\nu x^\nu + \sum_{|\nu| = d} c_\nu x^\nu \equiv Q(x) + R(x),$$

where $d > l$ and $\sum_{|\nu| = d} |c_\nu| \neq 0$.

Let $\psi \in \mathcal{D}_l$ be such that $\int \psi(x) R(x)\, dx \neq 0$, and let $r > 0$ be such that $\operatorname{supp} \psi \subseteq B_r(0)$. Let $\chi \in \mathcal{D}$ satisfy $\chi \equiv 1$ on $B_{2r}(0)$ and $\|\chi\|_\infty = 1$. For all $t > 1$,

$$\begin{aligned}
\langle P, \psi^t\rangle &= \langle T(y^\gamma), \psi^t\rangle \\
&= \langle T(\chi^t y^\gamma), \psi^t\rangle + \langle T[(1 - \chi^t) y^\gamma], \psi^t\rangle \\
&= \langle T(\chi^t y^\gamma), \psi^t\rangle + \langle T^* \psi^t, (1 - \chi^t) y^\gamma\rangle \\
&\equiv I + II.
\end{aligned}$$

By the WBP,

$$|I| \leqslant t^{|\gamma|}|T(\chi y^\gamma)^t, \psi^t\rangle| \leqslant C_{\psi,\chi} t^{n+l},$$

and, from (1.21) we have that

$$\begin{aligned}
|II| &= |\langle T^*\psi^t, (1-\chi^t)y^\gamma\rangle| \\
&= \left|\iint K(x,y)\psi^t(x)(1-\chi^t(y))y^\gamma\,dx\,dy\right| \\
&\leqslant \int_{|y|>2rt} 2|y|^{|\gamma|}\left|\int K(x,y)\psi^t(x)\,dx\right| dy \\
&\leqslant C(rt)^{n+l+\epsilon}\|\psi\|_\infty \int_{|y|>2rt} |y|^{-n-l-\epsilon+l}\,dy \\
&\leqslant C_\psi t^{n+l}.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\langle P, \psi^t\rangle &= \int P(x)\psi^t(x)\,dx \\
&= t^n \int P(tx)\psi(x)\,dx \\
&= t^n \int Q(tx)\psi(x)\,dx + t^n \int R(tx)\psi(x)\,dx \\
&= \sum_{|\nu|<d} c_\nu t^{n+|\nu|} \int x^\nu \psi(x)\,dx + t^{n+d}\int R(x)\psi(x)\,dx,
\end{aligned}$$

and, since $\int \psi(x)R(x)\,dx \neq 0$ and $d > l$, if we let t tend to infinity, we obtain a contradiction. Therefore, the degree of the polynomial $P(x)$ is at most l. This completes the proof of the theorem. □

References

[CZ] Calderón A. and Zygmund, A. On the Existence of Certain Singular Integrals, *Acta Math.*, **88**(1952), 85-139.

[CM] Coifman, R. and Meyer, Y. Au Delá des Operateurs Pseudo-Différentiels, *Astérisque,* **57**(1982).

[CW] Coifman, R. and Weiss, G. Extensions of Hardy Spaces and Their Use in Analysis, *Bull. Amer. Math. Soc.,* **83**(1977), 569-645.

[DJ] David, G. and Journé, J.-L. A Boundedness Criterion for Generalized Calderón-Zygmund operators, *Ann. of Math.,* **120**(1984), 371-397.

[FS] Fefferman, C. and Stein, E. H^p Spaces of Several Variables, *Acta Math.,* **129**(1972), 137-193.

[FHJW] Frazier, M., Han, Y. S., Jawerth, B. and Weiss, G. The $T1$ Theorem for Triebel-Lizorkin Spaces, to appear in *Proceedings of the Harmonic Analysis.* El Escorial, June 1987; to be published by Springer Verlag (Mathematical Notes series).

[FJ1] Frazier, M. and Jawerth, B. The φ-Transform and Applications to Distribution Spaces, *Lecture Notes in Math.* **1302**(1988), 223-246.

[FJ2] Frazier, M. and Jawerth, B. A Discrete Transform and Decomposition of Distribution Spaces, preprint.

[FW] Frazier, M. and Weiss, G. Calderón-Zygmund Operators on Triebel-Lizorkin Spaces, MSRI, Berkeley, 02221-88 (1988).

[GC-RF] García-Cuerva, J. and Rubio de Francia, J. L. *Weighted Norm Inequalities and related Topics*, North Holland Mathematics Studies, Amsterdam, 1985.

[L] Lemarié, P. Continuité sur les Espaces de Besov des Operateurs Définis par des Integrales Singuliérs, *Ann. Inst. Fourier, Grenoble,* **35**(1985), 175-187.

[MM] Meyer, M. Continuité Besov de Certains Opérateurs Integraux Singulières, Thèse de 3e Cycle, Orsay 1985.

[M1] Meyer, Y. Les Nouveaux Opérateurs de Calderón-Zygmund, *Astèrisque,* **131**(1985), 237-254.

[M2] Meyer, Y. Continuité sur les Espaces de Hölder et the Sobolev des Opérateurs Définis par des Intégrales Singulières, *Recent progress in Fourier Analysis*, I. Peral and J. L. Rubio de Francia ed., North Holland, (1985), 145-172.

[S] Stein, E. *Singular Integrals and Differentiability Properties of Functions*, Princeton Univ. Press (1970).

[TW] Taibleson, M. and Weiss, G. The Molecular Characterization of Certain Hardy Spaces, *Astérisque*, **77**(1980), 67-149.

[T] Torres, R. Washington University, Ph. D. thesis.

[Tr] Triebel, H. *Theory of Function Spaces, Monographs in Mathematics*, Birkhaus, 1983.

M. Frazier*, R. Torres and G. Weiss**
University of New Mexico
Alburquerque, NM 87131
and Washington University in St. Louis
St. Louis, MO 63130

* Supported in part by NSF Postdoctoral Fellowship DMS 8705935.
** Supported in part by NSF Grant MCS 8200884.

REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.° 1, 1988

Morceaux de Graphes Lipschitziens et Intégrales Singulières sur une Surface

Guy David

Introduction

Soit K un noyau, défini sur $\mathbb{R}^N\setminus\{0\}$, C^∞, impair et homogène de degré $-d$ pour un entier $d \leqslant N$ (on pense par exemple au cas où K est le noyau de Cauchy quand $N = 2$ et $d = 1$, ou le potentiel de double-couche quand $d = N - 1$). On peut se demander pour quels objets d-dimensionels $\Sigma \subset \mathbb{R}^N$ la formule

$$Tf(x) = \text{v.p.} \int K(x - y) f(y)\, d\sigma(y),$$

où σ est la mesure de surface d-dimensionelle sur Σ, définit un opérateur borné sur $L^2(\Sigma, d\sigma)$.

Un exemple important est donné par le théorème de Coifman, McIntosh et Meyer [CMM], qui dit que T est borné sur L^2 lorsque $K(z) = 1/z$ est le noyau de Cauchy et $\Sigma \subset \mathbb{C}$ est le graphe d'une fonction lipschitzienne. Plus généralement, lorsque $d = 1$, on connait les courbes rectifiables pour lesquelles T est borné sur L^2 pour tout K comme çi-dessus: ce sont les courbes «régulières» d'Ahlfors (voir [D1]).

Quand $d > 1$, il parait exclu pour le moment de trouver des conditions géométriques simples qui soient nécessaires et suffisantes pour que T soit borné. On connait toutefois quelques conditions suffisantes:

—Lorsque Σ est le graphe d'une fonction lipschitzienne F: $\mathbb{R}^d \to \mathbb{R}^{N-d}$, la continuité de T est une conséquence assez facile du théorème de Coifman, McIntosh et Meyer.

—Dans [S1], S. Semmes définit une classe d'hypersurfaces, qu'il appelle «corde-arc avec petite constante», et qui sont une généralisation à plusieurs dimensions des coubes corde-arc (avec une petite constante) au sens suivant. Ce sont les surfaces sur lesquelles la généralisation naturelle (en termes d'algèbres de Clifford) du noyau de Cauchy définit un opérateur continu sur L^2 et presque anti-autoadjoint. Semmes caractérise ces surfaces, notamment par le fait que le vecteur unitaire normal à Σ est dans BMO (Σ) avec une petite norme. Pour ces surfaces, on prouve la continuité de T assez aisément à partir de leur caractérisation géométrique.

—On peut généraliser la notion de courbe régulière, et obtenir des «surfaces régulières», qui sont définies à partir d'un paramétrage lipschitzien, et pour lesquelles T est encore borné. (Voir la définition 1 au paragraphe B-II, ou [D2]). Un graphe lipschitzien, l'image de $\mathbb{R}^d$ par une application bi-lipschitzienne de $\mathbb{R}^d$ dans $\mathbb{R}^N$ sont des exemples de surfaces régulières. La technique utilisée pour prouver la continuité de T dans ce cas est de trouver une décomposition de chaque surface régulière du type suivant. On prouve l'existence, pour toute boule B de rayon R centrée en un point de Σ, d'un sous-ensemble $E \subset \Sigma$, de surface $\geqslant \theta R^d$ pour un $\theta > 0$, et qui soit contenu dans une surface Σ' appartenant à une classe un peu plus restreinte de surfaces. Les méthodes classiques de variable réelle permettent alors de se ramener au cas où Σ fait partie de cette classe plus restreinte de surfaces. Après un nombre fini d'applications de ce stratagème, on est réduit au cas où Σ est un graphe lipschitzien.

—Dans [S2], Semmes prouve la continuité de T, lorsque l'hypersurface Σ vérifie certaines conditions géométriques relativement générales (voir la définition 2 du paragraphe B-IV), pour la plupart des noyaux K. L'idée de la démonstration est, cette fois, d'appliquer un critère de continuité d'un opérateur d'intégrale singulière (le «Théorème $T(b)$») à l'espace de nature homogène Σ.

Nous nous proposons de reprendre l'approche de [D2], et de prouver un résultat qui permette, dans certains cas, de trouver une décomposition de Σ du type suivant. On veut trouver, pour chaque $x \in \Sigma$ et chaque rayon $R > 0$, un ensemble $E \subset \Sigma \cap B(x, R)$, de surface $\geqslant \theta R^k$, et qui soit contenu dans l'image d'un graphe lipschitzien par une isométrie de $\mathbb{R}^N$. On demande que θ, ainsi que la constante de Lipschitz du graphe, ne dépende pas de x ni de R. L'existence d'une telle décomposition entraîne la continuité de T sur Σ par des techniques classiques de variable réelle.

Déterminer pour quels Σ de telles décompositions existent ressemble beaucoup à un problème de rectifiabilité de Σ (si Σ est «totalement rectifiable»,

on sait que Σ peut être recouvert par un nombre dénombrable de surfaces de classe C^1, et donc l'une d'entre elle aura une intersection de mesure positive avec $\Sigma \cap B(x, R)$). La propriété que nous demandons est plus forte, et son caractère quantitatif, ainsi que son invariance par l'effet des dilatations, limite toutefois l'utilisation de certaines techniques de la théorie de la mesure géométrique. L'objet de cet article est de donner une construction un peu systématique qui permette, dans certains cas, d'obtenir la décomposition souhaitée de Σ.

Pour illuster le lien entre les motivations de cet article et la théorie de la rectifiabilité, signalons une question à laquelle on aimerait pouvoir répondre (on en est sans doute encore loin): si $\Sigma \subset \mathbb{R}^2$ est tel que, pour tout $x \in \Sigma$ et tout $r > 0$,

$$C_0^{-1} r \leqslant |\Sigma \cap B(x, R)| \leqslant C_0 r$$

et aussi

$$\int_{[0,\,\theta]} |\pi_\theta(\Sigma \cap B(x, R))|\, d\theta \geqslant C_0^{-1} r$$

(où $|\pi_\theta(\Sigma \cap B(x, R))|$ est la mesure de la projection orthogonale de $\Sigma \cap B(x, R)$ sur une droite de pente θ), est-ce que Σ a une décomposition comme plus haut?

On montrera que, dans certaines circonstances, si Σ est une surface de dimension d dans $\mathbb{R}^N$ contenue dans la boule unité, et si $z: \Sigma \to \mathbb{R}^d$ est une application lipschitzienne telle que $|z(\Sigma)| \geqslant \delta > 0$, alors on peut trouver un ensemble $E \subset \Sigma$ de mesure $\geqslant \theta > 0$ et tel que la restriction de z à E soit bi-lipschitzienne. (C'est le Théorème 1, dont la démonstration par un argument de temps d'arrêt occupe toute la partie A.)

Dans le cas où $\Sigma = \mathbb{R}^d$, on obtient le résultat suivant: si z est une application lipschitzienne du cube unité $I_0 \subset \mathbb{R}^d$ dans $\mathbb{R}^d$, et si $|z(I_0)| \geqslant \delta > 0$, alors on peut trouver un ensemble $E \subset I_0$ tel que $|E| \geqslant \theta > 0$ et tel que la restriction de z à E soit bi-lipschitzienne (c'est le Théorème 2, qui est prouvé au paragraphe B-I). Signalons qu'il existe maintenant une démonstration nettement plus courte, due à P. Jones, de ce résultat (et même d'un résultat un peu meilleur). Voir l'article de P. Jones dans ce numéro.

On montre ensuite que l'on peut appliquer l'un des théorèmes précédents pour trouver une décomposition (comme plus haut) des surfaces régulières (paragraphe B-II), des hypersurfaces de Semmes, ou même de surfaces de dimension $d < N - 1$ satisfaisant à une condition semblable à celle de Semmes (paragraphes B-III et B-IV). L'idée est de prendre pour fonction z la projection orthogonale sur un d-plan judicieusement choisi, l'ensemble E sur lequel z est bi-lipschitzienne est alors contenu dans l'image par une transformation orthogonale d'un graphe lipschitzien. Pour les surfaces régulières et les hyper-

surfaces de Semmes, l'existence de la décomposition est nouvelle, mais la continuité de T était connue. Signalons, dans le cas des surfaces de Semmes, l'existence d'une démonstration plus courte, et aussi un peu plus performante, due à D. Jerison et à l'auteur.

Celui-ci a le plaisir de remercier R. Coifman, P. Jones, M. Reimann, et S. Semmes pour de nombreuses discussions qui l'ont aidé. Il est honoré de pouvoir s'associer à l'excellente initiative de la «Revista» de dédier un numéro spécial à José-Luis Rubio de Francia, dont nous regrettons tous la disparition.

A. L'ARGUMENT DE TEMPS D'ARRET

1. Enoncé du Théorème

On se donne un espace Σ, muni d'une distance que l'on notera dist (et pour laquelle Σ est, disons, localement compact), et d'une mesure de Radon positive σ. On se donne aussi un entier $d \geqslant 1$, que l'on considérera comme la dimension de Σ (on pense surtout à deux cas particuliers: celui de $\mathbb{R}^d$, muni de la mesure de Lebesgue, et celui d'une surface d-dimensionelle $\Sigma \subset \mathbb{R}^N$, munie de la mesure de surface). On suppose que

(1) si B est une boule de rayon r dans Σ, $\sigma(B) \leqslant C_0 r^d$, et
(2) si $x \in \Sigma$ et $r > 0$, alors $\sigma(B(x, r)) \geqslant C_0^{-1} r^d$.

Afin de pouvoir aisément procéder à un argument de temps d'arrêt, supposons que l'on dispose d'un analogue des décompositions de $\mathbb{R}^N$ en cubes dyadiques. Plus précisément, faisons l'hypothèse suivante:

(Q) Il existe une suite ϵ_j de réels > 0, tendant vers 0 et, pour chaque ϵ dans la suite des ϵ_j et pour chaque entier $n \geqslant 0$, une partition $Q(n)$ de Σ en ensembles R, $R \in Q(n)$, avec les propriétés suivantes (où l'on a fixé ϵ):
(3) si $n > n'$ et si $R \in Q(n)$, $R' \in Q(n')$ sont tels que $R \cap R' \neq \emptyset$, alors $R \subset R'$;
(4) pour tout $n \geqslant 0$ et tout $R \in Q(n)$, le diamètre de R est $\leqslant C_0 \epsilon^n$;
(5) pour tout $n \geqslant 0$ et tout $R \in Q(n)$, $\sigma(R) \geqslant C_0^{-1} \epsilon^{nd}$;
(6) si $n \geqslant 0$ et $R \in Q(n)$,

$$\sigma(\{x \in \Sigma, \operatorname{dist}(x, \partial R) \leqslant C\epsilon^{n+1}\}) \leqslant \tau \epsilon^{nd},$$

où, pour tout C, τ tend vers 0 quand ϵ tend vers 0.

Lorsque $\Sigma = \mathbb{R}^d$ et σ est la mesure de Lebesgue, nous choisirons bien entendu les cubes ϵ^{-1}-adiques, où ϵ est l'inverse d'un entier; dans le cas où Σ est contenu dans $\mathbb{R}^N$, nous construirons les $Q(n)$ au paragraphe B-III. Nous ne travaillerons en général qu'avec une seule valeur de ϵ à la fois (nous aurons

seulement besoin de savoir que τ est aussi petit que l'on veut), et nous appellerons les éléments de $Q(n)$ des «n-cubes».

On se donne encore une fonction $z: \Sigma \to \mathbb{R}^d$, que l'on supposera C_0-lipschitzienne. On supposera en outre que

(7) $|z(A)| \leqslant C_0\sigma(A)$ pour tout ensemble Borélien $A \subset \Sigma$.

(On vérifiera plus tard que, au moins dans le cas où σ est une mesure de Radon sur $\mathbb{R}^N$ et Σ est son support, (7) est en fait une conséquence des hypothèses précédentes; pour le moment, considérons (7) comme une hypothèse supplémentaire.)

Finalement, donnons-nous un 0-cube $I_0 \in Q(0)$ et supposons que, pour un $\delta > 0$, on ait

$$|z(I_0)| \geqslant \delta\sigma(I_0). \tag{8}$$

On souhaite trouver un ensemble $E \subset I_0$, avec $\sigma(E) \geqslant \theta > 0$, et tel que $z_{|E}$ soit bi-lipschitzienne. Notre hypothèse principale, qui nous permettra de nous ramener au cas où δ est plus grand, est la suivante.

(9) On peut, pour toute constante $C_1 \geqslant 0$, trouver des constantes C_2 et $\eta > 0$ telles que, si $x \in \Sigma$, $n \in \mathbb{N}$, et T est l'union des n-cubes qui rencontrent $B(x, C_2\epsilon^n)$, et si $|z(T)| \geqslant \delta\sigma(T)/2$, alors l'une des deux choses suivantes se produit:

—$z(T)$ contient la boule de centre $z(x)$ et de rayon $C_1\epsilon^n$;
—on peut trouver un n-cube $R \subset T$, tel que

$$|z(R)| \geqslant (1 + 2\eta)\sigma(R)|z(T)|\sigma(T)^{-1}.$$

(Dans ce qui précède, on impose que les constantes η et C_2 puissent être choisies indépendantes de ϵ assez petit.)

Théorème 1. *Soient Σ, σ, z (une fonction C_0-lipschitzienne) et I_0 (un 0-cube) tels que* (1), (2), (*Q*), (7), (8) *et* (9) *soient satisfaits. Alors, on peut trouver un fermé $E \subset I_0$ tel que $\sigma(E) \geqslant \theta$ et tel que $|z(x) - z(y)| \geqslant M^{-1}\operatorname{dist}(x, y)$ pour x, y dans E. Les constantes $\theta \geqslant 0$ et M ne dépendent que de C_0, δ, C_2 et η (où C_1 ne dépend que de C_0).*

La démonstration du théorème occupera le reste de la partie A.

2. Description du Temps D'Arrêt

La démonstration repose sur l'énoncé suivant, qui nous permettra de nous ramener, par une itération, au cas où δ est très grand.

Lemme 1. *Soient* Σ, σ, z et I_0 *comme dans le théorème. On peut trouver une constante* $\eta_0 > 0$ *telle que, si* ϵ *est assez petit et si l'on a de plus, pour un* $M_0 > 0$ *et un* $\theta_0 > 0$,

(10) *pour tout* $n \in \mathbb{N}$ *et tout n-cube Q tel que* $|z(Q)| \geqslant (1 + \eta_0)\delta\sigma(Q)$, *on peut trouver un fermé* $E(Q) \subset Q$ *tel que* $\sigma(E(Q)) \geqslant \theta_0 \epsilon^{nd}$ *et pour lequel* $|z(x) - z(y)| \geqslant M_0^{-1} \operatorname{dist}(x, y)$ *pour x et y dans* $E(Q)$,

alors on peut trouver un fermé $E \subset I_0$ *tel que* $\sigma(E) \geqslant \theta$ *et tel que*

$$|z(x) - z(y)| \geqslant M^{-1} \operatorname{dist}(x, y)$$

pour x, y dans E.

Avant de commencer la construction de l'ensemble E du lemme, signalons que nous utiliserons systématiquement la lettre C pour désigner des constantes qui dépendent de manière simple des constantes géométriques C_0, δ, C_2 et η. Plus précisément, les constantes seront choisies dans l'ordre suivant: C_1, puis C_4, C_2, η, C_3, η_1, η_2, η_0, enfin α et les constantes M et θ (tout cela vaudra dès que $\epsilon < \epsilon_0$, où ϵ_0 sera choisi en fonction de η_0).

Nous noterons $\partial(A) = |z(A)|/\sigma(A)$, pour tout $A \subset \Sigma$ tel que $\sigma(A) \neq 0$.

Nous allons définir une suite d'ensembles $E_0 = I_0 \supset E_1 \supset \cdots \supset E_k \supset \cdots$, où chaque E_{k+1} est obtenu à partir de E_k en lui retirant une union finie de n-cubes (nous noterons cette union Q_k). L'élimination des Q_k se fera en plusieurs générations; à la n-ième génération, les Q_k seront composés de n-cubes.

Supposons que nous ayons fini la $(n-1)$-ième génération d'éliminations, et décrivons la n-ième génération. A la fin de la génération $(n-1)$, on aura construit des ensembles $E(n-1)$ et $G(n-1)$, où $E(n-1)$ est le dernier E_k de la génération $n-1$ (à propos, nous appellerons $k(n-1)$ l'entier tel que $E(n-1) = E_{k(n-1)}$), et $G(n-1)$ est une partie de $E(n-1)$. Pour $n = 1$, on pose $E(0) = E_0$, $k(0) = 0$, et $G(0) = \varnothing$. Dans tous les cas, l'ensemble $G(n-1)$ sera une union de m-cubes R (où $m < n$) tels que

$$\partial(R) \geqslant (1 + \eta_0)\delta, \tag{11}$$

(12) $\operatorname{dist}(R, E(n-1)\setminus R) \geqslant C_3 \epsilon^m$ et $\operatorname{dist}(z(R), z[E(n-1)\setminus R]) \geqslant \epsilon^m$, où la valeur de la constante C_3 sera décidée plus tard.

Considérons l'ensemble $E(n-1)\setminus G(n-1)$ (qui est une union de $(n-1)$-cubes) comme une union de n-cubes. La n-ième génération d'élimination se passe en plusieurs étapes, que nous allons maintenant décrire.

Etape A. On considère les n-cubes R de $E(n-1)\setminus G(n-1)$ qui vérifient

$$\partial(R) \geqslant (1 + \eta_0)\delta. \tag{13}$$

On se donne un ensemble maximal de tels cubes R_j, $1 \leqslant j \leqslant m$, tels que $\operatorname{dist}(z(R_i), z(R_j)) > \epsilon^n$ et $\operatorname{dist}(R_i, R_j) > C_3\epsilon^n$ pour $i \neq j$. On élimine tous les n-cubes Q_k encore dans $E_k \setminus G(n-1)$ tels que $\operatorname{dist}(z(Q_k), z(R_j)) \leqslant \epsilon^n$ ou $\operatorname{dist}(Q_k, R_j) \leqslant C_3\epsilon^n$ pour l'un des R_j.

L'ensemble $G(n)$ sera la réunion de $G(n-1)$ et de l'union des n-cubes R vérifiant (13) qui n'ont pas été éliminés lors de l'étape A. On ne touchera plus à ces cubes au cours de la construction. Notons que, si $Q_k \subset E(n-1)$ est éliminé lors de cette étape, alors il restera un n-cube Q de $G(n)$ tel que $\operatorname{dist}(z(Q), z(Q_k)) \leqslant C'\epsilon^n$ (où C' ne dépend que de C_0 et C_3).

Etape B. On considère, pour $x \in \Sigma$, l'ensemble $T(x)$ formé de l'union des n-cubes qui rencontrent $B(x, C_2\epsilon^n)$. On élimine une famille maximale de $T_i = T(x_i)$ deux à deux disjoints, entièrement contenus dans $E_k \setminus G(n-1)$ à la fin de l'étape A, et tels que

$$\partial(T_i) \leqslant (1 - \eta/2)\delta. \tag{14}$$

Etape C. On élimine maintenant tous les n-cubes Q_k encore dans $E_k \setminus G(n-1)$ et tels que $\partial(Q_k) \leqslant (1-\eta)\delta$.

Etape D. On retire, aussi longtemps qu'on peut, un n-cube $Q_k \subset E_k \setminus G(n-1)$ tel que l'on puisse trouver un n-cube S, encore dans E_k, différent de Q_k, et pour lequel

$$|z(Q_k) \cap z(S)| \geqslant \eta_1 \delta\sigma(Q_k). \tag{15}$$

Etape E. On élimine, aussi longtemps qu'on peut, un n-cube Q_k tel que l'on puisse trouver un autre n-cube $R \subset E_k$, tel que $\operatorname{dist}(R, Q_k) \geqslant C_3\epsilon^n$, mais $\operatorname{dist}[z(R), z(Q_k)] \leqslant \epsilon^n$.

Notons qu'à la fin de cette étape, il ne reste plus de n-cubes R et S dans $E_k \setminus G(n-1)$, tels que $\operatorname{dist}(Q, R) \geqslant C_3\epsilon^n$ mais $\operatorname{dist}[z(Q), z(R)] \leqslant \epsilon^n$.

Nous venons de finir la n-ième génération de notre construction. On note $k(n)$ le dernier des indices k, et $E(n) = E_{k(n)}$. L'ensemble $G(n)$ a été défini à la fin de l'étape B; tout a été fait pour qu'il satisfasse encore aux conditions (11) et (12). Notons aussi que l'on n'a pas touché à $G(n-1)$, de sorte que l'on a bien $G(n) \supset G(n-1)$.

On décide de continuer indéfiniment le processus que nous venons de décrire, sauf si l'un des événements suivants se produit:

(i) $\sigma(G(n)) \geqslant \alpha$;
(ii) $\partial(E_k) \geqslant (1 + 2\eta_0)\delta$;
(iii) $\sigma(E_k) \leqslant \sigma(E_0)/2$ ou $\partial(E_k) \leqslant (1 - \eta_2)\delta$.

Nous verrons plus tard ce qu'il convient de faire lorsque (i) ou (ii) se produit. Nous allons d'abord prouver que (iii) n'arrive jamais en premier.

3. La Condition (iii) ne se produit pas

On se place dans le cas où aucun des temps d'arrêt ne s'est encore produit. En particulier, on suppose que

$$(1 - \eta_2)\delta \leqslant \partial(E_k) \leqslant (1 + 2\eta_0)\delta, \tag{16}$$

et, compte tenu de (4) et à condition que ϵ soit assez petit,

$$\sigma(E_0)/3 \leqslant \sigma(E_{k+1}) \leqslant \sigma(E_0), \tag{17}$$

ce qui donne

$$C^{-1}\delta \leqslant \partial(E_k)\sigma(E_{k+1}) \leqslant C\delta. \tag{18}$$

On va essayer d'évaluer jointement $\partial(E_k)$ et $\sigma(E_k)$ en prouvant que le premier augmente un peu chaque fois que l'on diminue le second. On veut donc, pour chaque k, évaluer la quantité $\mu(k) = \partial(E_{k+1})\partial(E_k)^{-1}$ en fonction de $\sigma(Q_k) = \sigma(E_k) - \sigma(E_{k+1})$. On sera amené à distinguer plusieurs cas; le lemme suivant résume les estimations que nous utiliserons.

Lemme 2. *Dans tous les cas, on a*

$$\mu(k) \geqslant 1 - C\sigma(Q_k). \tag{19}$$

Lorsque Q_k a disparu lors d'une des étapes B ou C,

$$\mu(k) \geqslant 1 + C^{-1}\eta\sigma(Q_k). \tag{20}$$

Lorsque Q_k a disparu lors de l'étape D,

$$\mu(k) \geqslant 1 + C^{-1}\eta_1\sigma(Q_k). \tag{21}$$

Finalement, si Q_k a disparu lors de l'étape E,

$$\mu(k) \geqslant 1 - C\eta_2\sigma(Q_k). \tag{22}$$

Commençons par prouver (19). On utilise l'estimation brutale suivante:

$$\begin{aligned} |z(E_{k+1})| &\geqslant |z(E_k)| - |z(Q_k)| \\ &\geqslant \partial(E_k)\sigma(E_k) - C_0\sigma(Q_k) \\ &\geqslant \partial(E_k)\sigma(E_{k+1}) - C_0\sigma(Q_k) \\ &\geqslant [\partial(E_k)\sigma(E_{k+1})][1 - C\sigma(Q_k)] \end{aligned} \tag{23}$$

grâce à (18), ce qui donne (19).

Supposons maintenant que Q_k ait été oté à l'étape B. On a

$$|z(Q_k)| \leqslant (1 - \eta/2)\delta\sigma(Q_k)$$

grâce à (14), de sorte que

$$\begin{aligned}|z(E_{k+1})| &\geqslant |z(E_k)| - |z(Q_k)| \\ &\geqslant \partial(E_k)\sigma(E_k) - (1 - \eta/2)\delta\sigma(Q_k) \\ &\geqslant \partial(E_k)\sigma(E_{k+1}) + \sigma(Q_k)[\partial(E_k) - (1 - \eta/2)\delta] \\ &\geqslant \partial(E_k)\sigma(E_{k+1}) + \sigma(Q_k)[(1 - \eta_2)\delta - (1 - \eta/2)\delta]\end{aligned}$$

(à cause de (16))

$$\begin{aligned}&\geqslant \partial(E_k)\sigma(E_{k+1}) + \sigma(Q_k)[(\eta/2) - \eta_2]\delta \\ &\geqslant \partial(E_k)\sigma(E_{k+1})[1 + C^{-1}\sigma(Q_k)[(\eta/2) - \eta_2]/2]\end{aligned}$$

(grâce à (18))

$$\geqslant \partial(E_k)\sigma(E_{k+1})[1 + C^{-1}\eta\sigma(Q_k)]$$

si η_2 est assez petit.

On en déduit (20) lorsque Q_k a été oté à l'étape B. Si Q_k a disparu au cours de l'étape C, on utilise l'inégalité $|z(Q_k)| \leqslant (1 - \eta)\delta\ \sigma(Q_k)$ au lieu de (14), et on obtient (20) de la même manière.

Passons à l'inégalité (21). Donnons-nous un Q_k qui a été retiré à l'étape D, et écrivons

$$\begin{aligned}|z(E_{k+1})| &= |z(E_k)| - |z(Q_k)| + |z(E_{k+1}) \cap z(Q_k)| \\ &\geqslant \partial(E_k)\sigma(E_k) - |z(Q_k)| + \eta_1\delta\sigma(Q_k).\end{aligned}$$

Notons que $\partial(Q_k) \leqslant (1 + \eta_0)\delta$, car sinon Q_k serait dans $G(n)$ ou aurait disparu lors de l'étape A. Par conséquent,

$$\begin{aligned}|z(E_{k+1})| &\geqslant \partial(E_k)\sigma(E_k) + \sigma(Q_k)[\eta_1\delta - (1 + \eta_0)\delta] \\ &\geqslant \partial(E_k)\sigma(E_{k+1}) + \sigma(Q_k)[(1 - \eta_2)\delta + \eta_1\delta - (1 + \eta_0)\delta] \\ &\geqslant \partial(E_k)\sigma(E_{k+1}) + \eta_1\delta\sigma(Q_k)/3\end{aligned}$$

si η_2 et η_0 sont assez petits par rapport à η_1. On en déduit (21) en divisant par $\partial(E_k)\sigma(E_{k+1})$ et en utilisant (18).

Pour montrer (22), utilisons le fait que, si Q_k a disparu à l'étape E, on a encore

$$|z(Q_k)| \leqslant (1 + \eta_0)\delta\sigma(Q_k).$$

Il vient

$$\begin{aligned}|z(E_{k+1})| &\geqslant |z(E_k)| - |z(Q_k)| \\ &= \partial(E_k)\sigma(E_k) - |z(Q_k)| \\ &\geqslant \partial(E_k)\sigma(E_k) - (1+\eta_0)\delta\sigma(Q_k) \\ &\geqslant \partial(E_k)\sigma(E_{k+1}) - \sigma(Q_k)(\eta_2+\eta_0),\end{aligned}$$

d'où l'on déduit aisément (22) si $\eta_0 \leqslant \eta_2$. Nous avons prouvé le Lemme 2.

Lorsque, au cours d'une des étapes A ou E, on retire un grand nombre de n-cubes Q_j (disons, pour $k \leqslant j < k+L$) tels que les $z(Q_j)$ soient contenus dans une même boule de rayon $\leqslant C\epsilon^n$, on utilisera l'estimation qui suit de préférence à celle du Lemme 2:

$$\begin{aligned}|z(E_{k+L})| &\geqslant |z(E_k)| - C'\epsilon^{nd} \\ &\geqslant \partial(E_k)\sigma(E_{k+L}) + \partial(E_k)[\sigma(E_k) - \sigma(E_{k+L})] - C'\epsilon^{nd} \\ &\geqslant \partial(E_k)\sigma(E_{k+L}) + (\delta/2)[\sigma(E_k) - \sigma(E_{k+L})]\end{aligned}$$

si les Q_j sont assez nombreux pour que

$$[\sigma(E_k) - \sigma(E_{k+L})] \geqslant 10\,C'\delta^{-1}\epsilon^{nd}. \tag{24}$$

Dans ce cas, on divise par $\partial(E_k)\sigma(E_{k+L})$; si l'on a de plus

$$[\sigma(E_k) - \sigma(E_{k+L})] \leqslant \sigma(E_0)/10, \tag{25}$$

on peut appliquer l'équivalent de (18), et il vient

$$\mu(k)\mu(k+1)\cdots\mu(k+L-1) \geqslant 1 + C^{-1}[\sigma(Q_k) + \cdots + \sigma(Q_{k+L-1})]. \tag{26}$$

Notons que (25) n'est en fait jamais une restriction: il suffit, s'il y a trop de cubes Q_j, de faire plusierurs paquets, pour lesquels on aura (25), et auquels on appliquera (26) séparément. D'autre part, quitte à modifier l'ordre dans lequel on retire les Q_j, on peut s'arranger pour que l'on puisse regrouper certains des Q_j en paquets de la forme $B_s = \{Q_j : k(s) \leqslant j < k(s) + L(s)\}$ tels que (24) et (25) soient vrais, tout cela de manière à ce que l'on ait la propriété suivante:

(27) pour chaque boule $B \subset \mathbb{R}^d$ de rayon ϵ^n, la mesure totale pour σ des ensembles Q_j qui ont été otés à l'étape A ou E, tels que $z(Q_j)$ rencontre B, et qui ne font partie d'aucun d'un des paquets B_s, est inférieur à $C_4\epsilon^{nd}$ (où C_4 dépend de C_0 et de δ).

Nous supposerons dans la suite que l'on a fait un tel choix de l'ordre des Q_j, et des B_s.

Nous utiliserons la convention suivante. Lorsque Q_k est dans l'un des paquets B_s, nous dirons que $k \in O$. Sinon, nous noterons $k \in A$ (resp. BC, D, E)

le fait que Q_k a disparu à l'étape A (resp. B ou C, D, E). Lorsqu'on peut appliquer l'une des inégalités (20), (21), ou (26), nous nous estimerons satisfaits parce que, dans ce cas, la diminution de $\sigma(E_k)$ est clairement liée à une augmentation de $\partial(E_k)$. Lorsque $k \in A$ ou $k \in E$, nous allons essayer d'associer à k un événement plus favorable.

Si Q_k est une victime de l'étape A qui ne fait pas partie d'un des blocs B_s, on note $T(k) = Q$, où Q est l'un des cubes de $G(n) \setminus G(n-1)$ qui ont été selectionnés lors de l'étape A, et qui est tel que $\operatorname{dist}(z(Q), z(Q_k)) \leqslant C'\epsilon^n$ (où C' est assez grand par rapport à C_0 et à C_3: voir la remarque qui suit la définition de l'étape A).

Le cas où $k \in E$ est un peu plus compliqué. Donnons-nous un tel k et soit T l'union des n-cubes qui rencontrent $B(x, C_2\epsilon^n)$ (où x est un point de Q_k pris au hasard). Nous allons distinguer cinq cas.

(i) T n'est pas contenu dans $E(n-1)$. Il se peut que ce soit parce que Q_k est à distance $\leqslant C\epsilon^n$ du bord de I_0. Dans ce cas, on notera $T(k) = \varnothing$. Sinon, Q_k est à distance $\leqslant C\epsilon^n$ d'un des cubes R_j qui ont été otés lors d'une génération précédente; alors on pose $T(k) = R_j$.

(ii) T est contenu dans $E(n-1)$, mais une partie de T a disparu lors de l'étape A. Dans ce cas, on note $T(k) = Q_j$, où Q_j est un n-cube de $G(n)$ qui a été selectionné à l'étape A, et tel que $\operatorname{dist}(z(Q_j), z(Q_k)) \leqslant C''\epsilon^n$ (où C'' dépend de C_0, C_2 et C_3).

(iii) T était encore contenu dans E_m à la fin de l'étape A, et $\partial(T) < (1 - \eta/2)\delta$. Alors, T n'a pas disparu non plus au cours de l'étape B (sinon Q_k ne serait plus là). On peut donc trouver un paquet T_i de cubes qui a été éliminé lors de la phase B, et qui est à distance $\leqslant CC_2\epsilon^n$ de Q_k. On prend $T(k) = T_i$.

(iv) T était encore contenu dans E_m à la fin de l'étape A, $\partial(T) \geqslant (1 - \eta/2)\delta$, et $z(T)$ ne contient pas la boule centrée en $z(x)$ et de rayon $C_1\epsilon^n/10$. Dans ce cas, on peut appliquer (9), et trouver un n-cube $R \subset T$, tel que

$$|z(R)| \geqslant (1 + 2\eta)\sigma(R)\partial(T) \geqslant (1 + \eta_0)\delta\sigma(R).$$

Comme R n'a pas disparu à l'étape A puisque T était encore intact à la fin de cette étape, R est dans $G(n)$. Ce cas est impossible, car si R était dans $G(n)$, on aurait éliminé tous ses voisins lors de l'étape A.

(v) Il reste le cas où T était encore contenu dans E_m à la fin de l'étape A, où $\partial(T) \geqslant (1 - \eta/2)\delta$, et $z(T)$ contient la boule centrée en $z(v)$ et de rayon $C_1\epsilon^n/10$. Soit S un des n-cubes encore contenu dans E_k tels que

$$\operatorname{dist}(Q_k, S) \geqslant C_3\epsilon^n \quad \text{et} \quad \operatorname{dist}(z(Q_k), z(S)) \leqslant \epsilon^n.$$

On choisit maintenant C_1 assez grand par rapport à C_0 pour que la dernière inégalité implique que $z(S)$ est contenu dans $z(T)$. Comme il y a moins de C n-cubes dans T, on peut trouver un cube Q dans T tel que

$|z(S) \cap z(Q)| \geqslant C^{-1}\delta\sigma(S) > \eta_1 \delta\sigma(S)$ si η_1 est assez petit. Comme S est encore là, c'est que Q a disparu avant la fin de l'étape D (donc au cours de l'une des étapes B, C, ou D). On prend $T(k) = Q$.

Nous avons maintenant défini $T(k)$ dans tous les cas qui nous intéressent, mais nous ne sommes pas encore sûrs que $T(k)$ corresponde vraiment à un événement plus favorable. Définissons donc un autre ensemble $U(k)$. Si $T(k) = \emptyset$, si $T(k)$ est l'un des cubes de $G(n)$, ou si $T(k)$ est l'un des cubes pour lesquels on peut appliquer l'une des inégalités favorables (20), (21), ou (26), on garde $U(k) = T(k)$. Sinon, $k \in E$ et l'on est dans le cas (i); on peut donc définir $U(k)$ de manière récursive (génération par génération) par $U(k) = U(j)$, où $T(k) = Q_j$.

Lemme 3. *Si R est un n-cube, ou $\emptyset$, notons $A(R)$ l'union des Q_k tels que $U(k) = R$; notons aussi $B(R)$ l'union des Q_k de $A(R)$ qui ne sont pas de la même génération que R (dans le cas de $\emptyset$, prenons $B(R) = A(R)$). Alors*

$$\sigma(B(R)) \leqslant \tau\sigma(R), \tag{28}$$

où τ est aussi petit qu'on veut pourvu que ϵ soit assez petit, et

$$\sigma(A(R)) \leqslant C\sigma(R). \tag{29}$$

Enfin, $\sigma(A(\emptyset)) \leqslant \tau$.

Pour prouver le lemme, commençons par estimer $\sigma(A(R) \setminus B(R))$. Si R est l'un des n-cubes de $G(n)$, tous les Q_k de la même génération que R tels que $T(k) = R$ sont tels que dist $(z(Q_k), z(R)) \leqslant C'\epsilon^n$; grâce à (27), leur mesure totale pour σ est inférieure à $CC_4\epsilon^{nd} \leqslant C\sigma(R)$ (à cause de (5)). Si R est l'un des Q_j qui ont disparu à la génération n, tous les Q_k de la même génération tels que $T(k) = R$ sont à distance $\leqslant C\epsilon^n$ de R, et par conséquent leur mesure totale pour σ est $\leqslant C\sigma(R)$ à cause de (1).

Il nous faut encore traiter les Q_k des générations précédentes. Si $T(k)$ est l'un des cubes S que nous venons de traiter, alors Q_k est contenu dans l'ensemble des points qui sont à distance $\leqslant C\epsilon^{n+1}$ du bord de S. L'hypothèse (Q) nous dit que la mesure pour σ de cet ensemble est $\leqslant \tau\epsilon^{nd}$, où τ est aussi petit que l'on veut pourvu que ϵ soit assez petit. La mesure totale de l'union des Q_k tels que $T(k)$ soit contenu dans l'un des S est donc $\leqslant C\tau\sigma(R)$. La même chose vaut pour les Q_m de la même génération que l'un des Q_k et tels que $T(m) \subset Q_k$ (on applique notre première majoration à chaque Q_k). On traite pareillement les cubes des générations suivantes dont l'une des images par T tombe dans R. On obtient ainsi (28) (avec la constante $C\tau$); (29) s'ensuit aussitôt. La dernière inégalité se montre comme (28). On en déduit le lemme 3.

Nous allons bientôt pouvoir estimer $P = \Pi\mu(k)$, où le produit porte sur $k < k_0$ pour un k_0 tel qu'aucun des temps d'arrêt ne s'est encore produit. On écrit $P = \Pi_1\Pi_2\Pi_3\Pi_4\Pi_5$, où

$$\Pi_1 = \prod_{k\in BC}\left[\mu(k)\prod_{j\in U^{-1}(Q_k)}\mu(j)\right], \qquad \Pi_2 = \prod_{k\in D}\left[\mu(k)\prod_{j\in U^{-1}(Q_k)}\mu(j)\right],$$

$$\Pi_3 = \prod_{k\in 0}\left[\prod_{j\in U^{-1}(Q_k)}\mu(j)\right], \qquad \Pi_4 = \prod_{R\in G(n)}\prod_{j\in U^{-1}(Q_k)}\mu(j),$$

$$\Pi_5 = \prod_{j\in U^{-1}(\varnothing)}\mu(j).$$

Bien entendu, on ne fait porter les produits que sur j et k inférieurs à k_0. Notons qu'aucun des termes du produit n'a été oublié: certains ont seulement été regroupés avec d'autres pour faciliter le calcul qui va suivre. Nous allons maintenant minorer chacun des Π_i séparément.

En ce qui concerne Π_1, notons $F(k)$, pour $k \in BC$, l'ensemble des $j \in U^{-1}(Q_k)$ tels que Q_j ait disparu à l'étape E. Pour ces Q_j, on utilisera (22). Si $k \in BC$ et si j n'est pas dans $F(k)$, alors Q_j est d'une génération différente de celle de k. On notera $H(k)$ l'ensemble des $j \in U^{-1}(Q_k)$ qui ne sont pas dans $F(k)$ (on leur appliquera (19)). Compte tenu de (20) et du Lemme 3,

$$\begin{aligned}\Pi_1 &\geqslant \prod_{k\in BC}[1 + C^{-1}\eta\sigma(Q_k)]\prod_{j\in F(k)}[1 - C\eta_2\sigma(Q_j)]\prod_{j\in H(k)}[1 - C\sigma(Q_j)]\\ &\geqslant \prod_{k\in BC}[1 + C^{-1}\eta\sigma(Q_k)][1 - C'\eta_2\sigma(Q_k)][1 - C'\tau\sigma(Q_k)]\\ &\geqslant \prod_{k\in BC}[1 + \eta\sigma(Q_k)/C'']\end{aligned}$$

à condition que η_2 et τ soient assez petits par rapport à η.

Le même calcul donne $\Pi_2 \geqslant \prod [1 + C^{-1}\eta_1\sigma(Q_k)]$, où le produit porte sur $k \in D$, et $\Pi_3 \geqslant \prod [1 + C^{-1}\sigma(Q_k)]$, avec un produit qui porte sur $k \in O$ (la seule différence est que l'on applique (21) au lieu de (20) dans le premier cas; dans le second cas, on regroupe les termes en paquets correspondant aux B_s, et on applique (26).

Lorsque $R \in G(n)$ et $U(j) = R$, on minore $\mu(j)$ en utilisant (19). Il vient, en utilisant (29), $\Pi_4 \geqslant \prod [1 - C\sigma(R)]$, où le produit porte sur tous les cubes de $G(n)$. Puisque le temps d'arrêt (i) ne s'est pas encore produit, la somme des $\sigma(R)$ est inférieure à α; en choisisant α assez petit, on obtient $\Pi_4 \geqslant (1 - \eta_2/10)$.

Pour minorer Π_5, on utilise à nouveau (19) et le Lemme 3. Il vient $\Pi_5 \geqslant 1 - C\tau \geqslant (1 - \eta_2/10)$ si ϵ est assez petit.

Nous sommes enfin en mesure de prouver que le temps d'arrêt (iii) ne se produit pas en premier. Tout d'abord, notons que $P = \prod \mu(k) \geqslant (1 - \eta_2/2)$, et par conséquent que (iii) ne peut se produire que si $\sigma(E_{k_0}) \leqslant \sigma(E_0)/2$, ce qui entraîne $\sum \sigma(Q_k) \geqslant \sigma(E_0)/2$. Si c'était le cas, on aurait encore, en utilisant le

Lemme 3 et si α est assez petit, $\sum' \sigma(Q_k) \geqslant \sigma(E_0)/C$, où la somme ne porte que sur les $k < k_0$ tels que $k \in BC$, $k \in D$ ou $k \in O$. Comme nous venons de voir que $P \geqslant (1 - \eta_2/2) \prod [1 + \eta_1 \sigma(Q_k)/C]$, où le produit porte justement sur $k \in (BC \cup D \cup O)$; on en déduit que $P \geqslant 1 + 3\eta_0$, si η_2 et η_0 sont assez petits par rapport à η_1. Dans ce cas, (ii) s'est produit en premier. Nous avons enfin montré que (iii) ne se produit pas.

4. Fin de la Preuve du Théorème

Finissons d'abord la démonstration du Lemme 1. Si le temps d'arrêt (i) se produit en premier, choisissons $E = \cup E(R)$, où l'union porte sur tous les cubes R composant $G(n)$, et $E(R)$ est donné par l'hypothèse (10). Comme $\sigma(G(n)) \geqslant \alpha$, il vient

$$\sigma(E) \geqslant \sum_{m \leq n} \sum_{R \in G(m) \setminus G(m-1)} \theta_0 \epsilon^{md} \geqslant \theta_0 C_0^{-1} \alpha = \theta.$$

Il ne reste plus qu'à vérifier que la restriction de z à E est M-bi-lipschitzienne. Nous ferons cette vérification un peu plus loin.

Supposons maintenant que le temps d'arrêt (ii) se produise au cours de la n-ième génération. Soient V l'union des n-cubes de $E_k \setminus G(n-1)$ tels que $\partial(Q) \geqslant (1 + \eta_0)\delta$, W l'union des cubes de $G(n-1)$, et X l'union des autres cubes de $E_k \setminus G(n-1)$. On a

$$\begin{aligned}(1 + 2\eta_0)\delta\sigma(E_k) &\leqslant \partial(E_k)\sigma(E_k) \\ &= |z(E_k)| \\ &\leqslant |z(V)| + |z(W)| + |z(X)| \\ &\leqslant |z(V)| + C\sigma(W) + (1 + \eta_0)\delta\sigma(X) \\ &\leqslant |z(V)| + C\sigma(W) + (1 + \eta_0)\delta\sigma(E_k),\end{aligned}$$

d'où on déduit que $\eta_0 \delta\sigma(E_k) \leqslant |z(V)| + C\sigma(W)$. Or $\sigma(E_k) \geqslant \sigma(E_0)/2$, et $\sigma(W) \leqslant \alpha$. Donc, si α est assez petit, $|z(V)| \geqslant \eta_0 \delta\sigma(E_0)/3$. On en déduit entre autres que $\sigma(V) > \alpha$, et par conséquent, le temps d'arrêt (ii) s'est produit au cours de l'étape A (sinon, (i) se serait produit en premier). On décide de finir l'étape A de la n-ième génération.

Nous avons vu lors de la définition de l'étape A que pour tout n-cube R dans V, on peut trouver un n-cube Q, qui reste selectionné à l'issu de l'étape A, et tel que $\operatorname{dist}(z(Q), z(R)) \leqslant C'\epsilon^n$. Donnons-nous une constante $C'' \geqslant C'$. Comme $|z(V)| \geqslant \eta_0 \delta\sigma(E_0)/3 \geqslant 1/C$, l'ensemble $z(V)$ ne peut pas être entièrement recouvert par moins de $C^{-1}(C''\epsilon^n)^{-d}$ boules de rayon $C''\epsilon^n$, ce qui implique, par un facile argument de recouvrement, que si C'' est assez grand, on peut trouver presque autant de n-cubes de V dont les images soient à distance

mutuelle $> 2C'\epsilon^n$. Il reste donc au moins $C^{-1}(C''\epsilon^n)^{-d}$ n-cubes dans $G(n)$ à la fin de l'étape A. Alors, si α est assez petit, $\sigma(G(n)) \geqslant \alpha$. On finit la construction comme quand (i) s'est produit en premier.

Finalement, dans le cas où aucun des temps d'arrêts ne se produit, nous pouvons nous permettre de jeter les $G(n)$, et de garder $E = \cap [E(n) \setminus G(n)]$. L'inégalité $\sigma(E) \geqslant \sigma(E_0)/3$ est due à ce que ni (i) ni (iii) ne se sont produits.

Il nous reste encore à vérifier que, dans les trois cas, la restriction de z à E est bi-lipschitzienne. Lors des étapes A et E de notre processus d'élimination, nous nous sommes assurés que si les n-cubes Q et R sont encore dans E_k à la fin de la n-ième génération, et sont tels que dist $(Q, R) \geqslant C_3\epsilon^n$, alors on a dist $(z(Q), z(R)) \geqslant \epsilon^n$. Donnons-nous deux points $x \neq y$ dans E. Soit m le plus petit entier tel que $|x - y| > 2C_3\epsilon^m$. Dans le cas où aucun temps d'arrêt ne se produit, x et y sont dans deux m-cubes lors de la génération m (parce qu'ils ne sont pas dans $G(m-1)$), et comme la distance entre ces cubes est $\geqslant C_3\epsilon^m$, il vient $|z(x) - z(y)| \geqslant \epsilon^m \geqslant c\epsilon^{-1}|x - y|$. Ce raisonnement reste valable dans les deux autres cas, à condition que le temps d'arrêt ne se soit pas encore produit avant la génération m, et que x et y soient dans des cubes de $E(m)$ différents. Si aucun des temps d'arrêt ne s'est produit avant la génération m et si x et y sont dans un même cube R de $E(m)$, alors ce cube est un des cubes de $G(m-1)$ et l'inégalité $|z(x) - z(y)| \geqslant M_0^{-1}|x - y|$ vient de la définition de $E(R)$. Il ne reste plus que le cas où l'un des temps d'arrêt s'est produit avant la génération m (disons, à la génération $n < m$). Dans ce cas, x et y sont tous les deux dans des cubes de $G(n)$, et l'on peut supposer grâce à l'argument qui précéde que ces cubes sont différents. Comme $|x - y| \leqslant 2C_3\epsilon^{m-1}$, ces cubes sont tous les deux des $(m-1)$-cubes (donc en fait, $n = m - 1$), et l'on s'est assuré lors de l'étape A que leurs images étaient distantes d'au moins ϵ^{m-1}. Dans ce cas aussi, donc, $|z(x) - z(y)| \geqslant c\epsilon^{-1}|x - y|$. Ainsi, la restriction de z à E est M-bi-lipschitzienne pour un certain M, et nous avons fini de démontrer le Lemme 1.

Nous voulons maintenant déduire le Théorème 1 du Lemme 1. Notons d'abord que, par simple changement d'échelle, le théorème est encore valable pour tout m-cube I_0 (même si $m \neq 0$) dès qu'il est vrai pour tout 0-cube I_0. Ainsi, le Lemme permet de réduire la démonstration du théorème au cas où $|z(I_0)| \geqslant (1 + \eta_0)\delta$. Une nouvelle application du Lemme 1, avec une nouvelle valeur $\delta_1 = (1 + \eta_0)\delta$ de δ, permet de se réduire au cas où $|z(I_0)| \geqslant (1 + \eta_0)^2\delta$. Après un nombre $N \leqslant C \operatorname{Log}(C_0\delta^{-1})/\operatorname{Log}(1 + \eta_0)$ d'applications du Lemme 1, nous pouvons nous contenter de prouver le théorème lorsque $|z(I_0)| \geqslant 2C_0^3$. (Tout cela est valable à condition que ϵ soit assez petit pour que l'on puisse appliquer le Lemme 1 chaque fois.) Or le théorème est trivial lorsque $|z(I_0)| \geqslant 2C_0^3$, parce que l'hypothèse n'en est jamais satisfaite (à cause de (1), (4) et (7)). Nous avons donc entièrement prouvé le Théorème 1.

B. Applications

1. Le cas de $\mathbb{R}^d$

Dans le cas où Σ est $\mathbb{R}^d$ et σ la mesure de Lebesgue, la situation est un peu plus simple que dans le cas général. Les conditions (1), (2), (7) et (Q) sont trivialement satisfaites pour toute fonction lipschitzienne z (on prend pour n-cubes les cubes ϵ^{-1}-adiques, où ϵ est l'inverse d'un entier assez grand). Nous allons aussi prouver que (9) est toujours vrai, ce qui donnera le résultat suivant.

Théorème 2. Soit I_0 *le cube unité de* $\mathbb{R}^d$. *Soit* $z: I_0 \to \mathbb{R}^d$ *une fonction lipschitzienne. On suppose que* $|z(I_0)| \geqslant \delta$ *pour un* $\delta > 0$. *Alors on peut trouver un ensemble compact* $E \subset I_0$, *tel que* $|E| > \theta$ *et, pour tous* x, y *dans* E,

$$|z(x) - z(y)| \geqslant M^{-1}|x - y|.$$

Les constantes $\theta > 0$ *et* M *ne dépendent que de* δ *et de la constante de Lipschitz de* z.

Pour prouver le théorème, on peut bien sûr se ramener au cas où z est 1-lipschitzienne. Nous ferons appel au Théorème 1, avec pour σ la mesure de Lebesgue. Les hypothèses (1), (2), (Q) et (7) sont satisfaites, (8) est l'hypothèse du Théorème 2, de sorte qu'il ne nous reste plus qu'à vérifier que (9) est toujours vrai. Nous utiliserons le lemme qui suit.

Lemme 4. *Pour tout* $\delta > 0$, *il existe* $\tau > 0$ *et* $\gamma > 0$ *tels que l'on ait le résultat suivant. Soit* Q *un cube de* $\mathbb{R}^d$ *et* $z: Q \to \mathbb{R}^d$ *une fonction* 1*-Lipschitzienne. Supposons que* $\partial(Q) = |z(Q)|/|Q| \geqslant \delta$. *Alors l'une des affirmations suivantes est vraie*:

(30) $z(Q)$ *contient une boule de centre* $z(v)$ *et de rayon* τr (*où* v *est le centre de* Q *et* r *en est le côté*);

(31) *on peut trouver un cube* R *de côté* $\geqslant \tau r$ *tel que* $\partial(R) \geqslant (1 + \gamma)\, \partial(Q)$.

Supposons pour le moment que nous ayions prouvé le Lemme 4, et montrons la propriété (9). Donnons-nous un $n \geqslant 0$, un point $x \in \mathbb{R}^d$, une constante C_2, et soit T l'union des n-cubes qui rencontrent $B(x, C_2\epsilon^n)$. Supposons qu'aucune des deux conclusions de (9) ne soit vraie. Nous voulons montrer que si C_2 est assez grand par rapport à C_1, et si η est assez petit, alors on a une contradiction.

D'abord, on peut supposer par changement d'échelle que $n = 0$. On peut aussi supposer que C_1 est aussi grand que l'on veut. Soit S un cube, centré en un point $v \in \mathbb{Z}^d$ tel que $|v - x| \leqslant d$, et de côté $2[d^{-1}C_2]$ (en notant $[\bullet]$ la partie entière). Il est clair que $z(S)$ ne contient pas la boule centrée en $z(v)$ et de rayon $C_1 + d$. D'autre part, puisque T ne vérifie pas la seconde conclusion et S est une union de 0-cubes,

$$|z(S)| \leqslant (1 + 2\eta)|S|\partial(T),$$

et aussi

$$|z(T \setminus S)| \leqslant (1 + 2\eta)|(T \setminus S)|\partial(T),$$

d'où

$$\begin{aligned} |z(S)| &\geqslant |z(T)| - |z(T \setminus S)| \\ &\geqslant [|T| - (1 + 2\eta)|(T \setminus S)|]\partial(T) \\ &\geqslant (1 - C\eta)|S|\partial(T). \end{aligned}$$

Par conséquent, $\partial(S) \geqslant \delta/3$ et, si R est un 0-cube contenu dans S, alors

$$\partial(R) \leqslant (1 + 2\eta)\partial(T) \leqslant (1 + 2\eta)(1 - C\eta)^{-1}\partial(S).$$

Soit maintenant R un cube quelconque contenu dans S, mais dont le côté est $\geqslant N$ (où N sera décidé dans un instant). On obtient, en approchant R par une union de 0-cubes contenus dans R,

$$|z(R)| \leqslant (1 + 2\eta)(1 - C\eta)^{-1}\partial(S)|R| + C|R|^{(d-1)/d} \leqslant (1 + C\eta)\partial(S)|R|$$

si l'on choisit $N \geqslant \eta^{-1}$, par exemple. En résumé, si r est le côté de S, tous les cubes de côté $\geqslant [CC_2^{-1}\eta^{-1}]r$ contenus dans S sont tels que

$$|z(R)| \leqslant (1 + C\eta)|R|\partial(S),$$

et d'autre part $z(S)$ ne contient pas la boule de centre $z(v)$ et de rayon $[CC_1C_2^{-1}]r$. On obtient une contradiction avec le Lemme 4 (appliqué avec $\delta/3$, avec l'hypothèse un peu plus faible que z est C_0-lipschitzienne, et après changement d'échelle) dès que $C\eta \leqslant \gamma$ et que C_2 est assez grand pour que $[CC_2^{-1}\eta^{-1}] + [CC_1C_2^{-1}] \leqslant \tau$. Donc (9) est bien une conséquence du Lemme 4.

Pour finir de prouver le Théorème 2, il ne nous reste plus qu'à établir le Lemme 4. On peut voir ce lemme comme une estimation de distortion, que nous allons prouver par contradiction et par compacité. L'idée est, grossièrement, que si le lemme n'est pas vrai, alors on peut trouver une application lipschitzienne z, telle que $|z(A)| = \delta|A|$ pour tout A (car (31) est faux), mais qui n'est pas un homéomorphisme (car (30) n'a pas lieu).

Par translation et changement d'échelle, on peut ramener la preuve du lemme au cas où Q est le cube unité I_0. On peut aussi supposer que $\partial(I_0) = \delta$ (sinon, on peut toujours considérer $z^*(x) = [\partial(I_0)^{-1}\delta]^{1/d} z(x)$).

Supposons que l'on ait une suite z_n de fonctions 1-lipschitziennes, définies sur le cube unité I_0, telles que $|z_n(I_0)| = \delta$, telles que la boule de centre $z_n(v)$ (si v est le centre de I_0) et de rayon $1/n$ ne soit pas contenue dans $z_n(I_0)$, et telles que, pour tout cube $R \subset I_0$ de côté $\geqslant 1/n$, on ait $|z_n(R)| \leqslant (1 + 1/n)\delta$. On veut arriver à une contradiction. Quitte à extraire une sous-suite, on peut supposer que la suite z_n converge uniformément vers une fonction z (qui est donc, elle aussi, 1-lipschitzienne).

Lemme 5. *Si $R \subset I_0$ est un cube (compact),* $\limsup |z_n(R)| \leqslant |z(R)|$.

DÉMONSTRATION. Pour tout $\epsilon > 0$, on peut trouver un ouvert U contenant $z(R)$, et tel que $|U| \leqslant |z(R)| + \epsilon$. Comme dist $(z(R), U^c) > 0$, $z_n(R)$ est contenu dans U pour n assez grand, et on en déduit le lemme.

Lemme 6. *Si R est un cube dyadique contenu dans I_0,* $|z(R)| \geqslant \delta|R|$.

DÉMONSTRATION. Compte tenu du Lemme 5, il suffit de montrer que, pour tout $\epsilon > 0$, $|z_n(R)| \geqslant \delta|R| - \epsilon$ pour n assez grand. Ecrivons $I_0 = \cup R_j$, où l'union porte sur tous les cubes dyadiques de la même taille que R et, disons, où $R = R_0$. Alors,

$$\delta = |z_n(I_0)| \leqslant \sum |z_n(R_j)| \leqslant |z_n(R)| + \sum_{j \neq 0} (1 + 1/n)\delta|R_j|,$$

donc $|z_n(R)| \geqslant \delta|R| - 1/n$, d'où on déduit le lemme.

Notons que, comme la fonction z est lipschitzienne, elle est différentiable presque-partout (voir [ST], p. 242).

Lemme 7. *Soit x_0 un point de différentiabilité de z, et $Jz(x_0)$ le déterminant Jacobien de z en x_0. Alors* $|Jz(x_0)| = \delta$.

DÉMONSTRATION. Si x_0 est un point de différentiabilité de z, le rapport $|z(R)| / |R|$ a une limite supérieure $\leqslant |Jz(x_0)|$ lorsque le cube R tend vers x_0. En particulier, à cause du Lemme 6, $|Jz(x_0)| \geqslant \delta$. Supposons que l'inégalité soit stricte. Donnons-nous $\epsilon > 0$, et un cube dyadique $R \subset I_0$ contenant x_0 et assez petit pour que, sur R, $|z(x) - z(x_0) - M(x - x_0)| \leqslant \epsilon|x - x_0|$ (où M est l'application linéaire tangente à z). Soit B l'image du cube $(1 - C\epsilon)R$ par l'application $x \to z(x_0) + M(x - x_0)$; pour tout point $Z \in B$, considérons l'applica-

tion $t \to (z(t) - Z)|z(t) - Z|^{-1}$, du bord ∂R de R dans la sphère unité. Si C est assez grand et ϵ assez petit,

$$|z(t) - Z| \geqslant |z(x_0) + M(t - x_0) - Z| - |z(t) - z(x_0) - M(t - x_0)| > 0$$

pour tout $t \in \partial R$, de sorte que cette application est bien définie, et de degré ± 1 (par continuité du degré, et parce que le résultat est vrai lorsqu'on remplace $z(t)$ par $z(x_0) + M(t - x_0)$). Comme dist $(B, z(\partial R) > 0$, l'application

$$t \to (z_n(t) - Z)|z_n(t) - Z|^{-1}$$

est encore définie et de degré ± 1 pour n assez grand et pour tout $Z \in B$. On en déduit que Z est dans $z_n(R)$, car sinon l'application

$$t \to (z_n(t) - Z)|z_n(t) - Z|^{-1}$$

pourrait être contractée en une constante (en utilisant une contraction de $z_n(R)$), ce qui est impossible. Comme $|Jz(x_0)| > \delta$, on peut choisir ϵ assez petit pour que $|B| > \delta|R|$, et ensuite n assez grand pour que

$$|z_n(R)| \leqslant (1 + 1/n)\delta|R| < |B|.$$

Ceci contredit le fait que $B \subset z_n(R)$; on en déduit le lemme.

Lemme 8. *Ou bien $Jz(x) = \delta$ p.p., ou bien $Jz(x) = -\delta$ p.p.*

Supposons qu'en un point x_0 de l'intérieur de I_0, z soit différentiable et $Jz(x_0) = \delta$; on veut montrer que Jz garde ce signe. Soit Q un petit cube centré en x_0; si Q est assez petit, l'application F, de ∂Q dans la sphère unité, définie par

$$F(t) = (z(t) - z(x_0))|z(t) - z(x_0)|^{-1},$$

est de degré 1. Nous aurons besoin du lemme suivant.

Lemme 9. *Si x_0 est un point différentiabilité de z intérieur à I_0 et si $y \neq x_0$, alors $z(y) \neq z(x_0)$.*

Nous prouverons ce lemme un peu plus tard; pour le moment, finissons la démonstration du Lemme 8. Le Lemme 9 nous assure que $z(t) - z(x_0)$ ne s'annule pas pour $t \neq x_0$. On peut par conséquent déformer la sphère topologique ∂Q en n'importe quelle sphère $S \subset I_0$ ayant x_0 dans son intérieur, et l'application G, de S dans la sphère unité, telle que

$$G(t) = [z(t) - z(x_0)]|z(t) - z(x_0)|^{-1} \quad \text{pour} \quad t \in S,$$

est encore de degré 1.

Si y_0 est un autre point intérieur où z est différentiable et si l'on peut trouver un chemin qui joigne $z(x_0)$ à $z(y_0)$ sans rencontrer $z(S)$, alors le degré de l'application H définie sur S par

$$H(t) = (z(t) - z(y_0))|z(t) - z(y_0)|^{-1}$$

est encore 1. En appliquant le début de notre argument à y_0 au lieu de x_0, on voit que $Jz(y_0) = \delta$.

Supposons maintenant que l'on puisse trouver un autre point intérieur x', tel que $Jz(x') = -\delta$. Par connexité, le signe de $Jz(x)$, restreint aux points de différentiabilité de z, ne peut pas être localement constant sur l'intérieur de I_0, et donc on peut trouver un point T qui soit dans l'adhérence à la fois de l'ensemble U des points de différentiabilité x tels que $Jz(x) = \delta$ et de l'ensemble V des points de différentiabilité y tels que $Jz(y) = -\delta$. Donnons-nous $\epsilon > 0$, et soient $x \in U \cap B(T, \epsilon)$ et $y \in V \cap B(T, \epsilon)$. Pour tout rayon r compris entre ϵ et dist $(T, \partial I_0)$, l'image par z du segment qui joint x à y rencontre l'image $z(S(T, r))$ de la sphère centrée en T et de rayon r, car sinon on pourrait appliquer la remarque précédente et on aurait $Jz(x) = Jz(y)$. Prenons des valeurs de r espacées de ϵ (on peut en trouver $(C\epsilon)^{-1}$, où C ne dépend que de T); pour chacune, on peut trouver un point t_r de $S(T, r)$ tel que $z(t_r)$ soit dans $B(z(T), \epsilon)$. Par conséquent, on peut trouver $(C\epsilon)^{-1}$ cubes dyadiques disjoints Q_r de côté $\geqslant C^{-1}\epsilon$ tels que $z(Q_r)$ soit contenu dans $B(z(T), 2\epsilon)$. Si n est assez grand, on a encore

$$z_n(Q_r) \subset B(z(T), 3\epsilon) \quad \text{et} \quad \left|z\left[\bigcup Q_r\right]\right| \leqslant (6\epsilon)^d,$$

alors que

$$\left|\bigcup Q_r\right| \geqslant C^{-1}\epsilon^{d-1};$$

si ϵ est assez petit, alors

$$\left|z\left[\bigcup Q_r\right]\right| \leqslant (\delta/2)\left|\bigcup Q_r\right|.$$

On peut conclure comme au Lemme 6: si

$$K = I_0 \setminus \bigcup Q_r,$$

on a

$$\delta \leqslant \left|z\left(\bigcup Q_r\right)\right| + |z(K)| < (\delta/2)\left|\bigcup Q_r\right| + (1 + 1/n)|K| < \delta$$

si n est assez grand. Finalement, le point T n'existait pas, et le Jacobien garde un signe constant.

Pour finir la démonstration du Lemme 8, il ne nous reste plus qu'à prouver le Lemme 9. Donnons-nous donc x_0 et y comme dans ce lemme. Choisissons

un cube Q centré en x_0, de diamètre $\leqslant |x-y|/2$, et sur lequel on ait

$$|z(x) - z(x_0) - M(x - x_0)| \leqslant \epsilon |x - x_0|$$

(où M est l'application linéaire tangente). Si ϵ est assez petit et si le point Z est assez près de $z(x_0)$, alors l'application $t \to (z(t) - Z)|z(t) - Z|^{-1}$, du bord ∂Q de Q dans la sphère unité, est de degré ± 1. La même chose vaut pour z_n si n est assez grand, et donc, par la même argument que lors de la preuve du Lemme 7, Z est dans $z_n(Q)$. On en déduit que, si Q' est un cube assez petit contenant y, alors $z_n(Q') \subset z_n(Q)$ pour n assez grand. La même chose est encore vraie si l'on remplace q par un cube dyadique plus grand qui ne rencontre pas Q', et Q' par un cube dyadique plus petit. On a encore une contradiction parce que

$$\delta = |z_n(I_0)| \leqslant \sum |z_n(R_j)|$$

(où la somme porte sur tous les cubes dyadiques R_j de la même taille que Q', sauf justement Q'), d'où on déduit que

$$\delta \leqslant \sum (1 + 1/n)\delta |R_j| < \delta$$

si n est assez grand.

Les Lemmes 8 et 9 sont enfin prouvés, et nous sommes maintenant en mesure de conclure. Comme z est ACL (et même lipschitzienne), et que sa dérivée est bornée et a presque-partout un Jacobien égal à δ (ou $-\delta$, auquel cas on peut toujours changer le signe d'une coordonnée), z est quasirégulière, et par conséquent est ouverte (voir [R], ou [MRV]).

Vérifions que z est un homéomorphisme sur l'intérieur de I_0, c'est à dire que $z(x) \neq z(y)$ pour $x \neq y$. Si ça n'était pas le cas, on pourrait trouver une boule B contenant y, mais à distance positive de x, telle que $z(B)$ contienne un voisinage de $z(x)$. En particulier, $z(B)$ contiendrait $z(x_0)$, où x_0 est un point de différentiabilité de z et n'est pas dans B. Le Lemme 9 nous assure que cela n'est pas possible. Donc, z est un homéomorphisme (et même quasiconforme).

Montrons pour finir que cela contredit notre supposition que la boule de centre $z_n(v)$ (où v est le centre de I_0) et de rayon $1/n$ n'est pas contenue dans $z_n(I_0)$. L'application

$$F(t) = (z(t) - z(v))|z(t) - z(v)|^{-1},$$

définie sur le bord de I_0, est de degré ± 1; soit $\epsilon > 0$ tel que $|z(t) - z(v)| > \epsilon$ pour tout t. Si n est assez grand pour que $|z_n(t) - z(t)| < \epsilon/4$ et si $u \in B(z(v), \epsilon/2)$, l'application définie par

$$G(t) = (z_n(t) - u)|z_n(t) - u|^{-1}$$

est encore de degré 1, d'où l'on conclut que u est dans l'image de z_n ce qui est une contradiction si n est assez grand).

Nous avons donc fini de prouver le Lemme 4 et, par conséquent, le Théorème 2.

Remarque. Le Théorème 2 peut être généralisé de la manière suivante. Donnons-nous un poids $\omega \in A_\infty(\mathbb{R}^d)$ dans la classe de Muckenhoupt. Pour x, y dans $\mathbb{R}^d$, soit

$$D(x, y) = \{\omega[B((x + y)/2, |x - y|)]\}^{1/d} = \left\{\int_{B((x+y)/2, |x-y|)} \omega(x)\, dx\right\}^{1/d}.$$

Théorème 2 Bis. *Soit $z: I_0 \to \mathbb{R}^d$ une fonction telle que*

$$|z(x) - z(y)| \leqslant CD(x, y)$$

(autrement dit, telle que $|\nabla f(x)| \leqslant C\omega(x)^{1/d}$), et telle que $|z(I_0)| \geqslant \delta\omega(I_0)$. Alors on peut trouver un ensemble compact $E \subset I_0$, tel que $\omega(E) > \theta\omega(I_0)$, et tel que

$$|z(x) - z(y)| \geqslant M^{-1}D(x, y)$$

pour x, y dans E. Les constantes θ et M ne dépendent que de C, δ, et de ω.

Nous nous contenterons d'esquisser la démonstration de ce résultat. On veut appliquer le Théorème 1 à $\mathbb{R}^d$, muni de la mesure $\omega\, dx$, et de $D(x, y)$ (le fait que D ne soit pas une distance n'est pas gênant). On choisit cette fois pour n-cubes des cubes dyadiques Q, pas nécessairement de rayon ϵ^n, mais tels que $\omega(Q) \approx \epsilon^{nd}$. Les propriétés (1), (2) et (7) sont des conséquences faciles du fait que $\omega \in A_\infty$, et le même argument que dans le cas où $\omega = 1$, combiné avec les propriétes élémentaires de A_∞, montre qu'on peut se contenter de prouver l'analogue suivant du Lemme 4.

Lemme 4 Bis. *On peut trouver $\tau > 0$ et $\gamma > 0$ (qui ne dépendent que de ω et de δ) tels que, si Q est un cube, $z: Q \to \mathbb{R}^d$ est une fonction telle que $|z(x) - z(y)| \leqslant D(x, y)$ pour x, y dans Q, et si $\partial(Q) = |z(Q)|/\omega(Q) \geqslant \delta$, l'une des affirmations suivantes est vraie:*

- *—$z(Q)$ contient une boule de centre $z(v)$ et de rayon $\tau\omega(Q)^{1/d}$ (où v est le centre de Q);*
- *—on peut trouver un cube R de côté $\geqslant \tau r$ (où r est le côté de Q) tel que $\partial(R) \geqslant (1 + \gamma)\partial(Q)$.*

On peut reprendre la preuve du Lemme 4, pratiquement sans modification. On commence par remarquer que, si les poids ω_n sont uniformément dans A_∞

et si $\omega_n(I_0) = 1$ pour tout n, alors on peut extraire une sous-suite des ω_n qui converge faiblement vers un poids $\omega \in A_\infty$. Si z_n est une suite de fonctions telle que $|z_n(x) - z_n(y)| \leqslant D_{\omega_n}(x, y)$ pour tout n, alors on peut extraire une sous-suite, convergeant uniformément vers une fonction z, qui vérifie encore $|z(x) - z(y)| \leqslant D(x, y)$, et qui est donc différentiable presque-partout (voir [ST], p. 242). On peut utiliser les mêmes arguments de degré pour montrer que $Jz(x_0) = \delta\omega(x_0)$ presque partout (ou bien $Jz(x_0) = -\delta\omega(x_0)$ p.p.). On en déduit pareillement que z est un homéomorphisme, ce qui donne la contradiction souhaitée, le Lemme 4 Bis et le Théorème 2 Bis.

2. Surfaces Régulières

Nous nous proposons d'appliquer le Théorème 2 pour trouver des morceaux de graphes lipschitziens dans toute «surface régulière». Rappelons ce que nous entendons par surface régulière.

Définition 1. *Soient $N \geqslant d \geqslant 1$ des entiers. On appellera «paramétrage régulier» une application $z: \mathbb{R}^d \to \mathbb{R}^N$ telle que, pour une constante $C_0 \geqslant 0$, z est C_0-lipschitzienne, et*

(32) *pour tout $R > 0$ et toute boule $B \subset \mathbb{R}^N$ de rayon R,*

$$|\{x \in \mathbb{R}^d : z(x) \in B\}| \leqslant CR^d.$$

L'image de $\mathbb{R}^d$ par z sera appelée une «surface régulière».

Proposition 1. *Soit $z: \mathbb{R}^d \to \mathbb{R}^N$ un «paramétrage régulier». Alors il existe des constantes M et $\theta > 0$ telles que, si $Q \subset \mathbb{R}^d$ est un cube, on puisse trouver un sous-ensemble compact $E \subset Q$ tel que $|E| \geqslant \theta|Q|$, une transformation orthogonale U sur $\mathbb{R}^N$, et une fonction M-lipschitzienne $F: \mathbb{R}^d \to \mathbb{R}^{N-d}$ telles que $z(E)$ soit contenu dans l'image par U du graphe de F, et que la restriction de z à E soit M-bi-lipschitzienne.*

Avant de démontrer la proposition, rappelons comment on peut utiliser un tel résultat (appelons cela une décomposition de $z(\mathbb{R}^d)$ en graphes lipschitziens). Donnons-nous un noyau $K(z)$, défini sur $\mathbb{R}^N \setminus \{0\}$, impair, homogène de degré $-d$, et indéfiniment dérivable. Alors, si z est régulière, on peut définir un opérateur borné sur $L^2(\mathbb{R}^d)$ par

$$Tf(x) = \text{v.p.} \int K(z(x) - z(y)) f(y)\, dy. \tag{33}$$

La démonstration de la continuité de T donnée dans [D2] utilise des décompositions semblables à celle de la Proposition 1, mais l'image de z n'y est pas

directement décomposée en graphes lipschitziens, mais en des classes successives de surfaces, dont chaque élément se décompose dans la classe suivante, la dernière classe étant celle des graphes lipschitziens. La Proposition 4 de [D2] dit justement que si $z(\mathbb{R}^d)$ admet une décomposition dans une classe de «surfaces régulières» $z^*(\mathbb{R}^d)$ (avec une constante C_0 qui ne dépend pas de z^*), la continuité de l'opérateur T défini par (33) découle de la continuité, uniforme en z^*, des opérateurs T^* associés aux z^*. On se ramène ainsi, en plusieurs étapes, au cas des graphes lipschitziens, pour lesquels la continuité de T est une conséquence assez facile du théorème de Coifman, McIntosh, et Meyer ([CMM]) sur la continuité de l'intégrale de Cauchy sur les graphes lipschitziens (voir par exemple [D2], Lemme 16).

La Proposition 1, en conjonction avec la Proposition 4 de [D2] et [CMM], donne une démonstration plus directe de la continuité de T lorsque z est un «paramétrage régulier». D'un autre côté, cette nouvelle preuve de la continuité de T a l'inconvénient d'utiliser (par deux fois) un argument de compacité. Nous n'insistons pas plus.

Pour prouver la proposition, considérons la fonction $Z = p \circ z$, où p est la projection orthogonale de $\mathbb{R}^N$ sur un sous-espace de dimension d de la forme $\{x \in \mathbb{R}^N, x_{i_1} = x_{i_2} = \cdots = x_{i_{N-d}} = 0\}$, où $1 \leqslant i_1 < i_2 < \cdots < i_{N-d} \leqslant N$ est un choix de $N - d$ coordonnées sur $\mathbb{R}^N$. Le lemme suivant va nous permettre d'appliquer le Théorème 2 à Z.

Lemme 10. *Si z est un paramétrage régulier et $Q \subset \mathbb{R}^d$ est un cube, on peut trouver $N - d$ coordonnées $1 \leqslant i_1 < \cdots < i_{N-d} \leqslant N$ telles que $|p(z(Q))| \geqslant \delta|Q|$, où δ ne dépend que de C_0.*

Supposons que nous ayions prouvé ce lemme. Le Théorème 2 nous donne un ensemble $E \subset Q$ tel que $|E| \geqslant \theta|Q|$, et tel que la restriction de Z à E soit bi-lipschitzienne. Cela signifie exactement que le paramétrage de $z(E)$ par sa projection est bi-lipschitziennement équivalent au paramétrage par z, et aussi que $z(E)$ est (après changement de l'ordre des coordonnées sur $\mathbb{R}^N$) contenu dans un graphe lipschitzien.

Pour prouver le lemme, on utilisera encore un argument de compacité. Supposons que l'on puisse trouver C_0 tel que δ n'existe pas. Cela veut dire que l'on peut trouver une suite de paramétrages z_n, réguliers avec la constante C_0, et de cubes Q_n, tels que pour tout choix de $i_1, \ldots, i_{N-d}$, on ait $|p(z_n(Q_n))| \leqslant 1/n$. Par changement d'échelle, on se ramène au cas où Q_n est, pour tout n, le cube unité. Quitte à extraire une sous-suite, on peut supposer que la suite z_n converge, uniformément sur le cube unité, vers une fonction z. Il est clair que z est encore un «paramétrage régulier». Montrons que, si x_0 est un point de différentiabilité de z, la dérivée de z en x_0 est de rang d. Sinon, pour ϵ assez petit, tout point de l'image de $B(x_0, \epsilon)$ serait à distance $\leqslant t(\epsilon)\epsilon$ d'une boule de

dimension $\leqslant d-1$ et de rayon $\leqslant C\epsilon$, où $t(\epsilon)$ tend vers 0 avec ϵ. On pourrait alors recouvrir $z(B(x_0, \epsilon))$ par moins de $Ct(\epsilon)^{-(d-1)}$ boules de rayon $Ct(\epsilon)\epsilon$ et, comme z est régulière, la mesure de l'image réciproque par z de chacune de ces boules serait $\leqslant C[t(\epsilon)\epsilon]^d$. La contradiction vient de ce que $B(x_0, \epsilon)$ devrait être contenue dans la réunion de ces images réciproques, dont la masse totale est pourtant inférieure à $Ct(\epsilon)\epsilon^d$. Donc, la dérivée de z en x_0 est de rang d, et l'on peut trouver des coordonnées $i_1, \ldots, i_{N-d}$ telles que la dérivée de $Z = p \circ z$ en x_0 soit bijective. Le même argument de degré que dans la démonstration du Lemme 7 montre que, pour n assez grand, $p \circ z_n(Q)$ contient un voisinage fixe de $p \circ z(x_0)$. Ceci est contradictoire avec notre supposition que $|p \circ z_n(Q)| \leqslant 1/n$. Nous avons fini de prouver le Lemme 10 et, du même coup, la Proposition 1.

Remarque. On peut appliquer le Théorème 2 Bis, et la démonstration qui précéde, pour prouver directement la continuité de l'opérateur d'intégrale singulière défini par le noyau K sur une «surface ω-régulière».

Nous appellerons «surface ω-régulière» l'image de $\mathbb{R}^d$ par une application $z: \mathbb{R}^d \to \mathbb{R}^N$, telle qu'il existe un poids $\omega \in A_\infty(\mathbb{R}^d)$ pour lequel

$$|z(x) - z(y)| \leqslant C_0 D(x, y)$$

pour $x, y \in \mathbb{R}^d$, où

$$D(x, y) = \{\omega[B((x+y)/2, |x-y|)]\}^{1/d},$$

et pour lequel $\omega[z^{-1}(B(w, R))] \leqslant C_0 R^d$ pour tout $w \in \mathbb{R}^N$ et tout $R > 0$. Comme précédemment, on dira que z est un «paramétrage ω-régulier».

On vérifie aisément que la classe des paramétrages ω-réguliers est stable par passage à la limite, de sorte que l'argument de compacité çi-dessus montre aussi que, si $Q \subset \mathbb{R}^d$ est un cube et z est un paramétrage ω-régulier, on peut trouver $N-d$ coordonnées $1 \leqslant i_1 < \cdots < i_{N-d} \leqslant N$ telles que, si p est la projection orthogonale sur le d-plan défini par $x_{i_1} = \cdots = x_{i_{N-d}} = 0$, alors $|p(z(Q))| \geqslant \delta\omega(Q)$. On peut appliquer le Théorème 2 Bis au cube Q et à l'application $Z = p \circ z$; on obtient la variante suivante de la Proposition 1.

Proposition 1 Bis. *Si $z: \mathbb{R}^d \to \mathbb{R}^N$ est un paramétrage ω-régulier, il existe des constantes M et $\theta > 0$ telles que, si $Q \subset \mathbb{R}^d$ est un cube, on peut trouver un sous-ensemble compact $E \subset Q$ tel que $\omega(E) \geqslant \theta\omega(Q)$, une transformation orthogonale U sur $\mathbb{R}^N$, et une fonction M-lipschitzienne $F: \mathbb{R}^d \to \mathbb{R}^{N-d}$ tels que $z(E)$ soit contenu dans l'image par U du graphe de F, et que la restriction à E de $\omega(Q)^{-1/d} z(x)$ soit M-bilipschitzienne.*

En fait, le Théorème 2 Bis donne un compact $E' \subset Q$ tel que $\omega(E') \geqslant \theta'\omega(Q)$ et tel que $|Z(x) - Z(y)| \geqslant D(x, y)/M'$ pour x, y dans E'. On utilise alors le fait

que, comme $\omega \in A_\infty$, si $F = \{x \in Q$: il existe $y \in Q$ tel que $D(x, y) > \lambda\omega(Q)^{1/d}$ ou $D(x, y) < \lambda^{-1}\omega(Q)^{1/d}\}$, alors $\omega(F)/\omega(Q)$ est aussi petit que l'on veut pourvu que λ soit assez grand. On choisit $E = E' \cap F^c$.

La Proposition 1 Bis, en conjonction avec la Proposition 4 de [D2], donne une nouvelle démonstration de la continuité sur $L^2(\mathbb{R}^d, \omega\, dx)$ de l'opérateur T défini par

$$Tf(x) = \text{v.p.} \int K(z(x) - z(y)) f(y)\omega(y)\, dy.$$

Nous n'entrons pas dans les détails (on pourra en trouver certains dans [D2]).

3. Le cas d'une surface dans $\mathbb{R}^N$: Construction des n-cubes

Nous nous proposons d'appliquer le Théorème 1 dans le cas où Σ est contenu dans $\mathbb{R}^N$. La situation est alors un peu plus simple que dans le cas général, et nous allons montrer que dans ce cas, les hypothèses (Q) et (7) sont vérifiées automatiquement.

Lemme 11. *Donnons-nous un entier $N \geqslant d$. Soit σ une mesure de Radon positive sur $\mathbb{R}^N$ (muni de la distance euclidienne), et soit Σ son support. Si σ satisfait à* (1) *et* (2), *alors les hypothèses* (Q) *et* (7) *sont vérifiées (peut-être avec une constante C_0 différente).*

Commençons par vérifier (7). Soit $A \subset \Sigma$ un Borélien. Donnons-nous un ouvert O de $\mathbb{R}^N$ contenant A, et tel que $\sigma(O) \leqslant 2\sigma(A)$. Pour chaque point x de A, donnons-nous une boule ouverte $B(x)$ contenue dans O et centrée en x. Le Théorème de Besicovitch (voir par exemple [G], p. 2) nous permet de trouver un sous-recouvrement de A par les $B_i = B(x_i)$, $i \in I$, tel que $\sum_i \chi_{B_i}(y) \leqslant C$ pour tout y. Alors, en notant $R(B_i)$ le rayon de B_i,

$$\begin{aligned} |z(A)| &\leqslant \sum_i |z(B_i)| \leqslant C \sum_i R(B_i)^d \leqslant C \sum_i \sigma(B_i) \\ &\leqslant C \int \sum_i \chi_{B_i}(y)\, d\sigma(y) \\ &\leqslant C\sigma(O) \leqslant C\sigma(A), \end{aligned}$$

ce qui prouve (7).

Montrons maintenant que (Q) est vraie. Prenons pour ϵ l'inverse d'un entier. Nous allons commencer par construire une première approximation des 0-cubes, en modifiant légèrement les cubes de côté 1 à sommets entiers.

Nous remplacerons les hyperplans $\{x_i = k\}$, où $1 \leqslant i \leqslant N$ et $k \in \mathbb{Z}$, par des hypersurfaces $S_{i,k}$, qui seront définies à l'aide de fonctions $f_{i,k}$ par

$$S_{i,k} = \{x \in \mathbb{R}^N : x_i = f_{i,k}(x^i)\},$$

où l'on utilise la notation

$$x^i = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_N).$$

Montrons que, si C' et τ sont donnés à l'avance et si ϵ est assez petit, on peut choisir $f_{i,k}$ telle que

(34) $\|\nabla f_{i,k}\|_\infty \leqslant \tau$, $|f_{i,k}(x^i) - k| \leqslant \tau/10$ et
(35) pour tout cube T de côté 2, $\sigma(\{x \in T : \operatorname{dist}(x, S_{i,k}) \leqslant 2C'\epsilon\}) \leqslant \tau/2N$.

On va chercher $f_{i,k}$ sous la forme

$$f_{i,k} = k + g_1 + \cdots + g_{8^{N-1}},$$

où les fonctions g_m, $1 \leqslant m \leqslant 8^{N-1}$, sont définies de la manière qui suit. On se donne des représentants $a_1, \ldots, a_{8^{N-1}}$ de $\mathbb{Z}^{N-1}/8\mathbb{Z}^{N-1}$, et une fonction $\varphi(x^i)$, de classe C^1, telle que $\|\varphi\|_\infty \leqslant 1$, telle que $\varphi(x^i) = 1$ sur J (où J est le cube de centre 0 et de côté 4), et $\varphi(x^i) = 0$ hors de $2J$. Pour $1 \leqslant m \leqslant 8^{N-1}$, la fonction g_m sera de la forme

$$\sum_j \lambda_{j,m} \varphi(x^i - j - a_m),$$

où la somme porte sur $j \in 8\mathbb{Z}^{N-1}$, et les $\lambda_{j,m}$ seront tels que

$$|\lambda_{j,m}| \leqslant C_1^{-m} \tau^m, \tag{36}$$

pour une constante C_1 que l'on choisira plus tard. Supposons que nous ayons déjà construit $g_1, \ldots, g_{m-1}$. Montrons que l'on peut choisir $\lambda_{j,m}$ tel que, si $T_{m,j}$ est l'ensemble des $x \in \mathbb{R}^N$ tels que x^i est dans le cube de côté 4 centré en $j + a_m$, alors

(37) $\sigma(\{x \in T_{m,j} : |x_i - k - g_1(x^i) - \cdots - g_m(x^i)| \leqslant 3C_1^{-m-1}\tau^{m+1}\}) \leqslant \tau/2N$.

Grâce à la définition de φ, les ensembles

$$\{x \in T_{m,j} : |x_i - k - g_1(x^i) - \cdots - g_m(x^i)| \leqslant 3C_1^{-m-1}\tau^{m+1}\}$$

qui correspondent à des valeurs de $\lambda_{j,m}$ qui diffèrent de plus de $7C_1^{-m-1}\tau^{m+1}$ sont disjoints et, à cause de (1), on ne peut pas en trouver plus de $CC_0(\tau/2N)^{-1}$ qui ne vérifient pas l'inégalité souhaitée; on en déduit que, si

$$C_1^{-m}\tau^{-m} \geqslant 7C_1^{-m-1}\tau^{m+1}CC_0(\tau/2N)^{-1}$$

(c'est-à dire si $C_1 \geqslant CC_0$), on peut choisir $\lambda_{j,m}$ tel que (36) et (37) soient vrais. Faisons, pour tout m, un tel choix des $\lambda_{j,m}$.

Vérifions que la fonction $f_{i,k}$ que nous venons de définir satisfait à (34) et (35). On déduit de (36) que $|g_m(x^i)| \leqslant C_1^{-m}\tau^m$, ce qui donne (34) si C_1 est assez grand. De plus, on déduit de (37) que, pour tout m et tout j,

$$\sigma(\{x \in T_{m,j}: |x_i - f_{i,k}(x^i)| \leqslant C_1^{-m-1}\tau^{m+1}\}) \leqslant \tau/2N,$$

ce qui donne (35) si ϵ est assez petit.

Pour tout $n \geqslant 0$, on peut faire la même construction à l'échelle ϵ^n. On trouve, pour $1 \leqslant i \leqslant N$ et $k \in \mathbb{Z}$, une fonction $f_{n,i,k}$ telle que

$$\|\nabla f_{n,i,k}\|_\infty \leqslant \tau, \qquad |f_{n,i,k}(x^i) - k\epsilon^n| \leqslant \tau\epsilon^n/10 \quad \text{et,} \tag{38}$$

si $S_{n,i,k}$ est définie par l'équation $x_i = f_{n,i,k}(x_i)$,

(39) pour tout cube T de côté $2\epsilon^n$,

$$\sigma(\{x \in T: \operatorname{dist}(x, S_{n,i,k}) \leqslant 2C'\epsilon^n\}) \leqslant \tau\epsilon^{nd}/2N.$$

On peut maintenant définir, pour $n \in \mathbb{N}$, une famille $Q_1(n)$ d'ensembles ressemblant à des cubes de taille ϵ^n: ce sera la famille des ensembles

$$\begin{aligned} R &= R_{n,k(1),\ldots,k(N)} \\ &= \{x \in \mathbb{R}^N: f_{n,i,k(i)}(x^i) \leqslant x_i < f_{n,i,k(i)+1}(x^i) \text{ pour } 1 \leqslant i \leqslant N\}, \end{aligned}$$

où $(k(1), \ldots, k(N)) \in \mathbb{Z}^N$. On déduit de (38) que

(40) si $R \in Q_1(n)$, on peut trouver un cube T de côté ϵ^n tel que

$$(1 - \tau/2)T \subset R \subset (1 + \tau/2)T.$$

Regroupons les éléments de $Q_1(n)$ en paquets. Chaque paquet S sera la réunion de moins de C éléments de $Q_1(n)$, choisis de manière à ce que le diamètre de S soit $\leqslant C\epsilon^n$. Appellons $Q_0(n)$ l'ensemble des S. Grâce à (2), si C est assez grand par rapport à C_0, on peut choisir les paquets S de façon que

$$\sigma(S) \geqslant C^{-1}\epsilon^n \quad \text{pour tout} \quad S \in Q_0(n), \tag{41}$$

et que les $S \cap \Sigma$, $S \in Q_0(n)$, forment une partition de Σ. Faisons un tel choix de $Q_0(n)$.

Donnons-nous maintenant une application φ, de l'ensemble des $Q_0(n+1)$ dans l'ensemble des $Q_0(n)$ telle que, si $Q \in Q_0(n+1)$, $\varphi(Q)$ soit l'un des $R \in Q_0(n)$ tels que $R \cap Q \neq \varnothing$. Pour chaque n et chaque $R \in Q_0(n)$, on définit une suite d'ensembles R_m par $R_0 = R$, et $R_m = \cup Q$, où l'union porte sur tous les $Q \in Q_0(n+m)$ tels que $\varphi^m(Q) = R$. Il est clair que la distance de Hauss-

dorff entre R_m et R_{m+1} est inférieure à $C\epsilon^{m+n+1}$, donc la suite des adhérences $(\overline{R_m})$ converge (pour la topologie de Haussdorff) vers un compact que nous noterons $A(R)$, et dont la distance de Haussdorff à R est inférieure à $C\epsilon^{n+1}$. Les $A(R)$, lorsque R parcourt $Q_0(n)$, sont presque notre partition $Q(n)$; nous devons encore les modifier un peu pour qu'ils forment une partition de Σ. Nous savons déjà que leur union contient Σ tout entier: si $x \in \Sigma$, x est pour tout m dans l'un des R_m, où R parcourt un ensemble fini; alors x est encore dans $A(R)$, où R est tel que $x \in R_m$ pour une infinité de m.

Pour retirer certains points des $A(R)$ de manière cohérente, donnons-nous un ordre (que nous noterons $<$) sur chaque $Q_0(n)$, de telle sorte que si Q et Q' sont dans un $Q_0(n+1)$ et si $\varphi(Q) < \varphi(Q')$, alors $Q < Q'$. On pose, pour $R \in Q_0(n)$,

$$B(R) = A(R) \cap \left\{ \bigcup_{R' \in Q(n),\, R' < R} A(R) \right\}^c \tag{42}$$

(notons que l'union en question est localement finie). Montrons que la famille $Q(n) = \{B(R) \cap \Sigma : R \in Q_0(n)\}$ convient.

Pour chaque n, $Q(n)$ est une partition de Σ. D'autre part, si $S \in Q_0(n)$ et $m \geqslant 1$, alors $A(S) = \cup A(Q)$, où l'union porte sur les $Q \in Q_0(n+m)$ tels que $\varphi^m(Q) = S$. Montrons maintenant que, si $R \in Q_0(n+m)$, alors $B(R) \subset B(\varphi^m(R))$: il est certainement contenu dans $A(\varphi^m(R))$ à cause de la remarque précédente et, s'il recontrait $A(R')$ pour un $R' \in Q_0(n)$ tel que $R' < \varphi^m(R)$, alors il rencontrerait aussi $A(R'')$ pour un $R'' \in Q_0(n+m)$ tel que $\varphi^m(R'') = R'$; dans ce cas, grâce à la manière dont on a choisi l'ordre $<$, on aurait $R'' < R$, et $B(R)$ ne pourrait pas rencontrer $A(R'')$ (contradiction). On en déduit (3) parce que, puisque $B(R) \subset B(\varphi^m(R))$, il ne rencontre aucun autre $B(Q)$, où $Q \neq \varphi^m(R)$ est dans $Q_0(n)$.

Si $S \in Q_0(n)$, S est une union d'objets $R \in Q_1(n)$, et pour chaque R on peut trouver un T comme dans (40), alors $A(S)$ est la réunion d'un nombre $\leqslant C$ d'ensembles R' tels que $(1-\tau)T \subset R' \subset (1+\tau)T$ pour des (vrais) cubes T, car la distance de Haussdorff de S à $A(S)$ est inférieure à $C\epsilon^{n+1}$. De plus, si T est associé à l'un des morceaux R de S et si $S' \neq S \in Q_0(n)$, alors $A(S')$ ne rencontre pas $(1-\tau)T$, de sorte que $B(S)$ est aussi une union d'ensembles R'' tels que $(1-\tau)T \subset R'' \subset (1+\tau)T$. On en déduit aussitôt (4). De plus, la frontière de $B(S)$ est à distance $\leqslant C\epsilon^{n+1}$ de celle de S, et par conséquent $B(S)$ vérifie (5) et (6). On a fini de prouver le Lemme 11.

Remarque. Notre dernière manipulation, destinée à transformer les $A(S)$ en une partition, n'est pas vraiment nécessaire: on vérifierait aisément que la frontière de $A(S)$ est de σ-mesure nulle.

Pour conclure ce paragraphe, résumons dans un lemme les résultats précédents.

Lemme 12. Soient σ *une mesure de Radon positive sur* $\mathbb{R}^N$ (*muni de la distance euclidienne*), Σ *son support, et* $z: \Sigma \to \mathbb{R}^d$ *une fonction lipschitzienne. Si* (1), (2) *et* (9) *sont satisfaites, et si* I_0 *est un* 0*-cube tel que* $|z(I_0)| \geqslant \delta\sigma(I_0)$, *alors on peut trouver* $E \subset I_0$, *tel que* $\sigma(E) \geqslant \theta$ *et tel que la restriction de* z *à* E *soit bi-lipschitzienne.*

4. Intégrales Singulières. Surfaces de S. Semmes

Donnons-nous une fonction K, définie et C^∞ sur $\mathbb{R}^N \setminus \{0\}$, impaire et homogène de degré $-d$. Etant donnée une mesure de Radon positive σ sur $\mathbb{R}^N$, on définit un opérateur maximal par

$$T_\sigma^* f(z) = \sup_{\epsilon > 0} \left| \int_{|z-w|>\epsilon} K(z-w) f(w)\, d\sigma(w) \right|. \tag{43}$$

On cherche à savoir quand cet opérateur est borné sur $L^2(\mathbb{R}^N, d\sigma)$. Nous avons déjà vu un exemple au Paragraphe II: on peut montrer que la continuité de l'opérateur défini par (33) lorsque z est un «paramétrage régulier» entraîne la continuité de T_σ^*, où σ est l'image de la mesure de Lebesgue sur $\mathbb{R}^d$ par z (autrement dit, σ est la mesure telle que

$$\int_{\mathbb{R}^N} f(w)\, d\sigma(w) = \int_{\mathbb{R}^d} f(z(x))\, dx$$

pour tout $f \in C_c(\mathbb{R}^N)$). (Voir [D2] pour plus de détails.) Un autre exemple est dû à S. Semmes [S2]. Nous nous proposons de redémontrer, en utilisant le Théorème 1, le résultat de Semmes.

Définition 2. *Soit* $\Sigma \subset \mathbb{R}^{d+1}$ *un ensemble de dimension* d, *muni d'une mesure de Radon positive* σ *supportée par* Σ. *Nous dirons que* Σ *est dans* $SS(d)$ *si* (1) *et* (2) *sont satisfaites, et si de plus*

(44) *Pour tout point* $x \in \Sigma$ *et tout* $R > 0$, *on peut trouver deux points* x_1 *et* x_2, *contenus dans des composantes connexes différentes de* $\mathbb{R}^{d+1} \setminus \Sigma$, *et tels que* $|x_i - x| \leqslant R$ *et* $\operatorname{dist}(x_i, \Sigma) \geqslant C_0^{-1} R$ *pour* $i = 1, 2$.

Notons au passage que nous n'avons gardé la condition (2) que par commodité: en fait, (2) est une conséquence facile de (44). La définition que nous venons de donner est très légèrement plus générale que celle utilisée dans [S2], où l'auteur devait, pour des raisons techniques, ajouter l'hypothèse qualitative que Σ est une surface orientée. Ici, nous ne supposons rien sur le nombre de composantes connexes du complémentaire de Σ, ou même sur le nombre de composantes de Σ. Modulo cette différence qualitative, Semmes a prouvé que, si $\Sigma \in SS(d)$, l'opérateur T_σ^* est borné sur $L^2(\Sigma, d\sigma)$ pour la plupart des

noyaux K (et en particulier lorsque K est le potentiel de double-couche, ou le noyau de Cauchy-Clifford). La preuve donnée par Semmes utilise le «Théorème $T(b)$», appliqué en prenant pour b le vecteur unitaire normal à Σ. La démonstration que nous allons donner maintenant est basée sur le critère suivant, qui utilise les techniques de variable réelle classiques dans l'étude des opérateurs de Calderón-Zygmund (par exemple, l'usage de fonctions maximales et d'inégalités aux bons λ).

Lemme 13. *Soit σ une mesure de Radon sur $\mathbb{R}^N$ qui vérifie* (1) *et* (2). *Supposons qu'il existe $M \geqslant 0$ et $\gamma > 0$ tels que, pour toute boule $B \subset \mathbb{R}^N$ centrée en un point du support Σ de σ, on puisse trouver une fonction M-lipschitzienne $F: \mathbb{R}^d \to \mathbb{R}^{N-d}$ et un compact $E \subset \Sigma$, tel que $\sigma(E) \geqslant \gamma\sigma(B)$, et qui soit contenu dans l'image du graphe de F par une transformation orthogonale de $\mathbb{R}^N$. Alors, l'opérateur maximal T_σ^* est borné sur $L^p(\mathbb{R}^N, d\sigma)$ pour $1 < p < \infty$.*

Pour prouver le lemme, on utilise la Proposition 4 Bis de [D2], qui dit que T_σ^* est borné sur L^p dès que σ satisfait à (1) et (2), et que de plus, il existe une constante $\gamma > 0$ telle que

(45) pour toute boule B centrée en un point du support de σ, on peut trouver une mesure μ, vérifiant (1) et (2) (avec une constante qui ne dépend pas de B), telle que, pour tout $1 < p < \infty$, l'opérateur maximal T_μ^* soit borné sur $L^p(\mathbb{R}^N, d\mu)$ uniformément en B, et un compact $E \subset \Sigma$ tel que $\sigma(E) \geqslant \gamma\sigma(B)$ et tel que $\chi_E \sigma \leqslant \mu$.

Nous allons appliquer ce résultat en prenant pour μ la mesure de surface d-dimensionelle sur l'image, par l'isométrie linéaire donné par l'hypothèse du lemme, du graphe de F. Le fait que, dans ce cas, l'opérateur T_μ^* est borné sur $L^p(\mathbb{R}^N, d\mu)$ est une conséquence facile du Théorème de Coifman, McIntosh et Meyer sur la continuité de l'opérateur défini par le noyau de Cauchy sur un graphe lipschitzien [CMM] (voir par exemple [D2], Lemme 16). Il ne reste donc plus, pour prouver le lemme, qu'à montrer que si E est le compact mentionné dans l'hypothèse, on a $\chi_E \sigma \leqslant C\mu$ (pour un C qui ne dépend pas de B).

Soit donc $A \subset E$, et soit B_i, $i \in I$ un recouvrement de A par des boules dont les rayons respectifs r_i sont tels que $\sum_i r_i^d \leqslant C\mu(A)$. (L'existence d'un tel recouvrement n'est rien de plus que la définition de la mesure de surface sur le graphe de Σ.) On déduit de (1) que, pour tout i, $\sigma(B_i) \leqslant Cr_i^d$; par conséquent, $\sigma(A) \leqslant C\mu(A)$, et on en déduit le lemme.

Proposition 2. *Soient $\Sigma \in SS(d)$, $x \in \Sigma$, et $r > 0$. Alors on peut trouver une fonction M-lipschitzienne $F: \mathbb{R}^d \to \mathbb{R}$ et un compact $E \subset \Sigma \cap B(x, R)$ tel que $\sigma(E) \geqslant \gamma\sigma(B(x, r))$ et tel que E soit contenu dans l'image, par une transformation orthogonale de $\mathbb{R}^{d+1}$, du graphe de F. Les constantes M et γ ne dépen-*

dent que de la constante C_0 de (1), (2) *et* (44). *Par conséquent, l'opérateur T_σ^* défini par* (43) *est borné sur $L^p(\mathbb{R}^{d+1}, d\sigma)$ pour $1 < p < \infty$.*

La deuxième partie de la proposition découle effectivement de la première et du Lemme 13. Pour trouver E et F, nous allons appliquer le Lemme 12, avec pour fonction z la projection orthogonale de Σ sur un hyperplan V de $\mathbb{R}^{d+1}$. Nous pouvons, par translation de Σ et changement d'échelle, supposer que $x = 0$ et $r = 2C_0$ (où C_0 est comme dans (4)). Choisissons pour V un hyperplan perpendiculaire à la droite qui joint les points x_1 et x_2 de (44), où l'on a pris $x = 0$ et $R = 1$ (voir la figure 1). Si $x_0 = z(x_1)$, où z est la projection orthogonale sur V, alors $z(\Sigma \cap B(0, 2))$ contient la boule $B(z(x_0), C_0^{-1}) \cap V$ car chaque segment de droite qui joint $B(x_1, C_0^{-1})$ à $B(x_2, C_0^{-1})$ rencontre Σ. Comme il n'y a qu'un nombre $\leqslant C$ de 0-cubes qui rencontrent $B(0, 2)$, on peut trouver un 0-cube I_0 tel que $|z(\Sigma \cap I_0 \cap B(0, 2))| \geqslant C^{-1}$. Notons que I_0 est contenu dans $B(x, r) = B(0, 2C_0)$, de sorte que si l'on trouve un compact $E \subset \Sigma \cap I_0$ tel que $\sigma(E) \geqslant \theta > 0$ et qui soit contenu dans l'image par une isométrie d'un graphe M-lipschitzien, nous aurons fini de prouver la Proposition 2.

Lemme 14. *La propriété* (9) *est vérifiée lorsque $\Sigma \in SS(d)$ et z est la projection orthogonale sur un hyperplan de $\mathbb{R}^{d+1}$.*

Avant de prouver ce lemme, vérifions qu'il entraîne la Proposition 2. Puisque (9) est vérifiée et $|z(\Sigma \cap I_0)| \geqslant C^{-1}$, on peut appliquer le Lemme 12 à I_0; on

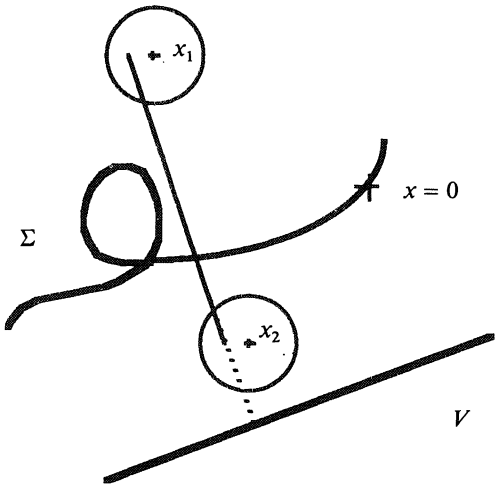

Figure 1.

trouve un compact $E \subset \Sigma \cap I_0$ tel que $\sigma(E) \geqslant \theta$, et tel que la restriction de la projection z à E soit bi-lipschitzienne. Il est clair que cela implique que E est contenu dans l'image par une isométrie d'un graphe lipschitzien. On en déduit la proposition.

Pour prouver le Lemme 14, donnons-nous $x \in \Sigma$, $n \in \mathbb{N}$, et soit T l'union des n-cubes qui rencontrent $B(x, C_2\epsilon^n)$. Supposons pour simplifier que $x = 0$, que $n = 0$, et que z est la projection sur l'hyperplan «horizontal» $\{x_{d+1} = 0\}$. Supposons aussi que $\partial(T) = |z(T)|\sigma(T)^{-1} \geqslant \delta/2$, et que la seconde conclusion de (9) ne soit pas satisfaite, c'est-à-dire que $\partial(R) \leqslant (1 + 2\eta)\partial(T)$ pour tout 0-cube R contenu dans T. Enfin, quitte à augmenter C_1, supposons que $C_1 \geqslant 1$.

Lemme 15. *Si C_2 est grand et η assez petit, on peut trouver un point x_1 sur le segment vertical joignant 0 à $(0, 0, \ldots, C_2/4)$ tel que la boule $B_1 = B(x_1, C_1)$ ne rencontre pas Σ.*

Supposons qu'il n'en soit pas ainsi. Alors, on pourrait trouver au moins $(10C_1)^{-1}C_2$ boules disjointes D_i, de rayon C_1, centrées en des points de $\Sigma \cap B(0, C_2/3)$, et telles que la projection $z(D_i)$ soit contenue dans $B(0, 2C_1)$. Soit A la réunion des 0-cubes qui rencontrent une de ces boules. On a $\sigma(A) \geqslant (CC_1)^{-1}C_2$, alors que $z(A) \subset B(0, 3C_1)$. Si C_2 est assez grand, on en déduit que $|z(A)| \leqslant \delta\sigma(A)/10$, et par conséquent que

$$\begin{aligned} |z(T \setminus A)| &\geqslant |z(T)| - |z(A)| \\ &\geqslant \partial(T)\sigma(T) - \delta\sigma(A)/10 \\ &\geqslant \partial(T)\sigma(T \setminus A) + \partial(T)\sigma(A) - \delta\sigma(A)/10 \\ &\geqslant \partial(T)\sigma(T \setminus A) + 4\delta\sigma(A)/10 \\ &> (1 + 2\eta)\partial(T)\sigma(T \setminus A) \end{aligned}$$

si η est assez petit pour que $4\delta(CC_1)^{-1}C_2 > 20\eta\partial(T)\sigma(T)$. On obtient la contradiction désirée en remarquant que $T \setminus A$ est composé de 0-cubes R tels que $|z(R)| \leqslant (1 + 2\eta)\sigma(R)\partial(T)$.

Donc le Lemme 15 est vrai; on prouverait de la même manière que l'on peut trouver un point x_2 sur le segment vertical joignant 0 à $(0, 0, \ldots, -C_2/4)$ tel que la boule $B_2 = B(x_2, C_1)$ ne rencontre pas Σ.

Supposons que x_1 et x_2 soient dans deux composantes connexes différentes de $\mathbb{R}^{d+1} \setminus \Sigma$. Alors tout segment vertical joignant B_1 à B_2 recontre Σ, et la projection $\Sigma \cap T)$ contient $B(0, C_1)$. Dans ce cas, donc, la conclusion de (9) est vérifiée.

Si x_1 et x_2 sont dans la même composante, utilisons (44). Il existe un point $x_3 \in B(0, 10^{-1}C_1)$ tel que dist $(x_3, \Sigma) \geqslant (10C_0)^{-1}C_1$ et qui n'est pas dans la même composante que x_1 et x_2 (voir la figure 2). Pour tout x dans $B(x_3, (20C_0)^{-1}C_1)$,

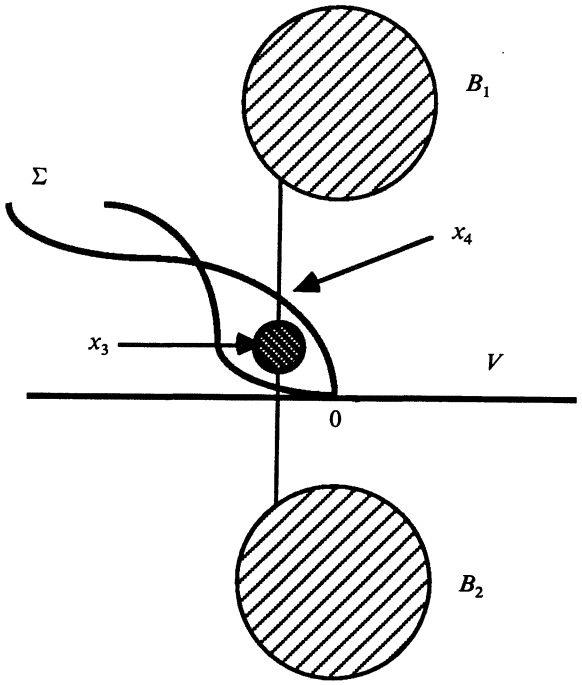

Figure 2.

les segment verticaux qui joignent x à B_1 et à B_2 rencontrent tous les deux Σ. Soit x_4 un point de Σ qui se trouve sur le segment vertical qui joint x_3 à B_1, et soit R le 0-cube contenant x_4. Quitte à agrandir C_1, on peut supposer que C_1 est assez grand par rapport à C_0 pour que R soit contenu dans $B(x_4, 100^{-1}C_1)$. Alors, $z(R)$ est contenu dans $z(T \setminus R)$ (si $y \in R$, le segment vertical joignant y à B_2 rencontre $\Sigma \cap T$ en un point qui est en dessous de la boule $B(x_3, 20^{-1}C_1)$, et qui ne peut donc pas être dans R). Par conséquent,

$$|z(T \setminus R)| = |z(T)| > (1 + 2\eta)\partial(T)\sigma(T \setminus R)$$

si η est assez petit, ce qui contredit le fait que l'on a supposé que tous les 0-cubes composant T sont tels que

$$|z(R)| \leqslant (1 + 2\eta)\sigma(R)\partial(T).$$

Ceci fini de prouver le Lemme 14, et, du même coup, la Proposition 2.

Nous nous proposons pour finir de généraliser la Proposition 2 à des surfaces de dimension d dans $\mathbb{R}^N$ (avec $N \geqslant d + 1$).

Définition 3. *Nous dirons que l'ensemble $\Sigma \subset \mathbb{R}^N$, muni de la mesure de Radon positive σ, est dans $SS(d, N)$ si* (1) *et* (2) *sont vérifiés, et si de plus*

(46) *pour tout $x_0 \in \Sigma$ et tout $r > 0$, on peut trouver un sous-espace affine W de dimension $N - d$ et une sphère S de dimension $N - d - 1$ contenue dans W, telle que $S \subset B(x_0, r)$, $\operatorname{dist}(S, \Sigma) \geqslant C_0^{-1} r$, et telle que S et Σ soient enlacés (au sens défini çi-dessous).*

Nous dirons que S et Σ sont enlacés si l'on ne peut trouver aucune homotopie $F_t(x)$, définie et continue pour $(t, x) \in [0, 1] \times \mathbb{R}^N$, et telle que

—$F_t(x) = x$ pour $t = 0$, ainsi que pour x hors de $B(x_0, 10r)$,
—$F_1(x) \in B(x_0, 10r)^c$ pour tout $x \in \Sigma$, et
—$F_t(x) \in S^c$ pour tout $x \in \Sigma$ et tout $0 \leqslant t \leqslant 1$.

Remarques. La définition de la notion d'enlacement que nous avons choisie n'est pas la plus naturelle: il aurait été plus agréable de dire que S et Σ sont enlacés si l'on ne peut pas déformer S en un point sans rencontrer Σ (au lieu de dire que l'on ne peut pas déformer Σ sans rencontrer S). Il est très probable que le résultat que nous avons en vue soit vrai avec cette définition plus naturelle de $SS(d, N)$, mais l'auteur ne sait pas le démontrer.

Le fait que S est une sphère n'est pas important; on peut se contenter de n'importe quel ensemble S contenu dans $W \cap B(x_0, r)$, à condition que la dimension de S soit $< N - d$ (la démonstration que nous allons donner conviendra encore dans ce cas). Enfin, nous avons encore demandé explicitement que (2) soit satisfaite, mais (2) est en fait une conséquence facile de (46).

Proposition 3. Soit $\Sigma \in SS(d, N)$. *Alors, il existe $\gamma > 0$ et $M \geqslant 0$ tels que, si $x_0 \in \Sigma$ et $r > 0$, on puisse trouver une fonction M-lipschitzienne F: $\mathbb{R}^d \to \mathbb{R}^{N-d}$ et un ensemble compact $E \subset \Sigma \cap B(x_0, r)$ tel que $\sigma(E) \geqslant \gamma\sigma(B(x, r))$ et tel que E soit contenu dans l'image du graphe de F par une transformation orthogonale de $\mathbb{R}^N$. Par conséquent, l'opérateur T_σ^* défini par* (43) *est borné sur $L^p(\mathbb{R}^N, d\sigma)$ pour $1 < p < \infty$.*

Le début de la démonstration de cette Proposition est calqué sur la preuve de la proposition 2. Comme lorsque $N = d + 1$, on veut appliquer le Théorème 1, et on prendra pour z la projection orthogonale sur un plan V de dimension d.

Si $x_0 \in \Sigma$ et $r > 0$, soient W et S le $(N - d)$-plan et la sphère donnés par (46). Soit x_1 un point de S, et soit V le d-plan orthogonal à W passant par x_1 (voir la figure 3). Supposons un instant que $z(B(x_0, 3r))$ ne contienne pas la boule de V centrée en x_1 et de rayon $(2C_0)^{-1} r$, et soit x_2 un point de cette boule tel que $z^{-1}(\{x_2\})$ ne rencontre pas $\Sigma \cap B(x_0, 3r)$. Appelons D le disque de dimension $N - d$, parallèle à W, centré en x_2, et de rayon $2r$. Alors, D ne ren-

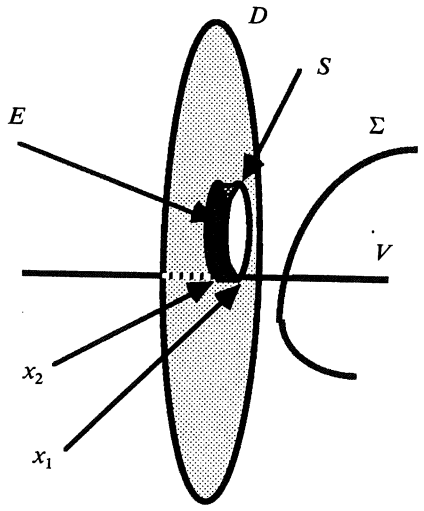

Figure 3. ($d = 1$, $N = 3$).

contre pas Σ, et si E est l'union de D et de tous les segments $[x, h(x)]$, où x est un point de S et $h(x)$ est sa projection orthogonale sur le $(N - d)$-plan passant par x_2 et parallèle à W, alors E ne rencontre pas Σ, parce que S est à distance $\geqslant C_0^{-1} r$ de Σ. Comme chaque $h(x)$ est dans D, on peut aisément trouver une homotopie qui fixe E et qui transforme le complémentaire d'un voisinage de E en $B(x_0, 10r)^c$ sans rencontrer E, ce qui contredit la définition de S. Ainsi, $z(B(x_0, 3r))$ contient une boule de rayon $(2C_0)^{-1} r$, de sorte que $|z(B(x_0, 3r))| \geqslant r/C$.

Le même argument que lorsque $N = d + 1$ nous permet maintenant d'affirmer que la Proposition 3 est une conséquence du lemme suivant.

Lemme 16. *La propriété* (9) *est vérifiée lorsque* $\Sigma \in SS(d, N)$, *et* z *est la projection orthogonale sur un sous-espace* V *de dimension* d.

Pour prouver ce lemme, donnons-nous une constante C_1, un point $x \in \Sigma$, un entier $n \geqslant 0$, et soit T l'ensemble des n-cubes qui rencontrent $B(x, C_2 \epsilon^n)$. Supposons que $|z(T)| \geqslant \delta\sigma(T)/2$ et aussi que, pour tout n-cube $R \subset T$, $\partial(R) \leqslant (1 + 2\eta)\partial(T)$; on veut en déduire que $z(T)$ contient la boule centrée en $z(x)$ et de rayon $C_1 \epsilon^n$ (à condition que C_2 soit assez grand et η assez petit).

On peut supposer que $x = 0$, $n = 0$, et que z est la projection sur le d-plan «horizontal»

$$V = \{x \in \mathbb{R}^N : x_{d+1} = \cdots = x_N = 0\}.$$

Soit

$$W = \{x \in \mathbb{R}^N : x_1 = \cdots = x_d = 0\}$$

le $(N-d)$-plan orthogonal.

Si $z(B(0, C_2))$ contient $V \cap B(0, C_1)$, nous avons prouvé (9). On peut donc supposer qu'il existe un point $x_0 \in V \cap B(0, C_1)$, tel que $z^{-1}(\{x_0\})$ ne rencontre pas $\Sigma \cap B(0, C_2)$. Le reste de la démonstration consistera à trouver une contradiction avec (46); très grossièrement, l'idée est que si $\partial(R) \leqslant (1 + 2\eta)\partial(T)$ pour tout 0-cube R, on ne peut jamais avoir deux 0-cubes R et R' dont les projections coincident, ce qui nous aidera à trouver une contraction de Σ qui ne rencontre pas une sphère S donnée à l'avance. Commençons par quelques notations: Q_0 sera le cube centré en 0 et de côté $(4N)^{-1}C_2$ (de sorte que $2Q_0$ est contenu dans $B(0, C_2)$), V_0 sera la projection $z(Q_0)$, et $W_0 = h(Q_0)$, où l'on note h la projection sur W. Quitte à changer un peu C_2, on peut supposer que les sommets de Q_0 sont dans $C_1\mathbb{Z}^N$. Recouvrons V_0 par les cubes compacts contenus dans V_0, de côté C_1, et dont les sommets sont dans $C_1\mathbb{Z}^N$ (nous noterons ces cubes A_i, $i \in I$). Enfin, nous noterons dorénavant (x, y), avec $x \in V$ et $y \in W$, le point courant de $\mathbb{R}^N$.

Lemme 17. *Il existe une constante C (qui ne dépend que de C_0, δ et C_1) telle que, pour tout $i \in I$, le nombre des cubes $Q \subset W_0$ dont les sommets sont dans $C_1\mathbb{Z}^{N-d}$ et qui sont tels que $A_i \times Q$ rencontre Σ, est $\leqslant C$.*

En effet, si n est ce nombre, on peut trouver $2^{-N}n$ boules disjointes, de rayon $10^{-1}C_1$, centrées en des points de $\Sigma \cap Q_0$, et dont la projection sur V est contenue dans le cube $2A_i$. Si n était trop grand, on pourrait conclure comme dans la démonstration du Lemme 15. Nous noterons $J(i)$ l'ensemble des cubes Q du Lemme 17.

Lemme 18. *Pour tout $i \in I$, on peut trouver un point x_i, dans l'intérieur de A_i, tel que si Q et Q' sont deux cubes de $J(i)$ et si $z(\Sigma \cap A_i \times Q)$ et $z(\Sigma \cap A_i \times Q')$ contiennent x_i, alors Q et Q' sont adjacents.*

En effet, si Q et Q' ne sont pas adjacents, ils sont à distance $\geqslant C_1$ l'un de l'autre; dans ce cas, si C_1 est assez grand, tout 0-cube R qui rencontre $A_i \times Q$ est disjoint de tout 0-cube R' qui rencontre $A_i \times Q'$. Donnons-nous un $\alpha > 0$. Si $|z(R) \cap z(R')| > \alpha$, alors $|z(T)| \leqslant \sum |z(R_j)| - \alpha$ (où la somme porte sur tous les 0-cubes $R_j \subset T$), donc

$$|z(T)| \leqslant (1 + 2\eta)\partial(T)\sigma(T) - \alpha < \partial(T)\sigma(T) = |z(T)|$$

si η est assez petit. Par conséquent, $|z(R) \cap z(R')| \leqslant \alpha$, et comme $A_i \times Q$ et $A_i \times Q'$ ne rencontrent qu'un nombre $\leqslant C$ de 0-cubes, on obtient

$$|z(\Sigma \cap A_i \times Q) \cap z(\Sigma \cap A_i \times Q')| < C\alpha.$$

Compte tenu du Lemme 17, l'ensemble des points $x \in A_i$ tels qu'il existe deux cubes non adjacents Q et Q' dans $J(i)$ pour lesquels

$$x \in z(\Sigma \cap A_i \times Q) \cap z(\Sigma \cap A_i \times Q')$$

a une mesure $\leqslant C\alpha < C_1^d$ (en choisissant α assez petit), ce qui prouve le Lemme 18.

Nous allons faire subir à Σ une première homotopie, qui déplacera les points parallèlement à V, afin d'agrandir les trous dont le Lemme 18 révéle l'existence.

Lemme 19. *On peut trouver une homotopie $F_t(x, y)$, définie et continue sur $[0, 1] \times \mathbb{R}^d \times \mathbb{R}^{N-d}$, et telle que*

(47) *$F_t(x, y) = (x, y)$ dès que $t = 0$ ou que (x, y) est hors de $2Q_0$;*
(48) *$|F_t(x, y) - (x, y)| \leqslant 2NC_1$ pour tout t et tout (x, y);*
(49) *$h(F_t(x, y)) = y$ pour tout t et tout (x, y);*
(50) *Notons Σ_1 l'image de Σ par F_1. Pour tout $i \in I$, on peut trouver un cube $B(i) \subset W_0$, de côté $4C_1$, tel que si $(x, y) \in \Sigma_1 \cap Q_0$, si $x \in A_i$, et si $y \notin B(i)$, alors $x \in \partial A_i$. De plus, il existe un $i_0 \in I$ tel que $A_{i_0} \subset B(0, CC_1)$ et tel que si $(x, y) \in \Sigma_1 \cap Q_0$ avec $x \in A_{i_0}$, alors $x \in \partial A_{i_0}$.*

Pour construire F_t, donnons nous pour chaque I un point x_i comme au Lemme 18, et choisissons un cube B de côté C_1 tel que $\Sigma \cap A_i \times Q$ ne rencontre pas $z^{-1}(\{x_i\})$ dès que $Q \in J(i)$ n'est pas contenu dans $3B$. On peut construire une homotopie $F_{i,t}$, définie sur $[0, 1] \times A_i \times W$, qui est l'identité sur $[0, 1] \times \partial A_i \times W$, qui ne change pas la coordonnée y, et qui transforme le complémentaire dans $A_i \times (W_0 \setminus 4B)$ d'un voisinage cylindrique de $\{x_i\} \times (W_0 \setminus 4B)$ en $\partial A_i \times (W_0 \setminus 4B)$. On choisit le voisinage cylindrique de façon à ce qu'il ne rencontre pas Σ; alors, en mettant bout-à-bout les homotopies $F_{i,t}$, on obtient une fonction F qui vérifie (47), (48), (49), et la première moitié de (50). La seconde moitié de (50) est encore plus facile à obtenir: on choisit i_0 tel que A_{i_0} contienne un point $x_0 \in B(0, C_1)$ tel que $z^{-1}(\{x_0\})$ ne rencontre pas $\Sigma \cap B(0, C_2)$ (on a supposé plus haut qu'un tel point existe), et on déforme $3A_{i_0} \times W$ de manière à transformer le complémentaire d'un voisinage tubulaire de $\{x_0\} \times W$ en $\partial(3A_{i_0}) \times W$. On obtient, en recollant cette homotopie aux $F_{i,t}$ construites plus haut, la fonction F_t cherchée.

Nous allons maintenant déformer la surface Σ_1, en déplaçant les points parallèlement à W, pour obtenir une nouvelle surface Σ_2 telle que, pour tout $x \in V_0$, $\Sigma \cap z^{-1}(\{x\})$ soit composé d'au plus C points.

Lemme 20. *On peut trouver une fonction $G_t(x, y)$, définie et continue sur $[0, 1] \times \mathbb{R}^d \times \mathbb{R}^{N-d}$, et telle que*

(50) *$G_t(x, y) = (x, y)$ quand $t = 0$ ou quand (x, y) est hors de $2Q_0$;*
(51) *$|G_t(x, y) - (x, y)| \leqslant C$ pour tout t et tout (x, y);*
(52) *$z(G_t(x, y)) = x$ pour tout t et tout (x, y);*
(53) *si Σ_2 est l'image de Σ_1 par G_1, alors, pour tout $x \in V_0$, $\Sigma_2 \cap z^{-1}(\{x\})$ a au plus C éléments.*

Pour chaque $i \in I$ et chaque cube $Q \in J(i)$, notons c_Q le centre de Q, et soit $P(Q)$ l'ensemble des points $(x, y) \in \mathbb{R}^N$ tels que $\operatorname{dist}(x, A_i) < 1$, et tels que y est dans le cube de centre c_Q et de côté $(1 - \operatorname{dist}(x, A_i))C_1$. Notons P l'union, pour $i \in I$ et $Q \in J(i)$, des $P(Q)$; pour chaque $x \in 2V_0$, soit $P(x)$ l'ensemble des points $y \in W_0$ tels que $(x, y) \in P$. On définit une semi-distance D sur W par

$$D(y, y') = \inf_\gamma \left\{ \int_{[0,1]} |\gamma'(u)| \chi_{\gamma(u) \notin P(x)} \, du \right\},$$

où la borne inférieure est prise sur l'ensemble des chemins γ de classe C^1 tels que $\gamma(0) = y$ et $\gamma(1) = y'$. Grâce à la forme particulière des $P(Q)$, la distance D est une fonction lipschitzienne de x. On définit maintenant la fonction G_t par $z(G_t(x, y)) = x$ et

$$h(G_t(x, y)) = \lambda_0 y + \Sigma_Q \lambda_Q c_Q, \tag{54}$$

où la somme porte sur les cubes Q tels que $x \in z(P(Q))$, et où λ_0 et les λ_Q sont donnés par les formules suivantes:

$$\lambda_0 = M^{-1}\mu_0 \quad \text{et} \quad \lambda_Q = M^{-1}\mu_Q, \quad \text{où} \tag{55}$$

$$\mu_0 = 1 - t \sup_Q \{ [1 - \operatorname{dist}(x, A_i)]_+ f[D(y, c_Q)] \} \tag{56}$$

(où i est l'indice tel que $Q \in P(i)$),

$$\mu_Q = t[1 - \operatorname{dist}(x, A_i)]_+ f[D(y, c_Q)], \tag{57}$$

où f est une fonction C^∞ décroissante, telle que $f(0) = 1$ et $f(u) = 0$ pour tout $u \geqslant 1$, et où

$$M = \mu_0 + \Sigma_Q \mu_Q. \tag{58}$$

Vérifions que G_t est bien la fonction annoncée dans le lemme. Les coefficients μ_0 et μ_Q ne sont jamais simultanément nuls, et donc G_t est une fonction

continue (en particulier de x). D'autre part, si $t = 0$ ou si (x, y) n'est pas dans $2Q_0$, tous les μ_Q sont nuls et $\mu_0 = 1$, donc $G_t(x, y) = (x, y)$. L'inégalité (51) découle de ce que tous les cubes Q tels que $\mu_Q \neq 0$ sont à distance (euclidienne) $< C$ de y, de sorte que $h(G_t(x, y))$ est une moyenne de points situés à distance $\leqslant C$ de y.

Il ne reste plus qu'à vérifier (53). Fixons $i \in I$ et $x \in A_i$. A cause de notre définition de D, le coefficient μ_0 est nul lorsque $t = 1$ et y est dans l'un des cubes $Q' \in J(i)$; de plus, les μ_Q sont constants sur chaque composante connexe de $P(x)$, de sorte que $h(G_t(x, y))$ ne prend qu'un nombre $\leqslant C$ de valeurs lorsque y décrit l'union des Q' (qui contient $\Sigma \cap z^{-1}(\{x\})$ par définition de $J(i)$, et $\Sigma_1 \cap z^{-1}(\{x\})$ car on a choisi l'homotopie F_t de façon à ce que $z[F_t(x, y)]$ soit dans A_i quand $x \in A_i$). On en déduit le Lemme 20.

Appelons $L = \Sigma_2 \cap Q_0 \cap z^{-1}\left(\bigcup_i \partial A_i\right)$. Compte tenu de (53), L est un ensemble de dimension $\leqslant d - 1$ (et il nous sera facile de modifier une homotopie de manière à lui faire éviter un ensemble donné S de dimension $N - d - 1$). Par ailleurs, grâce à (51) et au Lemme 19, $L' = (\Sigma_2 \cap Q_0) \setminus L$ est contenu dans l'union des $A_i \times CB(i)$, $i \neq i_0$, qui rencontrent Σ_2 (où i_0 et $B(i)$ sont comme au Lemme 19).

Pour obtenir notre contradiction, donnons-nous une sphère S de dimension $N - d - 1$, contenue dans $Q_0/2$, dont la distance à Σ est $\geqslant C^{-1}C_2$, et qui est entrelacée avec Σ au sens de la Définition 3. Notons que, si C_2 est assez grand, l'image de Σ par les homotopies successives que nous avons construites ne rencontre pas S.

Lemme 21. *On peut trouver une homotopie H_t, telle que $H_t(x, y) = (x, y)$ pour $t = 0$ et pour (x, y) hors de $2Q_0$, telle que $H_1(L \cup L')$ soit contenu dans ∂Q_0, telle que $H_t(L \cup L') \cap S = \varnothing$ pour tout t.*

Rangeons les cubes A_i en deux catégories: nous dirons que A_i est plein si $z(\Sigma_2 \cap [A_i \times CB(i)])$ contient au moins un point de l'intérieur de A_i; sinon, nous dirons que A_i est vide. Notons que, grâce à la seconde partie de (50), on est sûr qu'au moins un des cubes intérieurs de V_0 est vide. Soit A_i un cube plein dont l'un des voisins (appelons-le A_j) est vide. Alors, on peut trouver une homotopie de $\mathbb{R}^N$ qui contracte $[A_i \cup A_j \setminus V] \times CB(i)$ (où V est un petit voisinage tubulaire du point x_j) en une partie de $\partial[A_i \cup A_j] \times CB(i)$, sans rien changer hors de $(A_i \cup A_j) \times 2CB(i)$. Nous venons de trouver une homotopie à l'issu de la laquelle le cube A_j est devenu vide. On peut répéter cette opération aussi longtemps que nécessaire, et obtenir une homotopie qui transforme L' en un ensemble L'' dont la projection est contenue dans l'union des ∂A_i. On peut encore appliquer la technique du Lemme 20 pour obtenir, après nouvelle homotopie (que nous appellerons $M_t(x, y)$), un ensemble $L^{(3)}$ qui est de

dimension $\leqslant d-1$. Notons que chacune des homotopies que nous venons de composer n'est différente de l'identité que dans des boules de rayon C contenant un point de Σ_2, et, même dans ces boules, diffère de l'identité par moins de C. Si C_2 est assez grand, l'image de $L \cup L'$ par M_t ne rencontre pas S.

Appliquons une dernière homotopie N_t, que l'on choisit de classe C^1, et qui contracte le complémentaire dans Q_0 d'une petite boule de l'intérieur de Q_0 ne rencontrant pas $M_1(L \cup L')$, en une partie de ∂Q_0. Si l'on mettait M_t et N_t bout-à-bout, on obtiendrait une homotopie qui convient presque pour le Lemme 21, sauf que la trajectoire de $M_1(L \cup L')$ par N_t pourrait rencontrer S. Cette difficulté est aisément circonvenue, car comme S est de dimension $< N-d$ et la trajectoire par N_t de $M_1(L \cup L')$ est de dimension $\leqslant d$, on peut trouver une petite perturbation de N_t telle que $N_t[M_1(L \cup L')]$ ne rencontre jamais S. Si nous composons les homotopies des Lemmes 19, 20 et 21, nous obtenons une homotopie qui contredit la définition de S (en fait, il faudrait encore modifier un peu notre construction pour obtenir une homotopie qui soit l'identité hors de $B(0, 5|Q_0|^{1/N})$, au lieu de $2Q_0$; nous omettrons ce détail facile). Nous avons enfin prouvé le Lemme 16 et la Proposition 3.

Signalons pour finir que la figure 4 représente un exemple de surface qui a une grande projection sur le plan horizontal, mais qui ne contient pas de gros morceaux de graphe lipschitzien. On pourrait aussi prendre le produit d'un ensemble de Cantor de dimension 1 (l'exemple de Garnett-Ivanov) par un segment. L'exemple Σ de la figure ne satisfait pas, bien entendu, la condition 9 (la projection de Σ sur le plan horizontal préserve la mesure à une constante près, mais est loin d'être un homéomorphisme). Par ailleurs, la démonstration de P. Jones du Théorème de Garnett s'applique aussi à Σ: le noyau de Cauchy-Clifford ne définit pas un opérateur borné sur $L^2(\Sigma)$ (voir [J]).

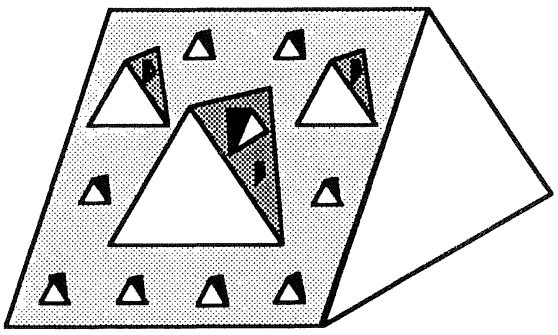

Figure 4. Un contre-exemple.

Références

[CMM] Coifman, R. R., McIntosh, A. et Meyer, Y. L'intégrale de Cauchy définit un opérateur borné sur L^2 pour les courbes lipschitziennes, *Ann. of Math.* **116**(1982), 361-387.

[D1] David, G. Opérateurs intégraux singuliers sur certaines courbes du plan complexe, *Ann. Sci. Ec. Norm. Sup.* **17**(1984), 157-189.

[D2] David, G. Opérateurs d'intégrale singulière sur les surfaces régulières, *Ann. Sci. Ec. Norm. Sup.*, **21** (1988), 225-258.

[G] De Guzmán, M. Differentiation of integrals in $\mathbb{R}^n$. *Lect. Notes in Math.* 481, Springer Verlag, 1975.

[J] Jones, P. Square functions, Cauchy integrals, analytic capacity, and harmonic measure. A paraître, Actes du Séminaire d'Analyse Harmonique, El Escorial, Juin 1986.

[MRV] Martio, O., Rickman, s. et Väisälä, J. Definitions for quasiregular mappings, *Ann. Acad. Sci. Fenn., Ser. Al,* **448**(1969), 1-40.

[R] Resetnjak, J. G. A condition for boundedness of index for mappings with bounded distortion (en russe), *Sibirsk, Mat. Z.* **9**(1968), 368-374.

[S1] Semmes, S. Chord-arc surfaces with small contants I et II, preprint.

[S2] Semmes, S. A criterion for the boundedness of singular integrals on hypersurfaces, à paraître, *Trans. Amer. Math. Soc.*

[St] Stein, E. M. Singular integrals and differentiability properties of functions, Princeton University Press, 1970.

Guy David
Centre de Mathématiques
Ecole Polytechnique
91128 Palaiseau Cedex, France

REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.° 1, 1988

Lipschitz and bi-Lipschitz Functions

Peter W. Jones

Let $Q_0 = [0, 1]^n$ be the unit cube in $\mathbb{R}^n$ and let $f: Q_0 \to \mathbb{R}^m$, $m \geqslant n$, have Lipschitz norm bounded by one,

$$\|f\|_{\text{Lip}} = \sup_{\substack{x,y \in Q_0 \\ x \neq y}} \frac{|f(x) - f(y)|}{|x - y|} \leqslant 1. \tag{1}$$

Then classical results (see *e.g.* Federer [2] or Stein [5]) assert that

$$Df = \left(\frac{\partial f_i}{\partial x_j} \right)_{i,j}$$

is defined almost everywhere on Q_0, and f may be recovered from Df via integration along line segments parallel to the axes. We also recall two classical qualitative results. Let $\mathcal{H}(\cdot)$ denote n dimensional Hausdorff measure and let h denote n dimensional Hausdorff content,

$$h(E) = \inf \sum_{j=1}^{\infty} c_n r_j^n$$

where the infimum is taken over all coverings of E by balls $B(x_j, r_j)$ with no restrictions on the radii r_j. Then $h(E) \leqslant \mathcal{H}(E)$ and if $E \subset \mathbb{R}^n$, $h(E) = \mathcal{H}(E)$. Sard's theorem asserts that $\mathcal{H}(\{f(x): \text{Rank}\,(Df(x)) < n\}) = 0$. A slightly stronger result is that one can decompose

$$Q_0 = G \cup \bigcup_{j=1}^{\infty} K_j$$

where $\mathcal{H}(f(G)) = 0$ and f is bi-Lipschitz on each K_j, *i.e.* there are constants c_j such that

$$|f(x) - f(y)| \geqslant c_j|x - y|, \qquad x, y \in K_j.$$

A qualitative version of the last result was first given by Guy David in [1].

Theorem (Guy David [1]). *Suppose* $f: Q_0 \to \mathbb{R}^n$ *satisfies* $\|f\|_{\text{Lip}} = 1$ *and* $\mathcal{H}(f(Q_0)) \geqslant \epsilon > 0$. *Then there is* $\delta = \delta(\epsilon) > 0$ *and* $K \subset Q_0$ *such that* $\mathcal{H}(K)$, $\mathcal{H}(f(K)) \geqslant \delta$ *and*

$$|f(x) - f(y)| \geqslant \delta|x - y|, \qquad x, y \in K.$$

David's result was used to prove boundedness properties for singular integrals on certain surfaces $\mathcal{S} \subset \mathbb{R}^m$. If $\mathcal{S} = f(\mathbb{R}^n)$ where f is Lipschitz and satisfies some other criterion, the above theorem can be used to show that for all $x_0 \in \mathcal{S}$ and all $r > 0$, $\mathcal{S} \cap \{x \in \mathbb{R}^m: |x - x_0| \leqslant r\}$ contains a subset $K = K(x_0, r)$ such that $\mathcal{H}(K) \geqslant cr^n$ and such that singular integrals are known to be bounded operators on $L^2(K)$. Real variales methods are then used to show that singular integrals are bounded on $L^2(\mathcal{S})$. In this note we present a generalization and strengthening of David's theorem. Our proof is also shorter than David's.

Theorem. *Suppose* $f: Q_0 \to \mathbb{R}^m$ *satisfies* $\|f\|_{\text{Lip}} = 1$. *Then for each* $\delta > 0$ *there is* $M(\delta) < \infty$ *and there are closed sets* $K_1, \ldots, K_M \subset Q_0$, $M \leqslant M(\delta)$, *such that*

$$h\left(f\left(Q_0 \setminus \bigcup_{j=1}^{M} K_j\right)\right) < \delta$$

and such that

$$|f(x) - f(y)| \geqslant \frac{\delta}{2}|x - y|, \qquad x, y \in K_j, \qquad 1 \leqslant j \leqslant M.$$

By using truncation methods, the theorem can be seen to have L^p analogues.

Corollary. *Suppose* $f = (f_1, \ldots, f_m): Q_0 \to \mathbb{R}^m$ *is such that each* f_j *is in the Sobolev space* $W^{1, n+\epsilon}$ *(one derivative in* $L^{n+\epsilon}$*) with* $\|f_j\|_{W^{1, n+\epsilon}} \leqslant 1$, $1 \leqslant j \leqslant m$. *Then the conclusions of the theorem hold with* $M = M(\epsilon, \delta)$.

PROOF. Fix $N < \infty$ and build F such that $\|F\|_{\text{Lip}} \leqslant N$ and $\mathcal{H}(\{x: F(x) \neq f(x)\})$ $\leqslant cN^{-(n+\epsilon)}$. Then use the theorem plus the fact that for any $G \subset Q_0$, $\mathcal{H}(f(G))$ $\leqslant C\mathcal{H}(G)^{(\epsilon/n)/(1+\epsilon/n)}$. □

The proof of the theorem is given in section 2. The main tool is a Littlewood-Paley inequality. The theorem can be used to obtain a different approach to Guy David's results on singular integrals; this will appear elsewhere.

It is with great sorrow that I dedicate this paper to the memory of my good friend José-Luis Rubio de Francia.

2. Proof of the Theorem

Let $F(x)$ be a real valued function on $\mathbb{R}^n$ and let $F(x, y)$ denote its Poisson (harmonic) extension to $\mathbb{R}^{n+1}_+ = \{(x, y): x \in \mathbb{R}^n, y > 0\}$. Also let

$$\nabla F(x, y) = (F_{x_1}(x, y), \ldots, F_{x_n}(x, y), F_y(x, y))$$

denote the gradient of F. Then if $Q \subset \mathbb{R}^n$ is any cube with sidelength $\ell(Q)$, and if $R(Q) = Q \times (0, \ell(Q)]$, ∇F satisfies the well-known BMO type estimate

$$\iint_{R(Q)} |\nabla F|^2 y \, dx \, dy \leqslant C\mathcal{H}(Q)\|F\|^2_{L^\infty(\mathbb{R}^n)}. \tag{2.1}$$

See Fefferman-Stein [3] or Garnett's book [4] page 240 for the proof.

We denote by $\mathfrak{D}$ the collection of all dyadic cubes in $\mathbb{R}^n$, *i.e.* the collection of all cubes Q of form $\prod_{j=1}^n [a_j 2^{-k}, (a_j + 1)2^{-k}]$ where a_j and k lie in $\mathbb{Z}$. For such a cube Q we denote by $\ell(Q) = 2^{-k}$ the sidelength of Q. From now on, all cubes will be dyadic. We also let

$$T(Q) = Q \times \left[\frac{1}{2}\ell(Q), \ell(Q)\right]$$

denote the top half of $R(Q)$. If $Q, Q' \in \mathfrak{D}$, we say that Q and Q' are semi-adjacent if $\ell(Q) = \ell(Q')$, $Q \cap Q' = \varnothing$, and there is $Q'' \in \mathfrak{D}$ with $\ell(Q'') = \ell(Q)$, such that $Q \cap Q'' \neq \varnothing$, $Q' \cap Q'' \neq \varnothing$. Then

(2.2) For each $Q \in \mathfrak{D}$ there are exactly $5^n - 3^n$ semi-adjacent cubes Q'.

Let f satisfy the hypotheses of the theorem; by Whitney's extension theorem (see [2] or [5]) we may assume f is defined on all of $\mathbb{R}^n$ and $\|f\|_{\text{Lip}} \leqslant 1$ there. Write $f = (f_1, \ldots, f_m)$ and $Df = \left(\frac{\partial f_j}{\partial x_k}\right)$. Let $F_{j,k}$ be the harmonic extension of $\frac{\partial f_j}{\partial x_k}$ to $\mathbb{R}^{n+1}_+$ and let

$$|\nabla Df| = \left(\sum_{j,k} |\nabla F_{j,k}|^2\right)^{1/2}.$$

Our next lemma says that if Q and Q' are semi-adjacent, $h(f(Q))$ is large but $f(Q)$ and $f(Q')$ are not well separated, then $|\nabla Df|$ must be large somewhere in $T(Q)$.

Lemma 2.1. *Suppose Q and Q' are semi-adjacent and $h(f(Q)) \geqslant \delta \mathcal{H}(Q)$. If there are $x \in Q$, $x' \in Q'$ such that*

$$|f(x) - f(x')| \leqslant \frac{\delta}{2}|x - x'|,$$

then

$$\iint_{T(Q)} |\nabla Df|^2 y \, dx \, dy \geqslant c(\delta)\mathcal{H}(Q),$$

where $c(\delta) > 0$ is a constant depending only on δ.

PROOF. Since the hypotheses and conclusions are dilation invariant, it is sufficient to treat the case where $Q = Q_0$ is the unit cube. Suppose that the lemma is false, so that there is a sequence of functions f_j satisfying the hypotheses of the lemma but such that

$$\iint_{T(Q_0)} |\nabla Df|^2 y \, dx \, dy \leqslant 2^{-j}.$$

By Arzelà-Ascoli we may assume the f_j converge uniformly to f on compacta. Then $\|f\|_{\text{Lip}} \leqslant 1$, and since $h(f(Q_0)) \geqslant \liminf h(f_j(Q_0))$, f satisfies the hypotheses of the lemma. On the other hand,

$$\iint_{T(Q_0)} |\nabla Df|^2 y \, dx \, dy = 0,$$

so by the uniqueness principle for harmonic functions Df is constant a.e. on $\mathbb{R}^n$, and consequently f is linear on $\mathbb{R}^n$.

However, there is $x' \in Q'$ such that

$$|f(x) - f(x')| \leqslant \frac{\delta}{2}|x - x'|$$

for some $x \in Q_0$. This is not possible for a linear map satisfying $h(f(Q_0)) \geqslant \delta$ and $\|f\|_{\text{Lip}} \leqslant 1$. □

Let

$$G_1 = \left\{ Q \in \mathfrak{D} : Q \subset Q_0, h(f(Q)) \leqslant \frac{\delta}{2}\mathcal{H}(Q) \right\}$$

and let

$$G_1 = \bigcup_{Q \in G_1} Q$$

so that

$$h(f(G_1)) \leqslant \frac{\delta}{2}. \tag{2.3}$$

The set G_1 is a «garbage» set to be thrown out. Let $\mathfrak{F}$ denote the collection of all unordered pairs of semi-adjacent cubes Q, Q' such that both $Q, Q' \notin G_1$, and such that Q, Q' satisfy the hypotheses of Lemma 2.1. We enumerate the collection $\mathfrak{F}$ by $\{(Q_1^1, Q_2^1), (Q_1^2, Q_2^2), (Q_1^3, Q_2^3), \dots\}$ where $\ell(Q_1^j) \geqslant \ell(Q_1^{j+1})$. Let $c(\delta)$ be the constant of Lemma 2.1, and let

$$G_2 = \left\{ x: \sum_{j=1}^{\infty} \sum_{k=1}^{2} \chi_{Q_k^j}(x) \geqslant C_1 \delta^{-1} c(\delta)^{-1} \right\},$$

where C_1 is a constant to be fixed later.

Lemma 2.2. $\mathcal{H}(G_2) \leqslant \delta/2$.

PROOF. By Chebychev's inequality,

$$\begin{aligned}
\mathcal{H}(G_2) &\leqslant C_1^{-1} \delta c(\delta) \sum_{j=1}^{\infty} \sum_{k=1}^{2} \mathcal{H}(Q_k^j) \\
&\leqslant C_1^{-1} \delta \sum_{j=1}^{\infty} \sum_{k=1}^{2} \iint_{T(Q_k^j)} |\nabla Df|^2 y \, dx \, dy \\
&= C_1^{-1} \delta \iint |\nabla Df|^2 \sum_{j,k} \chi_{T(Q_k^j)}(x, y) y \, dx \, dy \\
&\leqslant C_1^{-1} \delta C(n) \iint_{R(Q_0)} |\nabla Df|^2 y \, dx \, dy \\
&\leqslant C_1^{-1} \delta C(n) C.
\end{aligned}$$

The penultimate inequality follows from remark (2.2), while the final inequality results from Lemma 2.1. The proof is concluded by choosing $C_1 \geqslant 2C(n)C$. □

Setting $G = G_1 \cup G_2$, it follows from (2.3) and Lemma 2.2 that $h(f(G)) \leqslant \delta$. We now divide $Q_0 \setminus G$ into M disjoint compact a K_α so that f is bi-Lipschitz on each K_α. To do this we define inductively indices α_j. To each x will correspond $\alpha_j(x)$ and we will define

$$K_\alpha = \left\{ x \in Q_0 \setminus G: \lim_{j \to \infty} \alpha_j(x) = \alpha \right\}.$$

Each $\alpha_j(x)$ will be a finite string of zeros and ones, $\alpha_j(x) = \{\epsilon_1^j(x), \ldots, \epsilon_k^j(x)\}$, where $\epsilon_l^j = 0$ or 1.

At stage zero we define $\alpha_0(x) = \{0\}$ for all $x \in Q_0$. At stage one, let Q_1^1 and Q_2^1 be the first pair of cubes in $\mathcal{F}$. Define

$$\begin{aligned} \alpha_1(x) &= \{0,0\} \quad \text{on} \quad Q_1^1 \\ \alpha_1(x) &= \{0,1\} \quad \text{on} \quad Q_2^1 \\ \alpha_1(x) &= \{0\} \quad \text{on} \quad Q_0 \setminus (Q_1^1 \cup Q_2^1). \end{aligned}$$

We suppose by induction that $\alpha_0, \alpha_1, \ldots, \alpha_{k-1}$ have been defined and that α_{k-1} is constant on each $Q \in \mathcal{D}$ with $\ell(Q) \leqslant \ell(Q_1^{k-1}) = \ell(Q_2^{k-1})$. Let Q_1^k and Q_2^k be the k^{th} pair of cubes in $\mathcal{F}$, and suppose $\alpha_{k-1}(x) = \{\epsilon_1, \ldots, \epsilon_s\}$ on Q_1^k, $\alpha_{k-1}(x) = \{\epsilon_1', \ldots, \epsilon_t'\}$ on Q_2^k.

Case 1. $s = t$. Define

$$\begin{aligned} \alpha_k(x) &= \{\epsilon_1, \ldots, \epsilon_s, 0\} \quad \text{on} \quad Q_1^k \\ \alpha_k(x) &= \{\epsilon_1', \ldots, \epsilon_t', 1\} \quad \text{on} \quad Q_2^k \\ \alpha_k(x) &= \alpha_{k-1}(x) \quad \text{on} \quad Q_0 \setminus (Q_1^k \cup Q_2^k). \end{aligned}$$

Case 2. $s > t$. Define $\epsilon_{t+1}' = 2 - 2^{1+\epsilon_t}$ and

$$\begin{aligned} \alpha_k(x) &= \alpha_{k-1}(x) \quad \text{on} \quad Q_0 \setminus Q_2^k \\ \alpha_k(x) &= \{\epsilon_1', \ldots, \epsilon_t', \epsilon_{t+1}'\} \quad \text{on} \quad Q_2^k. \end{aligned}$$

Case 3. $s < t$. Reverse the roles of Q_1^k and Q_2^k and apply Case 2.

The procedure guarantees that $\alpha_l(x)$ distinguishes between points in Q_1^k and Q_2^k whenever $l \geqslant k$. More precisely, if $l \geqslant k$, $Q_1 \subset Q_1^k$, $Q_2 \subset Q_2^k$, $\alpha_l(x) = \{\epsilon_1, \ldots, \epsilon_u\}$ on Q_1, $\alpha_l(x) = \{\epsilon_1', \ldots, \epsilon_v'\}$ on Q_2, then $u, v \geqslant k+1$ and there is $j \leqslant k+1$ such that

$$\epsilon_j \neq \epsilon_j'. \tag{2.4}$$

Let C_1 be the constant in the definition of G_2, and for $\alpha = \{\epsilon_1, \ldots, \epsilon_s\}$ define $\rho(\alpha) = s$ to be the «length» of α. Then by the definition of G_2 and the α_j's,

$$\rho(\alpha_j(x)) \leqslant 1 + C_1 \delta^{-1} c(\delta)^{-1}$$

for all $x \in Q_0 \setminus G_2$. If $x \in Q_0 \setminus G$ we see therefore that $\alpha(x) = \lim_{j \to \infty} \alpha_j(x)$ is well defined and

$$\rho(\alpha(x)) \leqslant 1 + C_1 \delta^{-1} c(\delta)^{-1}.$$

For $s \leqslant 1 + C_1 \delta^{-1} c(\delta)^{-1}$ and $\alpha = \{\epsilon_1, \ldots, \epsilon_s\}$ a string of zeros and ones, define

$$K_\alpha = \{x \in Q_0 \setminus G : \alpha(x) = \alpha\}.$$

Then there are at most $M(\delta)$ sets K_α, so we need only check that f is bi-Lipschitz on each K_α. To this end, suppose $x, y \in K_\alpha$, but

$$|f(x) - f(y)| < \frac{\delta}{2} |x - y|.$$

Then $x \in Q$, $y \in Q'$ where Q and Q' are semi-adjacent. Since $x, y \notin G_1$, $h(f(Q)), h(f(Q')) \geqslant \ell(Q)$. Therefore the pair (Q, Q') must show up in $\mathfrak{F}$ as a pair (Q_1^k, Q_2^k). By (2.4), $\alpha_l(x) \neq \alpha_l(y)$ whenever $l \geqslant k$, so $\alpha(x) \neq \alpha(y)$. This contradiction completes the proof of the theorem. □

References

[1] David, G. Morceaux de graphes Lipschitziens et integrales singulières sur une surface. *Rev. Mat. Iberoamericana*, **4**(1988), pp. 73-114.
[2] Federer, H. *Geometric measure theory*, Springer-Verlag, 1969.
[3] Fefferman, C. and Stein, E. M. H^p spaces of several variables, *Acta Math.* **129**(1972), 137-193.
[4] Garnett, J.*B*. *Bounded analytic functions*, Academic Press, 1981.
[5] Stein, E. M. *Singular integrals and differentiability properties of functions*, Princeton University Press, 1970.

Peter W. Jones
Department of Mathematics
Yale University
New Haven, CT. 06520, U.S.A.

N.S.F. Grant DMS-86-02500.

REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.° 1, 1988

Maximal and Area Integral Characterizations of Hardy-Sobolev Spaces in the Unit Ball of $\mathbb{C}^n$

Patrick Ahern and Joaquim Bruna

Dedicated to the memory of Professor José Luis Rubio de Francia

Abstract

In this paper we deal with several characterizations of the Hardy-Sobolev spaces in the unit ball of $\mathbb{C}^n$, that is, spaces of holomorphic functions in the ball whose derivatives up to a certain order belong to the classical Hardy spaces. Some of our characterizations are in terms of maximal functions, area functions or Littlewood-Paley functions involving only complex-tangential derivatives. A special case of our results is a characterization of H^p itself involving only complex-tangential derivatives.

1. Introduction. Statement of Results

Let B^n denote the unit ball of $\mathbb{C}^n$. For a holomorphic function f on B^n with homogeneous expansion $f(z) = \sum_k f_k(z)$, the *radial fractional derivative* of order $s > 0$ is defined

$$R^s f(z) = \sum_k (k+1)^s f_k(z)$$

(so for $s = 1$, $Rf(z) = f(z) + Nf(z)$, where $N = \sum z_j(\partial/\partial z_j)$ is the normal field). The *Hardy-Sobolev* space H^p_s is defined

$$H^p_s = \left\{ f: \sup_r \int_S |R^s f(r\zeta)|^p \, d\sigma(\zeta) = \|f\|^p_{p,s} < +\infty \right\}.$$

Here S denotes the boundary of B^n and $d\sigma$ its normalized Lebesgue measure.

It is a well-known general principle that holomorphic functions in B^n behave twice as well in the complex-tangential directions. Our first goal is to make precise how this principle applies to the Hardy-Sobolev spaces H^p_s. To state our results we need to introduce some definitions. A vector field

$$T = \sum_{j=1}^{n} a_j(z) \frac{\partial}{\partial z_j}, \qquad a_j \in C^\infty(\bar{B}^n)$$

is called *complex-tangential* if

$$\sum_{j=1}^{n} a_j(z)\bar{z}_j = 0.$$

If T is a complex-tangential vector field and f is holomorphic then Tf is generally no longer holomorphic. For this reason it is more natural to deal with real-variable characterizations of H^p itself. These are in terms of the following quantities, defined for a smooth function f:

(*a*) The *radial maximal function*

$$f^+(\eta) = \sup \{ |f(r\eta)| : 0 \leqslant r < 1 \}, \qquad \eta \in S.$$

(*b*) The *admissible maximal function*

$$f^*_\alpha(\eta) = \sup \{ |f(z)| : z \in D_\alpha(\eta) \}.$$

Here $D_\alpha(\eta)$ denotes the *admissible approach region*

$$D_\alpha(\eta) = \left\{ z \in B^n : |1 - \bar{z}\zeta| < \frac{\alpha}{2}(1 - |z|^2) \right\},$$

with $\alpha > 1$.

(*c*) The *admissible area function*

$$A_\alpha f(\eta) = \left\{ \int_{D_\alpha(\eta)} |Rf(z)|^2 (1 - |z|)^{1-n} \, dm(z) \right\}^{1/2}.$$

(*d*) The *Littlewood-Paley g-function*

$$g(f)(\eta) = \left\{ \int_0^1 |Rf(r\eta)|^2 (1 - r) \, dr \right\}^{1/2}.$$

It will be convenient to introduce the notations, for $t \in \mathbb{R}^+$ and $\gamma > 0$

$$A^t_{\alpha,\gamma} f(\eta) = \left\{ \int_{D_\alpha(\eta)} |R^t f(z)|^2 (1 - |z|)^{2\gamma - n - 1} \, dm(z) \right\}^{1/2}$$

$$g^t_\gamma f(\eta) = \left\{ \int_0^1 |R^t f(r\eta)|^2 (1 - r)^{2\gamma - 1} \, dr \right\}^{1/2}$$

(so that $A_\alpha = A^1_{\alpha,1}$ and $g = g^1_1$).

If $\varphi(f)$ denotes any of the functions in (a)-(d), then $f \in H^p$ if and only if $\varphi(f) \in L^p(S)$, (see [8]). It then follows by standard techniques (see [7, pp. 214-216] and Lemma 3.6 below) that the area function

$$A^t_{\alpha,t-s} f(\eta) = \left\{ \int_{D_\alpha(\eta)} |R^t f(z)|^2 (1 - |z|)^{2t - 2s - n - 1} \, dm(z) \right\}^{1/2}$$

with any $t > s$, characterizes H^p_s, *i.e.*, $f \in H^p_s$ if and only if $A^t_{\alpha,t-s} f \in L^p(S)$. In a similar way, the g-function

$$g^t_{t-s} f(\eta) = \left\{ \int_0^1 |R^t f(r\eta)|^2 (1 - r)^{2t - 2s - 1} \, dr \right\}^{1/2}$$

characterizes H^p_s if $t > s$.

Our first result states that if $t > s$, if k is a positive integer with $k \leqslant 2t$ and $T_1, \ldots, T_k$ are complex-tangential vector fields, then $f \in H^p_s$ implies

$$(1) \qquad \left\{ \int_{D_\alpha(\eta)} |T_1 \cdots T_k R^{t - k/2} f(z)|^2 (1 - |z|)^{2t - 2s - n - 1} \, dm(z) \right\}^{1/2} \in L^p(S)$$

$$(2) \qquad \left\{ \int_0^1 |T_1 \cdots T_k R^{t - k/2} f(r\eta)|^2 (1 - r)^{2t - 2s - 1} \, dr \right\}^{1/2} \in L^p(S).$$

Also, if $f \in H^p_s$ and $k \leqslant 2s$, the radial and admissible maximal functions of $T_1 \cdots T_k R^{s - k/2} f$ are in $L^p(S)$. That is, one can formally replace each $R^{1/2}$ by one complex-tangential vector field and still get a function in $L^p(S)$.

In case of the area function, this result again follows by the standard techniques in [7, pp. 214-216] and [6, Th. 12], but for the g-function and the two maximal functions it depends on some special properties of complex-tangential derivatives. We comment briefly on these. Each complex-tangential vector field is a linear combination of the T_{ij}

$$T_{ij} = \bar{z}_i \frac{\partial}{\partial z_j} - \bar{z}_j \frac{\partial}{\partial z_i}, \qquad i, j = 1, \ldots, n, \qquad i \neq j.$$

If f is holomorphic, then $T_{ij} f$ is no longer holomorphic, but on every complex line is annihilated by $\bar{\partial}^2$. In the same way, if the $T_1, \ldots, T_k$ are chosen among the T_{ij} (which we may assume without loss of generality), then $T_1 \cdots T_k f$ is annihilated by $\bar{\partial}^{k+1}$ on every complex line. We show that such functions share many properties of holomorphic functions, such as mean-value inequalities,

Cauchy inequalities for derivatives etc., which are on the basis for the transition from the area function to the others.

Our second result states that the above property gives in fact a characterization of H^p_s. Namely, with k fixed, if (1) or (2) holds for all possible choices of $T_1 \cdots T_k$ among the T_{ij} then $f \in H^p_s$. In the same way, if $(T_1 \cdots T_k R^{s-k/2} f)^+ \in L^p$ for all $T_1 \cdots T_k$ (k fixed, $k \leqslant 2s$), then $f \in H^p_s$. Choosing k such that $k > 2s$ and $t = k/2$ we thus obtain characterizations of H^p_s in terms of area functions or g-functions involving only complex-tangential derivatives. For instance, as a particular case we obtain that $f \in H^p$ if and only if for all T_{ij}

$$\left\{\int_0^1 |T_{ij} f(r\eta)|^2 \, dr\right\}^{1/2} \in L^p(S).$$

Also, when $s = k/2$, $f \in H^p_s$ if and only if $(T_1 \cdots T_k f)^+ f \in L^p(S)$. We point out that these characterizations do not involve the $\bar{T}_{ij}$. The same result but allowing as well T_{ij} and $\bar{T}_{ij}$ would be much easier to prove because the normal derivative appears as the Lie bracket of T_{ij} and $\bar{T}_{ij}$'s.

We want to point out a basic difference that occurs, as in the real variable theory of Hardy spaces, between the area functions and the others. Indeed, for the area functions there are pointwise estimates, in the sense that the integral appearing in (1) and the $A^t_{\alpha, t-s}$ are pointwise equivalent, (up to replacement of the aperture α), but there are no such estimates for the others, *e.g.* it is not possible to bound $(T^2_{ij} f)^+$ by $(Rf)^+$ or $(Rf)^*$, pointwise.

For example, with $n = 2$, let $f(z, w) = w^2 \log(1 - z)$ then Rf is bounded in B^2 but $T^2 f$ is not (here $T = \bar{z}(\partial/\partial w) - \bar{w}(\partial/\partial z)$). For this reason it is easier to deal with the area functions.

We also obtain analogous results for the Bergman-Sobolev spaces

$$B^p_{s,\gamma} = \left\{f \text{ holomorphic: } \int_{B^n} |R^s f(z)|^p (1 - |z|)^\gamma \, dm(z) < +\infty\right\},$$

$p > 0$, $\gamma > -1$. Namely, one can replace $R^{k/2}$ by k complex-tangential vector fields to define these spaces. The paper [1] deals with one aspect of this.

This paper is organized as follows. In the second section we collect all properties that will be needed of functions annihilated by some power of $\bar{\partial}$ on every complex line. In the third section we prove our results for the case of area functions and in the fourth section we deal with the g-functions and maximal functions. The fifth section is more technical and deals with the following question, in case $s = k/2$: if $R^{k/2} \in H^p$ then it has a distribution boundary value which belongs to the space H^p_{at} of [4]. Is this also true for $T_1 \cdots T_k f$? We show in Section 5 that this is the case, in spite of the fact that $T_1 \cdots T_k f$ is not holomorphic. Finally in section 6 we prove our result on Bergman-Sobolev spaces.

In the remaining of this section we introduce some more notations. First, as it is easily checked, the operators T_{ij} commute with the ordinary Laplacian

Δ, so that the functions $T_1 \cdots T_k f$ are harmonic. It will be then convenient to consider R^s as defined in the class of harmonic functions. If u is harmonic and $u = \sum_k P_k$ is its expansion in spherical harmonics, then we define as well $R^s u = \sum_k (k+1)^s P_k$.

The operator I^s of fractional integration is defined

$$I^s f(z) = \frac{1}{\Gamma(s)} \int_0^1 \left(\log \frac{1}{t} \right)^{s-1} f(tz)\, dt.$$

Then

$$I^s P_k = (k+1)^{-s} P_k$$

if P_k is homogeneous of degree k. Thus I^s is the inverse of R^s. This simple expression of the inverse of R^s is the technical reason for which we use R instead of N. When $s = 1$ we will write simply I. It is easily checked that I commutes with the T_{ij} and the $\bar{T}_{ij}$.

We will consider differential operators X appearing as composition

$$Xf = X_1 \cdots X_k f$$

where each X_i is R, a T_{ij} or a $\bar{T}_{ij}$. For such an operator its *weight* is defined to be $\sum_1^k w(X_k)$, the weight of R being 1 and 1/2 the weight of each T_{ij} and $\bar{T}_{ij}$. The functions Xf are most easily estimated in terms of $d^k f(X_1, \ldots, X_k)$. This is the function whose value at $z \in B^n$ is

$$\frac{\partial^k f}{\partial v_1 \cdots \partial v_k}(z)$$

with $v_j = X_j(z)$. It is easily checked by induction that Xf can be written as a linear combination of $d^j f(Y_1, \ldots, Y_j)$ with $j \leqslant k$, and

$$\sum_1^j w(Y_i) \leqslant \sum_1^k w(X_i),$$

so that when proving estimates on Xf involving its weight we will instead estimate $d^k f(X_1, \ldots, X_k)$, which is simpler because the X_j are «frozen» at z.

One final remark is in order. As said before some of our results depend on the special properties that the $T_{ij} f$ have for an holomorphic f in the ball and thus do not generalize to other domains. It is probably the case that our results hold for more general domains, but this would require other methods, as maybe the «freezing coefficient» technique. The technical difficulty involved would be similar as the one encountered when trying to generalize the Fefferman-Stein paper on real H^p spaces to a general smooth domain.

2. On Functions Annihilated by Powers of $\bar{\partial}$

If f is holomorphic in B^n and $T_1, \ldots, T_k$ are complex-tangential vector fields then of course $F = T_1 \cdots T_k f$ is no longer holomorphic. Indeed if we fix $z_0 \in B^n$, $\zeta_0 \in \mathbb{C}^n$ then the function $u(\lambda) = F(z_0 + \lambda\zeta_0)$ need not be harmonic not even subharmonic. Nevertheless we will see that the function u does satisfy some mean value inequalities that we can exploit to obtain «non-isotropic» mean value estimates for the function F. It is clear that u does satisfy the differential equation $(\partial^k/\partial\bar{\lambda}^{k+1})u \equiv 0$.

Definition. *Let $U \subseteq \mathbb{C}$ be an open set; by $H_k(U)$ we mean the set of functions u defined on U so that $(\partial^k/\partial\bar{\lambda}^k)u(\lambda) \equiv 0$ in U. By $H_k^p(U)$ we mean those u in $H_k(U)$ so that*

$$\int_U |u|^p \, dm = \|u\|_p^k < \infty.$$

It is easily proved, by induction on k, that if $u \in H_k(U)$ and $\lambda_0 \in \mathbb{C}$, then u has a unique expression

$$u(\lambda) = \sum_{j=0}^{k-1} (\bar{\lambda} - \bar{\lambda}_0)^j f_j(\lambda)$$

where f_j is holomorphic in U.

Lemma 2.1. *Given non-negative integers l, m there is a constant $C = C(l, m)$ such that if $u \in H_k^1(D(\lambda_0, r))$, then*

$$\left| \frac{\partial^{l+m}}{\partial\lambda^l \partial\bar{\lambda}^m} u(\lambda_0) \right| \leqslant \frac{C}{r^{l+m+2}} \int_{D(\lambda_0, r)} |u(\lambda)| \, dm(\lambda).$$

PROOF. By a translation and dilation it is enough to prove this when $\lambda_0 = 0$ and $r = 1$. Moreover we may assume that $u \in H_k(D(0, R))$ for some $R > 1$. For a given l, m we will find a polynomial $p = p_{l,m}$ so that

$$\frac{\partial^{l+m}}{\partial\lambda^l \partial\bar{\lambda}^m} u(0) = \int_{|\lambda|<1} u(\lambda) p(\lambda) \, dm(\lambda).$$

The result will follow with

$$C = \sup_{|\lambda|<1} |p(\lambda)|.$$

As observed above we have

$$u(\lambda) = \sum_{j=0}^{k-1} \bar{\lambda}^j f_j(\lambda),$$

with f_j holomorphic. Clearly we may assume that $m \leqslant k-1$. Then

$$\frac{\partial^{l+m}}{\partial\lambda^l \partial\bar{\lambda}^m} u(0) = m! f_m^{(l)}(0).$$

We try

$$p(\lambda) = \lambda^m \bar{\lambda}^l \sum_{\nu=0}^{k-1} \alpha_\nu |\lambda|^{2\nu},$$

where the numbers α_ν are to be determined.

$$\int_{|\lambda|<1} u(\lambda)p(\lambda)\, dm\,(\lambda) = \sum_{j,\nu=0}^{k-1} \alpha_\nu \int \bar{\lambda}^j \lambda^m \bar{\lambda}^l |\lambda|^{2\nu} \alpha_\nu f_j(\lambda)\, dm\,(\lambda).$$

Each f_j has a power series expansion

$$f_j(\lambda) = \sum_{n=0}^{\infty} \hat{f}_j(n)\lambda^n.$$

If we insert this above we obtain, after using polar coordinates,

$$\begin{aligned}
&\sum_{j,\nu=0}^{k-1} \alpha_\nu \int_0^1 r^{1+j+m+l+2\nu} \int_0^{2\pi} e^{i(m-j-l)\theta} \sum \hat{f}_j(n) r^n e^{in\theta} \frac{d\theta}{2\pi}\, dr \\
&\qquad = \sum_{j,\nu=0}^{k-1} \alpha_\nu \hat{f}_j(j+l-m) \int_0^1 r^{1+j+m+l+2\nu+j+l-m}\, dr \\
&\qquad = \sum_{j=0}^{k-1} \left\{ \sum_{\nu=0}^{k-1} \frac{\alpha_\nu}{2(1+j+l+\nu)} \right\} \hat{f}_j(j+l-m).
\end{aligned}$$

So, we wish to choose the numbers $\{\alpha_\nu\}$ so that

$$\sum_{\nu=0}^{k-1} \frac{\alpha_\nu}{1+j+\nu+l} = \begin{cases} 2m!l! & \text{when } j = m, \\ 0 & j \neq m. \end{cases}$$

For a fixed $l \geqslant 0$, the finite Hilbert matrix

$$(H_{j\nu}) = \left(\frac{1}{1+j+\nu+l} \right)_{j,\nu=0}^{k-1}$$

is always invertible so this is possible. The proof of the lemma is complete. □

As a simple consequence of the lemma we get a boundary growth estimate.

Corollary 2.2. *There is a constant $C = C(l, m)$ so that if $u \in H_k^1(U)$ and $\lambda_0 \in U$ then*

$$\left|\frac{\partial^{l+m}u}{\partial\lambda^l\,\partial\bar{\lambda}^m}(\lambda_0)\right| \leqslant \frac{C}{\delta^{k+l+2}}\,\|u\|_1$$

where $\delta = \operatorname{dist}(\lambda_0, \partial U)$.

PROOF. Just apply the lemma to a disc of radius δ about λ_0. □

Next we prove an analogue of a well-known theorem of Hardy and Littlewood. In fact with the above lemma and its corollary we can use their proof. We include the details for completeness.

Lemma 2.3. *Suppose* $0 < p < \infty$; *then there is a constant* C_p *such that if* $u \in H_k(D(\lambda_0, r))$ *then*

$$|u(\lambda_0)|^p \leqslant \frac{C_p}{r^2}\int_{D(\lambda_0, r)} |u(\lambda)|^p\, dm(\lambda).$$

PROOF. The case $p = 1$ is just the lemma with $l = m = 0$. The case $p > 1$ follows from the case $p = 1$ by Hölder's inequality. For the case $0 < p < 1$ we may assume that

$$\lambda_0 = 0, \qquad r = 1 \quad \text{and} \quad u \in H_k(D(0, R)), \qquad R > 1.$$

Also assume that

$$\int_{|\lambda|<1} |u(\lambda)|^p\, dm(\lambda) = 1$$

and $|u(0)| > 1$. Similarly to [3] we let

$$m_p^p(r) = \int_{|\lambda|<r} |u(\lambda)|^p\, dm(\lambda)$$

and $m_\infty(r) = \sup_{|\lambda|\leq r} |u(\lambda)|$. Since $0 < p < 1$,

$$\begin{aligned} m_1(r) &= \int_{|\lambda|<r} |u(\lambda)|\, dm(\lambda) \\ &\leqslant \int_{|\lambda|<r} |u(\lambda)|^p |u(\lambda)|^{1-p}\, dm(\lambda) \\ &\leqslant m_\infty^{1-p}(r) m_p^p(r) \\ &\leqslant m_\infty^{1-p}(r). \end{aligned}$$

Now if we apply Corollary 2.2 with $l = m = 0$ to $D(0, r)$ with $r < 1$ we obtain

$$m_\infty(\rho) \leqslant C\frac{m_1(r)}{(r-\rho)^2}, \qquad 0 < \rho < r < 1.$$

So we have

$$m_\infty(\rho) \leqslant C \frac{m_\infty^{1-p}(r)}{(r-\rho)^2}, \qquad 0 < \rho < r < 1.$$

We take $\rho = r^\alpha$ with $\alpha > 1$ to be determined, take logarithms and integrate

$$\int_{1/2}^1 \log m_\infty(r^\alpha) \frac{dr}{r} \leqslant \log C + 2 \int_{1/2}^1 \log \frac{1}{(r - r^\alpha)} \frac{dr}{r} + (1-p) \int_{1/2}^1 \log m_\infty(r) \frac{dr}{r}.$$

Letting $t = r^\alpha$ in the left hand side and rearranging we get

$$\left[\frac{1}{\alpha} - (1-p)\right] \int_{1/2}^1 \log m_\infty(r)\, dr \leqslant \frac{1}{2} \log C + 2 \int_{1/2}^1 \log \frac{1}{r - r^\alpha} \frac{dr}{r}.$$

Now choose α so close to 1 that the coefficient on the left hand side is positive. Then we get

$$\log m_\infty(1/2) \leqslant 4 \int_{1/2}^1 \log m_\infty(r)\, dr \leqslant C = C_p. \quad \square$$

Next we see that there is a version of Lemma 2.1 valid for all p.

Lemma 2.4. *If m, l are non-negative integers and $0 < p < \infty$ are given then there is a constant C such that if $u \in H_k^1(D(\lambda_0, r))$ then*

$$\left| \frac{\partial^{l+m}}{\partial \lambda^l \partial \bar{\lambda}^m} u(\lambda_0) \right|^p \leqslant \frac{C}{r^{2+p(l+m)}} \int_{D(\lambda_0, r)} |u(\lambda)|^p\, dm(\lambda).$$

PROOF. Let Du denote $\dfrac{\partial^{l+m}}{\partial \lambda^l \partial \bar{\lambda}^m} u$ and $\nu = l + m$. By Lemma 2.1, (the case $p = 1$) we have

$$|Du(\lambda_0)| \leqslant \frac{C}{(r/2)^{2+\nu}} \int_{D(\lambda_0, r/2)} |u(\lambda)|\, dm(\lambda).$$

For each $\lambda \in D(\lambda_0, r/2)$ we have by Lemma 2.3

$$|u(\lambda)|^p \leqslant \frac{C}{(r/2)^2} \int_{D(\lambda, r/2)} |u(w)|^p\, dm(w) \leqslant \frac{C}{r^2} \int_{D(\lambda_0, r)} |u(w)|^p\, dm(w).$$

If we insert this into the above estimate for $Du(\lambda_0)$ we have

$$|Du(\lambda_0)| \leqslant \frac{C}{r^{2+\nu}} \int_{D(\lambda_0, r/2)} \left(\frac{1}{r^2} \int_{D(\lambda_0, r)} |u(w)|^p \, dm(w) \right)^{1/p} dm(\lambda)$$

$$= \frac{C}{r^{2/p+\nu}} \left(\int_{D(\lambda_0, r)} |u(\lambda)|^p \, dm(\lambda) \right)^{1/p},$$

which completes the proof of the lemma. □

We want to point out that what the above arguments actually show is the following: Suppose u is defined in a domain U and there is a constant C such that whenever $D(\lambda_0, r) \subseteq U$ we have

$$|u(\lambda_0)| \leqslant \frac{C}{r^2} \int_{D(\lambda_0, r)} |u(\lambda)| \, dm(\lambda).$$

Then for each $p, 0 < p < \infty$, there is a constant C_p such that

$$|u(\lambda_0)|^p \leqslant \frac{C_p}{r^2} \int_{D(\lambda_0, r)} |u(\lambda)|^p \, dm(\lambda)$$

whenever $D(\lambda_0, r) \subseteq U$.

We let $H_k(B^n)$ be the class of C^∞ functions in B^n that are annihilated by $\bar{\partial}^k$ on every complex line. More precisely we say that $u \in H_k(B^n)$ if for every $z \in B^n$ and every $0 \neq \zeta \in \mathbb{C}^n$ we have

$$\frac{d^k}{d\bar{\lambda}^k} u(z + \lambda\zeta) \equiv 0.$$

Let u be a C^∞ function in B^n. Take $z \in B^n$, and $\zeta \in \mathbb{C}^n$. Then we may calculate that

$$\frac{d^k}{d\bar{\lambda}^k} [u(z + \lambda\zeta)]_{\lambda = 0} = \sum_{|\beta| = k} \bar{D}^\beta u(z) \bar{\zeta}^\beta.$$

From this we deduce that a C^∞ function u lies in $H_k(B^n)$ if and only if $\bar{D}^\beta u = 0$ in B^n for all multi-indexes β with $|\beta| = k$. If follows that if $u \in H_k(B^n)$ then $\bar{D}^\beta D^\gamma u \in H_k(B^n)$ for any β and γ.

In the next lemma, for $z \in B^n$ and $\delta > 0$ the polydisc $P(z, \delta)$ is defined as follows. If $z = r\zeta$, pick $\zeta_2, \ldots, \zeta_n$ so that $\{\zeta, \zeta_2, \ldots, \zeta_n\}$ is an orthonormal basis for $\mathbb{C}^n$. Then

$$P(z, \delta) = \left\{ w = r\zeta + \lambda\zeta + \sum_{j=2}^{n} \lambda_j \zeta_j : |\lambda| < \delta, |\lambda_j| < \delta^{1/2}, j = 2, \ldots, n \right\}.$$

Lemma 2.5. *Suppose $0 < p < \infty$, $1 \leqslant k$ and a differential operator X of weight m are given. Then there is a constant C such that*

$$|Xu(z)|^p \leqslant \frac{C}{\delta^{n+1+mp}} \int_{P(z,\delta)} |u(w)|^p \, dm(w)$$

for all $u \in H_k(B^n)$, as long as $P(z, \delta) \subset B^n$.

PROOF. For convenience we assume $z = (r, 0, \ldots, 0)$. By the discussion in the introduction it is enough to obtain the above estimate with Xu replaced by

$$\frac{\partial^N}{\partial z_1^{k_1} \partial \bar{z}_1^{l_1} \cdots \partial z_n^{k_n} \partial \bar{z}_n^{l_n}} u(\lambda, 0)$$

where

$$\sum_{j=1}^{n} (k_j + l_j) = N$$

and

$$k_1 + l_1 + \frac{1}{2} \sum_{j=2}^{n} (k_j + l_j) \leqslant m.$$

Recalling that if $u \in H_k(B^n)$ then the same is true of any partial derivative of u we may apply Lemma 2.4 successively, one variable at a time, to obtain the desired result. □

3. Characterizations of H_s^p in Terms of Area Functions

The main purpose of this section is to obtain characterizations of H_s^p in terms of area functions, some of them involving only complex-tangential derivatives. Our starting point for this and next section is the following account of different characterizations of H^p itself, already mentioned in the introduction.

Theorem 3.1 ([8]). *The following are equivalent, (with an aperture α fixed):*

(*a*) $f \in H^p$.
(*b*) $A_\alpha(f) \in L^p(S)$.
(*c*) $g(f) \in L^p(S)$.
(*d*) $f_\alpha^* \in L^p(S)$.
(*e*) $f_\alpha^+ \in L^p(S)$.

The Theorem is well-known. The equivalence of (*a*), (*d*) and (*e*) is Koranyi's result. The conditions (*b*) and (*c*) are also known (see [8]), though a detailed

proof seems to be lacking in the literature. A proof can be obtained by adapting the methods of Fefferman and Stein. This requires as in [3] the use of a version of Green's theorem for regions that appear as unions of admissible aproach regions, particularly in the equivalence of (*b*) and (*d*). We point out that for the particular case of the ball an easy proof can be obtained as follows. That (*b*) implies (*c*) is a consequence of a pointwise estimate (see Lemma 4.3 below), (*c*) implies (*a*) can be obtained by slice integration of the same result in dimension 1 and (*a*) $\Rightarrow$ (*b*) can be proved using the atomic decomposition of Garnett and Latter ([4]).

To state our result we need to introduce some more notations. For an operator X as in the introduction

$$Xf = X_1 \cdots X_k f$$

and $\gamma > 0$ we define the area function

$$A^X_{\alpha,\gamma} f(\eta) = \left\{ \int_{D_\alpha(\eta)} |Xf(z)|^2 (1 - |z|)^{2\gamma - 1 - n}\, dm(z) \right\}^{1/2}.$$

We will consider in particular operators $T = T_1 \cdots T_k$ where the $T_1, \ldots, T_k$ are chosen among the T_{ij} (not the $\bar{T}_{ij}$'s). We denote by $\{T_\delta\}$, $\delta \in C_k$, the collection of such operators and define the *k-th complex-tangential gradient* as

$$\nabla^k_T f = \sum_{\delta \in C_k} |T_\delta f|.$$

In this section we will prove

Theorem 3.2. *Let $f \in H^p_s$ and let X be a differential operator as above of weight $m > s$. Then $A^X_{\alpha, m-s} f \in L^p(S)$ for any α.*

Theorem 3.3. *If $l \geqslant 0$ and $k \in Z$ are such that $l + k/2 > s$, then $f \in H^p_s$ if and only if*

$$\left\{ \int_{D_\alpha(\eta)} |\nabla^k_T R^l f(z)|^2 (1 - |z|)^{2l + k - 2s - 1 - n}\, dm(z) \right\}^{1/2} \in L^p(S).$$

We will need a version of Hardy's inequality for which we have not found a reference so we include it here.

Lemma 3.4. *If $\alpha, \beta > 0$, $1 \leqslant p < \infty$ there is a constant C so that for all $f \geqslant 0$ we have,*

$$\int_0^1 (1 - r)^{\alpha - 1} \left(\int_0^r (r - t)^{\beta - 1} f(t)\, dt \right)^p dr \leqslant C \int_0^1 (1 - r)^{\alpha - 1 + \beta p} f(r)^p\, dr.$$

PROOF. We give a sketch. After changing variables the left hand side of the above becomes

$$\int_0^1 r^{\alpha-1}\left(\int_r^1 (t-r)^{\beta-1} f(1-t)\,dt\right)^p dr.$$

Letting $t = rs$, this becomes

$$\int_0^1 r^{\alpha-1}\left(\int_1^{1/r} r^{\beta}(s-1)^{\beta-1} f(1-rs)\,ds\right)^p dr.$$

If we now use the continuous form of Minkowski's inequality and change variables again we have that the above is at most

$$\left[\int_1^\infty \frac{(x-1)^{\beta-1}}{x^{\beta+\alpha/p}}\,dx\right]^p \int_0^1 (1-r)^{\alpha-1+p\beta} f(r)^p\,dr.$$

The integral occurring in the first factor is convergent since $\alpha, \beta > 0$. □

In the next geometric lemma, for an approach region $D_\alpha(\eta)$ and $0 < r < 1$, $S_\alpha(r, \eta)$ denotes the region

$$S_\alpha(r,\eta) = \left\{z \in D_\alpha(\eta): \frac{1}{2}(1-r^2) < 1 - |z|^2 < 2(1-r^2)\right\}.$$

Lemma 3.5. *For each α, β, $1 < \alpha < \beta$, there is an $\epsilon = \epsilon(\alpha, \beta)$ such that if $\eta \in S$ and $z = r\zeta \in D_\alpha(\eta)$ then the polydisc $P(z, \epsilon(1-r^2))$ (as defined before Lemma 2.5) is contained in $S_\beta(r, \eta)$.*

PROOF. Take

$$w = (r+\lambda)\zeta + \sum_{j=2}^n \lambda_j \zeta_j \in P(z,\delta),\ \delta = \epsilon(1-r^2);$$

then

$$|w|^2 = |r+\lambda|^2 + \sum_{j=2}^n |\lambda_j|^2 \leqslant r^2 + |\lambda|^2 + 2|\lambda| + (n-1)\delta \leqslant r^2 + (n+2)\delta.$$

Also

$$|w|^2 \geqslant (r - |\lambda|)^2 = r^2 - 2r|\lambda| + |\lambda|^2 \geqslant r^2 - 2|\lambda| \geqslant r^2 - 2\delta.$$

So

$$(1-r^2) - (n+2)\delta \leqslant 1 - |w|^2 \leqslant (1-r^2) + 2\delta$$

so we have

$$\frac{1}{2}(1-r^2) \leqslant 1-|w|^2 \leqslant 2(1-r^2)$$

for sufficiently small ϵ. Next we calculate

$$\begin{aligned}
|1-\langle w,\eta\rangle| &= \left|1-r\langle\zeta,\eta\rangle+\lambda\langle\zeta,\eta\rangle+\sum_{j=2}^{n}\lambda_j\langle\zeta_j,\eta\rangle\right| \\
&\leqslant \frac{\alpha}{2}(1-r^2)+\delta+\sum_{j=2}^{n}\delta^{1/2}|\langle\zeta_j,\eta\rangle| \\
&= \frac{\alpha}{2}(1-r^2)+\delta+\sum_{j=2}^{n}\delta^{1/2}|\langle\zeta_j,\eta-r\zeta\rangle| \\
&\leqslant \frac{\alpha}{2}(1-r^2)+\delta+n\delta^{1/2}|\eta-r\zeta| \\
&\leqslant \frac{\alpha}{2}(1-r^2)+\delta+n\delta^{1/2}\sqrt{2}\,|1-r\langle\zeta,\eta\rangle|^{1/2} \\
&\leqslant \frac{\alpha}{2}(1-r^2)+\delta+n\sqrt{2}\left(\frac{\alpha}{2}(1-r^2)\right)^{1/2}\delta^{1/2} \\
&= \left(\frac{\alpha}{2}+\epsilon+n\alpha^{1/2}\epsilon^{1/2}\right)(1-r^2).
\end{aligned}$$

Also from the first part of the proof

$$\frac{\beta}{2}(1-|w|^2) \geqslant \frac{\beta}{2}[1-(n+2)\epsilon](1-r^2)$$

and so

$$|1-\langle w,\eta\rangle| < \frac{\beta}{2}(1-|w|^2)$$

for all $w \in P(z, \epsilon(1-r^2))$ for ϵ sufficiently small. □

Lemma 3.6. *Fix $1<\alpha<\beta$, $q \geqslant 1$, $l \geqslant 0$, $\gamma > 0$ and let X be an operator of weight m. Then there is a constant C such that for all $f \in H_q(B^n)$ and all $\eta \in S$ we have*

$$\begin{aligned}
(3)\quad &\int_{D_\alpha(\eta)} |I^l Xf(z)|^2(1-|z|)^{2\gamma-n-1}\,dm(z) \\
&\qquad \leqslant C\int_{D_\beta(\eta)} |f(z)|^2(1-|z|)^{2(\gamma+l-m)-n-1}\,dm(z).
\end{aligned}$$

PROOF. As already mentioned, the proof follows the same lines as in [7, pp. 214-216]. Combining Lemma 2.5 with Lemma 3.5 we see that if $r\zeta \in D_\alpha(\eta)$ then

$$|Xf(r\zeta)| \leqslant \left\{ \frac{c}{(1-r)^{n+1+2m}} \int_{S_\beta(r,\eta)} |f(w)|^2 \, dm(w) \right\}^{1/2} = J_r. \tag{4}$$

Assume $l > 0$. In this case,

$$|I^l Xf(r\zeta)| \leqslant \frac{c}{r} \int_0^r (r-t)^{l-1} J_t \, dt, \qquad r > 0.$$

Having obtained a bound for $I^l Xf(r\zeta)$ which just depends on r, we integrate in polar coordinates in $D_\alpha(\eta)$, using the fact that for fixed r,

$$\sigma\{\zeta \in S: r\zeta \in D_\alpha(\eta)\} \leqslant c(1-r)^n.$$

This gives that the left hand side of (3) is bounded by

$$C \int_0^1 (1-r)^{2\gamma-1} \left(\int_0^r (r-t)^{l-1} J_t \, dt \right)^2 dr$$

and, using Lemma 3.4, by

$$C \int_0^1 (1-r)^{2\gamma-1+2l} J_r^2 \, dr. \tag{5}$$

Inserting the definition of J_r we obtain the bound

$$\int_0^1 (1-r)^{2(\gamma+l-m-1)-n} \int_{S_\beta(r,\eta)} |f(w)|^2 \, dm(w) \, dr.$$

If $w \in S_\beta(r,\eta)$, then $1-|w|$ is comparable to $1-r$ hence the above integral is dominated by

$$\int_{D_\beta(\eta)} (1-|z|)^{2(\gamma+l-m)-n-1} |f(z)|^2 \, dm(z).$$

For $l = 0$ we use the same argument to pass directly from (4) to (5). □

PROOF OF THEOREM 3.2. If X is as in Theorem 3.2, we apply Lemma 3.6 to $R^{1+s}f$ instead of f, $l = 1+s$ and $\gamma = m-s$ to obtain (recall that I^l commutes with the T_{ij})

$$A^X_{\alpha,m-s} f(\eta) \leqslant c A^{1+s}_{\beta,1} f(\eta) = c A_\beta(R^s f)(\eta).$$

By Theorem 3.1, $A_\beta(R^s f) \in L^p(S)$ and so part (*a*) is proved. □

PROOF OF THEOREM 3.3. Now, taking in Lemma 3.6 as X powers of R it is easily checked by symmetry the following

Corollary 3.7. *For $t \in \mathbb{R}$ and $\gamma > 0$, recall the notation*

$$A^t_{\alpha,\gamma} f(\eta) = \left\{ \int_{D_\alpha(\eta)} |R^t f(z)|^2 (1 - |z|)^{2\gamma - n - 1}\, dm(z) \right\}^{1/2}.$$

(where $R^t = I^{-t}$ if $t < 0$). Then if, $t, l \in \mathbb{R}$ and $\gamma, r > 0$ are such that $l - t = r - \gamma$, the area functions $A^t_{\alpha,\gamma} f$ and $A^l_{\alpha,r} f$ are pointwise equivalent, up to replacement of the apertures.

Now, by Theorem 3.1, $f \in H^p_s$ is equivalent to $A_\alpha(R^s f) = A^{1+s}_{\alpha,1}(f)$ being in $L^p(S)$. By Corollary 3.7, it follows that $f \in H^p_s$ if and only if $A^l_{\alpha,r} f \in L^p(S)$, for each α and l, τ, $\tau > 0$, with $l - \tau = s$, which proves Theorem 3.3 in case $k = 0$. To finish the proof of Theorem 3.3 we must show that in this characterization we can replace $R^{k/2}$ by $T_1 \cdots T_k$. This will be a consequence of

Theorem 3.8. *Fix $\gamma > 0$ and let E be the space of all functions such that*

$$\left\{ \int_{D_\alpha(\eta)} |f(z)|^2 (1 - |z|)^{2\gamma - n - 1}\, dm(z) \right\}^{1/2} \in L^p(S)$$

for all α. Let k be any positive integer. Then, if $f \in H_q(B^n)$ for some q, f belongs to E if and only $\nabla^k_T I^{k/2} f \in E$.

Note that we just need this lemma for an holomorphic f, but one is lead to the above statement by the non-analyticity of the $T_1 \cdots T_k I^{k/2} f$.

Choosing $X = T_1 \cdots T_k$ and $I^l = I^{k/2}$ in Lemma 3.6, we obtain that $f \in E$ implies $T_1 \cdots T_k I^{k/2} f \in E$. For the reverse implication, we need in turn several Lemmas. In the first one, $H(p, q)$ denotes the space of harmonic polynomials of bidegree (p, q).

Lemma 3.9. $\sum \bar{T}_{ij} T_{ij} = -2p(q + n - 1)Id$ *on $H(p, q)$, the sum being extended over all pairs (i, j) $1 \leqslant i, j \leqslant n$.*

PROOF. It is convenient for the proof *not* to exclude the case $i = j$, even though $T_{ii} = 0$. With this in mind we calculate that

$$\bar{T}_{ij} T_{ij} = |z_i|^2 \frac{\partial^2}{\partial z_j \partial \bar{z}_j} + |z_j|^2 \frac{\partial^2}{\partial z_i \partial \bar{z}_i} - z_i \bar{z}_j \frac{\partial^2}{\partial z_i \partial \bar{z}_j} - \bar{z}_i z_j \frac{\partial^2}{\partial \bar{z}_i \partial z_j}$$
$$- z_i \frac{\partial}{\partial z_i} - z_j \frac{\partial}{\partial z_j} + \delta_{ij} \left[z_i \frac{\partial}{\partial z_i} + z_j \frac{\partial}{\partial z_j} \right].$$

Now if we add on *all* i and j we have

$$2|z|^2 \Delta - 2 \sum_{i,j=1}^{n} z_i \bar{z}_j \frac{\partial^2}{\partial z_i \partial \bar{z}_j} - 2n \sum_{i=1}^{n} z_i \frac{\partial}{\partial z_i} + 2 \sum_{i=1}^{n} z_i \frac{\partial}{\partial z_i}.$$

Now if we apply this to a function in $H(p, q)$ the part involving Δ is zero so we need only check the action of the rest of the terms on an element of $H(p, q)$. Now each element of $H(p, q)$ is a linear combination of terms of the form $z^\alpha \bar{z}^\beta$ where $|\beta| = p$, $|\beta| = q$. Applying the terms not involving Δ to $z^\alpha \bar{z}^\beta$ yields

$$\left\{-2\sum_{i,j=1}^{n}\alpha_i\beta_j - 2(n-1)\sum_{i=1}^{n}\alpha_i\right\}z^\alpha\bar{z}^\beta = -2(|\alpha|\,|\beta| + 2(n-1)|\alpha|)z^\alpha\bar{z}^\beta$$

$$= -2p(q+n-1)z^\alpha\bar{z}^\beta. \quad \square$$

In the statement of next Lemma, if $T_\delta = T_i \cdots T_1$ we define $\overline{T_\delta} = \overline{T_k} \cdots \overline{T_1}$.

Lemma 3.10.

$$\tau_k = \sum_{\delta \in C_k} \bar{T}_\delta T_\delta = (-2)^k \frac{p!}{(p-k)!} \frac{(q+k+n-2)!}{(q+n-2)!} \quad \text{on} \quad H(p, q).$$

PROOF. We do induction on k. The case $k = 1$ is just Lemma 3.9.

Take $u \in H(p, q)$, then

$$\tau_k u = \sum_{\delta \in C_{k-1}} \bar{T}_\delta \sum_{i,j} \bar{T}_{ij} T_{ij} T_\delta u.$$

Now $T_\delta u \in H(p-k+1, q+k-1)$ and so

$$\sum_{i,j} \overline{T_{ij}} T_{ij} T_\delta u = -2(p-k+1)(q+k+n-1)T_\delta u,$$

by Lemma 3.9. We see that

$$\tau_k u = -2(p-k+1)(q+k+n-1)\tau_{k-1}u$$

$$= (-2)^{k-1}(-2)\frac{p!}{p-(k-1))!}\frac{(q+k+n-2)!}{(q+n-2)!}(p-k+1)(q+k)u$$

$$= (-2)^k \frac{p!}{(p-k)!}\frac{(q+k+n-2)!}{(q+n-2)!}u$$

as promised. $\square$

Lemma 3.11. *For f holomorphic in B^n we have*

$$\sum_{\delta \in C_k} I^{k/2}\bar{T}_\delta T_\delta I^{k/2} = \sum_{j=0}^{k} d_j I^j f$$

for some constants $d_0, \ldots, d_k$ depending only on k, $d_0 \neq 0$.

PROOF. Take an holomorphic monomial $z^\alpha \in H(p, 0)$, $p = |\alpha| > k$. Then

$$\tau_k z^\alpha = (-2)^k \frac{p!}{(p-k)!} \frac{(k+n-2)!}{(n-2)!} z^\alpha = c_k \frac{p!}{(p-k)!} z^\alpha.$$

Now $p!/(p-k)!$ is a polynomial of degree k in p whose coefficients depend only on k. If we express this polynomial as a function of $p+1$ we obtain

$$\tau_k z^\alpha = C_k(p+1)^k z^\alpha + C_{k-1}(p+1)^{k-1} z^\alpha + \cdots + C_0 z^\alpha = \left(\sum_{j=0}^{k} C_j R^j \right) z^\alpha.$$

This shows that

$$\tau_k f = \sum_{j=0}^{k} c_j R^j f, \qquad c_k \neq 0$$

for an holomorphic function f. We now just apply I^k to both sides, using again the fact that the I^k commute with the T_δ. □

Lemma 3.12. *Let*

$$P(x) = \sum_{j=1}^{k} c_j x^j$$

be a polynomial and denote

$$P(I)f = \sum_{j=1}^{k} c_j I^j f.$$

Then there are c, r_0 depending on α and P, c, $r_0 < 1$ such that for f holomorphic

$$\int_{D_\alpha(\eta)} |f(z)|^2 (1-|z|)^{2\gamma-n-1}\, dm(z) \leqslant c \int_{D_\alpha(\eta)} |f + P(I)f(z)|^2 (1-|z|)^{2\gamma-n-1}\, dm(z) + c \sup\{|f(z)|^2 : |z| \leqslant r_0\}.$$

PROOF. Note that $\{f : f + P(I)f = 0\}$ is a finite dimensional subspace, hence all norms are equivalent on it and thus it really does not matter what norm is considered in the last term. We will prove that

$$(6) \quad \int_{D_\alpha(\eta)} |If(z)|^2 (1-|z|)^{2\gamma-n-1}\, dm(z) \leqslant c \int_{D_\alpha(\eta)} |f(z)|^2 (1-|z|)^{2\gamma-n}\, dm(z).$$

This implies by iteration that associated to P there is a constant M such that

$$\int_{D_\alpha(\eta)} |P(I)f(z)|^2 (1-|z|)^{2\gamma-n-1}\, dm(z) \leqslant M \int_{D_\alpha(\eta)} |f(z)|^2 (1-|z|)^{2\gamma-n}\, dm(z)$$

Hence

$$\int_{D_\alpha(\eta)} |f(z)|^2(1-|z|)^{2\gamma-n-1}\,dm(z)$$
$$\leqslant c\int_{D_\alpha(\eta)} |f+P(I)f(z)|^2(1-|z|)^{2\gamma-n-1}\,dm(z)$$
$$+ Mc\sup\{|f(z)|^2: |z|\leqslant r_0\}$$
$$+ M(1-r_0)\int_{D_\alpha(\eta)} |f(z)|^2(1-|z|)^{2\gamma-n-1}\,dm(z).$$

and the lemma will follow choosing r_0 such that $M(1-r_0)<1$.

Note that (6) is similar to (3) but the aperture of the admissible approach region is not changed in (6). To prove it, we use polar coordinates in $D_\alpha(\eta)$ and

$$|If(r\zeta)|^2 \leqslant \frac{1}{r}\int_0^r |f(t\zeta)|^2\,dt$$

to bound the left hand side of (6) by

$$\int_0^1 (1-r)^{2\gamma-n-1}r^{2n-2}\,dr\int_{A(r,\eta)} d\sigma(\zeta)\int_0^r |f(t\zeta)|^2\,dt.$$

Here $A(r,\eta)=\{\zeta\in S: r\zeta\in D_\alpha(\eta)\}$. Now we use Fubini's theorem: denoting

$$I(t,r)=\int_{A(r,\eta)} |f(t\zeta)|^2\,d\sigma(\zeta)$$

and using that $I(t,r)\leqslant I(t,t)$ for $t\leqslant r$ (because $A(r,\eta)\subset A(t,\eta)$), we obtain the bound

$$\int_0^1 (1-r)^{2\gamma-n-1}\,dr\int_0^r I(t,t)\,dt = c\int_0^1 I(t,t)(1-t)^{2\gamma-n}\,dt. \tag{7}$$

On the other hand by the same argument, the right hand side equals

$$c\int_0^1 (1-t)^{2\gamma-n}t^{2n-1}I(t,t)\,dt.$$

Thus the part of (7) corresponding to the integral for t far from 0 is right. The other part, for $t\leqslant\epsilon$, is harmless: we bound it by $\sup\{|f(z)|^2: |z|\leqslant\epsilon\}$, that is in turn bounded by the right-hand side of (6), choosing ϵ so that $D_\alpha(\eta)$ contains the ball $|z|\leqslant 2\epsilon$. □

As a consequence of the above, we note that if

$$N=\sum_j z_j\frac{\partial}{\partial z_j}$$

is the true radial derivative, then, since $Nf = Rf - f = Rf - IRF$, we see

$$\int_{D_\alpha(\eta)} |Nf(z)|^2(1-|z|)^{2\gamma-n-1}\,dm(z) \leqslant c\int_{D_\alpha(\eta)} |Rf(z)|^2(1-|z|)^{2\gamma-n-1}\,dm(z)$$
$$\leqslant c\int_{D_\alpha(\eta)} |Nf(z)|^2(1-|z|)^{2\gamma-n-1}\,dm(z)$$
$$+ \sup\{|f(z)|: |z| \leqslant r_0\}.$$

(The first inequality is elementary and the second is a particular case of Lemma 3.12.) Thus, it does not matter to use R or N in the definitions of the area functions.

END OF PROOF OF THEOREM 3.8 (AND THEOREM 3.3). We have to show that if $T_\delta I^{k/2} f \in E$ for all $\delta \in C_k$ then $f \in E$. Applying Lemma 3.6 to $T_\delta I^{k/2} f$ (this is why we need Lemma 3.6 for functions more general than holomorphic) $X = \bar{T}_\delta$, $l = k/2$, we obtain

$$\int_{D_\alpha(\eta)} |I^{k/2}\bar{T}_\delta T_\delta I^{k/2} f(z)|^2(1-|z|)^{2\gamma-n-1}\,dm(z)$$
$$\leqslant C\int_{D_\alpha(\eta)} |\nabla_T^k I^{k/2} f(z)|^2(1-|z|)^{2\gamma-n-1}\,dm(z).$$

Adding on $\delta \in C_k$, using Lemma 3.11 and then Lemma 3.12 the proof of Theorem 3.8 is finished. □

4. Characterizations of H_s^p in Terms of Littlewood-Paley Functions and Maximal Functions

In this section we will obtain criteria for $f \in H_s^p$ is terms of g-functions and maximal functions, some of them just depending on complex-tangential derivatives.

For an operator X and $\gamma > 0$ we consider the g-function

$$g_\gamma^X f(\eta) = \left\{\int_0^1 |Xf(r\eta)|^2(1-r)^{2\gamma-1}\,dr\right\}^{1/2}.$$

In this section we will prove results corresponding to Theorems 3.2 and 3.3.

Theorem 4.1. *If $f \in H_s^p$ and X has weight $m > s$, then $g_{m-s}^X f \in L^p(S)$.*

Theorem 4.2. *The following are equivalent:*

(*a*) $f \in H_s^p$.

(*b*) $\left\{\int_0^1 |\nabla_T^k R^l f(r\eta)|^2(1-r)^{2l+k+2s-1}\,dr\right\}^{1/2} \in L^p(S)$, *for some* l, k, $l + k/2 > s$ (*and hence for all*).

(*c*) $(\nabla_T^k I^{k/2} R^s f)^+ \in L^p(S)$, $k \leqslant 2s$.
(*d*) $(\nabla_T^k I^{k/2} R^s f)^* \in L^p(S)$, $k \leqslant 2s$.
(*e*) $\int_S |\nabla_T^k I^{k/2} R^s f(r\eta)|\, d\sigma(\eta) \leqslant C$, $0 \leqslant r \leqslant 1$.

The proof of Theorem 4.1 (and part (*a*) $\Rightarrow$ (*b*) of Theorem 4.2) is a consequence of the next Lemma and Theorem 3.2.

Lemma 4.3. *Let $f \in H_q(B^n)$ for some q, let $\gamma > 0$, and let α be an aperture. Then*

$$\int_0^1 |f(r\eta)|^2 (1-r)^{2\gamma-1}\, dr \leqslant c \int_{D_\alpha(\eta)} |f(z)|^2 (1-|z|)^{2\gamma-1-n}\, dm(z).$$

PROOF. We combine Lemma 2.5 (with $p = 2$) with Lemma 3.5 to obtain

$$|f(r\eta)|^2 \leqslant \frac{C}{(1-r)^{n+1}} \int_{S_\alpha(r,\eta)} |f(z)|^2\, dm(z) = c(1-r)^{-n-1} J_r.$$

Recall that

$$S_\alpha(r,\eta) = \left\{ z \in D_\alpha(\eta) : \frac{1}{2}(1-r^2) < 1 - |z|^2 < 2(1-r^2) \right\}.$$

Hence

$$\int_0^1 |f(r\eta)|^2 (1-r)^{\gamma+n}\, dr \leqslant c \int_0^1 (1-r)^{\gamma-1} J_r\, dr.$$

and Fubini's theorem finishes the proof. □

That (*c*) implies (*d*) follows from the next Lemma applied to each $T_\delta I^{k/2} R^s f$, $\delta \in C_k$.

Lemma 4.4. *Let $u \in H_q(B^n)$ for some q. Then $\|u_\alpha^*\|_p \leqslant c\|u^+\|_p$.*

PROOF. The proof is very much alike the corresponding real variable result in [3], and depends on the Hardy-Littlewood result, Lemma 2.5. For $z \in D_\alpha(\eta)$ let P_z be a polydisc as in Lemma 3.5, with any $\beta > \alpha$. By Lemma 2.5,

$$|u(z)|^{p/2} \leqslant c(1-r)^{-n-1} \int_{P_z} |u(w)|^{p/2}\, dm(w).$$

If we use now polar coordinates, $w = r\zeta$, and bound $|u(w)|$ by $u^+(\zeta)$ we obtain

$$|u(z)|^{p/2} \leqslant c(1-r)^{-n-1} \int_{\{r\zeta \in P_z\}} |u^+(\zeta)|^{p/2}\, dr\, d\sigma(\zeta).$$

If $K(\eta, r)$ denotes the Koranyi ball $K(\eta, r) = \{\zeta \in S: |1 - \bar{\zeta}\eta| < r\}$, the projection of P_z on S is contained in $K(\eta, c(1-r))$, because $z \in D_\alpha(\eta)$, for some constant c. Hence

$$|u(z)|^{p/2} \leqslant c(1-r)^{-n} \int_{K(\eta, c(1-r))} |u^+(\zeta)|^{p/2}\, d\sigma(\zeta).$$

Since $\sigma(K(\eta, c(1-r))) \simeq (1-r)^n$, this shows that

$$|u(z)|^{p/2} \leqslant cM[(u^+)^{p/2}](\eta).$$

i.e. $(u^*)^{p/2} \leqslant cM[(u^+)^{p/2}]$, where M denotes the Hardy-Littlewood maximal operator (defined with Korányi balls). The fact that M is bounded in L^2 then finishes the proof. □

For the remaining implications we need to recall some facts from the real variable theory of Hardy spaces, which we state as follows for the reader's convenience.

Lemma 4.5. *Let u be harmonic in B^n and for $\ell \in \mathbb{R}$, $\gamma > 0$ recall the notation*

$$g_\gamma^\ell u(\eta) = \left\{ \int_0^1 |R^\ell u(r\eta)|^2 (1-r)^{2\gamma - 1}\, dr \right\}^{1/2}.$$

Then, if $\ell, m \in \mathbb{R}$, $\gamma, r > 0$ and $m - \tau = \ell - \gamma$, the conditions $g_\gamma^\ell u \in L^p(S)$ and $g_\tau^m u \in L^p(S)$ are equivalent. In case $\ell = \gamma$, they are equivalent to $u^+ \in L^p(S)$.

PROOF. The proof depends on a result like Corollary 3.7 for the area functions in the real variable theory

$$S_{\alpha,\gamma}^\ell u(\eta) = \left\{ \int_{\Gamma_\alpha(\eta)} |R^\ell u(z)|^2 (1 - |z|)^{2(\gamma - n)}\, dm(z) \right\}^{1/2}, \qquad \gamma > 0.$$

Here $\Gamma_\alpha(\eta)$ is the cone of aperture α and vertex η. Namely, for $\gamma, r > 0$ and $m - r = \ell - \gamma$, $S_{\alpha,\gamma}^\ell$ and $S_{\alpha,\tau}^m$ are pointwise equivalent up to replacement of the apertures. This is proved as in [7, pp. 214-216] (the proof of Lemma 3.6 was precisely an adaptation of this).

Now, there is a pointwise estimate

$$g_\gamma^\ell u(\eta) \leqslant c_\alpha S_{\alpha,\gamma}^\ell u(\eta)$$

proved as in Lemma 4.3. On the other hand, the same argument as in [3, Cor. 3 p. 171], but usig the Hilbert space $L^2(0,1)$ with the measure $(1-r)^\gamma\, dr$, shows that if $g_\gamma^\ell u \in L^p(S)$ then $S_{\alpha,\gamma+1}^{\ell+1} u \in L^p(S)$, which proves the first part of the Lemma. (We remark that in this case there are no pointwise estimates among the g_γ^ℓ.)

In case $\ell = \gamma$, one can assume $\ell = \gamma = 1$ and then we have the situation of the real Hardy spaces described in [3]. □

PROOF OF $(b) \Rightarrow (c)$. According to Lemma 4.5, condition (b) is equivalent to (recall that $T_1 \cdots T_k R^\ell f$ are harmonic)

$$\left\{\int_0^1 |R^q \nabla_T^k R^\ell f(r\eta)|^2 (1-r)\, dr\right\}^{1/2} \in L^p(S)$$

with $2q - 1 = 1 + 2s - 2\ell - k$, *i.e.* $q = 1 + s - \ell - k/2$. Writing

$$R^q \nabla_T^k R^\ell = R I^{1-q} \nabla_T^k R^\ell = R \nabla_T^k I^{k/2} R^s$$

(note that $q - 1 < 0$ and recall that I^ℓ commutes with the T_δ) and applying last part of Lemma 4.5 we conclude that (c) holds.

For the implication $(d) \Rightarrow (a)$ we need

Lemma 4.6. *Let u be an harmonic function in the ball and assume that $u_\alpha^* \in L^p(S)$. Then, if $\beta < \alpha$, and*

$$A_\beta u(\eta) = \left\{\int_{D_\beta(\eta)} |Ru(z)|^2 (1 - |z|)^{1-n}\, dm(z)\right\}^{1/2}$$

then $A_\beta u \in L^p(S)$.

PROOF. This is proved exactly as in the first part of [3, Thm. 8], but replacing the regions appearing as unions of cones by regions which are unions of admissible approach ones, and using Green's theorem (also as in [7]). □

We notice that the lemma has somewhat an unnatural statement, because the notion of harmonicity, a real variable one, is mixed with complex variable notions such as the one of admissible approach region. It simply says that among harmonic functions, maximal functions control area functions. The converse would not be true for a general harmonic function. The lemma is most easily understood when $p = 2$. In this case it does no matter which kind of area function, the non-tangential Su or the admissible Au is used, because as it is easily checked their L^2-norms are comparable to the L^2-norm of u if u is harmonic, and their finiteness is thus equivalent to the non-tangential maximal function being in L^2. So it is clear that $u_\alpha^* \in L^2$ is a strictly stronger condition.

PROOF OF $(d) \Rightarrow (a)$. By the lemma just proved, from (d) we conclude that

$$\left\{\int_{D_\beta(\eta)} |R \nabla_T^k I^{k/2} R^s f(z)|^2 (1 - |z|)^{1-n}\, dm(z)\right\}^{1/2}$$

belongs to $L^p(S)$. For $\delta \in C_k$, $RT_\delta I^{k/2} R^s = T_\delta R^{1+s-k/2}$ and thus we get that the condition in Theorem 3.3 holds with $\ell = 1 + s - k/2$, and hence $f \in H^p_s$.

It remains to prove the equivalence of (*c*) and (*e*). Obviously (*c*) implies (*e*). We prove the converse for $k = 1$ showing that if g is holomorphic and $T_{ij} g(r\eta)$, $0 \leqslant r \leqslant 1$ have bounded L^p norms, then their supremum is in L^p, once again as if $T_{ij} g$ were holomorphic. If $1 < p$ this is clear because $T_{ij} g$ is harmonic and for $p \leqslant 1$ it is shown as follows. We may assume that g is holomorphic in a neighbourhood of the closed ball. Let $h(\eta) = \sup \{ |T_{ij} g(r\eta)| : 0 \leqslant r \leqslant 1 \}$. Then

$$\int_0^{2\pi} |h(e^{i\theta}\eta)|^p \, d\theta \leqslant \int_0^{2\pi} \sup_r \left| \bar{\eta}_i \frac{\partial g}{\partial \eta_j}(e^{i\theta} r\eta) - \bar{\eta}_j \frac{\partial g}{\partial \eta_i}(e^{i\theta} r\eta) \right|^p d\theta.$$

By the one variable result for H^p applied to the function in the last integral, which is holomorphic in $re^{i\theta}$, this is less than some constant times the same integral but without the supremum and with $r = 1$. Now the result follows integrating in η on the sphere. The proof for $k > 1$ is similar. □

We also point out here that if $s = k/2$ where k is an integer then we have proved

$$\sup_{0<r<1} \int_S |\nabla_T^k f(r\zeta)|^p \, d\sigma(\zeta) \leqslant C \int_S |R^{k/2} f(\zeta)|^p \, d\sigma(\zeta).$$

This will be used in Section 6.

5. The Boundary Distribution of Complex Tangential Derivatives

One consequence of the results of the previous sections is the following: if k is a positive integer and if f is holomorphic in B^n with $R^{k/2} f \in H^p(B^n)$ then $T_\delta f$, $\delta \in C_k$ has admissible maximal function in $L^p(S)$. For $0 < p \leqslant 1$ we may inquire more closely about the boundary behaviour of $T_\delta f$, $\delta \in C_k$. If $0 < p \leqslant 1$, since $T_\delta f$ is harmonic it follows that $T_\delta f$ has a boundary distribution which lies in the real Hardy space, $\operatorname{Re} H^p(S)$ and hence has an atomic decomposition as a sum of $\operatorname{Re} H^p$ atoms. We have seen however that the functions $T_\delta f$ share many properties with holomorphic functions (non-isotropic mean value inequalities for example). We see here that the boundary distribution of $T_\delta f$ also has a «holomorphic nature» by showing that it also lies in the non-isotropic Hardy space H^p_{at} of Garnett and Latter [4]. We do this by showing that the mapping $K = T_\delta I^{k/2}$ which leads from $R^{k/2} f$ to $T_\delta f$ can be

realized as a non-isotropic singular integral operator. Then we apply the method of Coifman and Weiss [2] to show that it maps the non-isotropic Hardy space H^p_{at} to itself. Then we just have to check that the boundary distribution of the harmonic function $T_\delta f$ is the same as the distribution $KR^{k/2}f$ determined by the singular integral theory.

We will prove the following

Theorem 5.1. *There is a singular integral operator K which maps H^p_{at} to H^p_{at} for all $0 < p$ such that if $f \in H^p(B^n)$, then $Kf = T_\delta I^{k/2} f$ in the following sense: we regard f as an element of H^p_{at} and apply K to obtain a distribution Kf in H^p_{at}; on the other hand the harmonic function $T_\delta I^{k/2} f$ also has a boundary distribution; these distributions are the same.*

PROOF. First we show that $\nabla_T^k I^{k/2}$ can be realized as a singular integral. To see this let $T_1, \ldots, T_k$ be complex-tangential operators and take $F \in H^1(B^n)$. If we use the definition of the operator $I^{k/2}$ and the Cauchy integral formula for F we see that

$$(I^{k/2}F)(z) = \frac{1}{\Gamma(k/2)} \int_S F(\zeta) \int_0^1 \left(\log \frac{1}{t}\right)^{\beta - 1} (1 - t\langle z, \zeta\rangle)^{-n}\, dt\, d\sigma(\zeta).$$

If we now apply $T_1 \cdots T_k$ and differentiate under the integral sign we find that

$$(KF)(z) = \int F(\zeta) K(z, \zeta)\, d\sigma(\zeta),$$

where

$$K(z, \zeta) = C\psi_1(z, \zeta) \cdots \psi_k(z, \zeta) \int_0^1 \left(\log \frac{1}{t}\right)^{\frac{k}{2} - 1} (1 - t\langle z, \zeta\rangle)^{-n-k} t^k\, dt.$$

Here C is a constant and each of the factors $\psi_1, \ldots, \psi_k$ is of the form $\overline{z_k}\overline{\zeta_l} - \overline{z_l}\overline{\zeta_k}$ for some $l \neq k$. In particular we have $\psi_j(\zeta, \zeta) \equiv 0$. Now it is not hard to see that the integral factor in the expression for $K(z, \zeta)$ behaves essentially like $(1 - \langle z, \zeta\rangle)^{-n-k/2}$. However, rather than getting an asymptotic development of this integral and keeping track of the error terms it seems easier to deal with it directly.

We state now some elementary lemmas.

Lemma 5.2. *There are constants C, A such that if $|1 - \langle \zeta, \zeta_0\rangle| < \delta$ and $|1 - \langle \zeta_1, \zeta_0\rangle| \geq C\delta$, then for $0 < r < 1$, we have*

(i) $|1 - \langle r\zeta_1, \zeta\rangle| \geq \delta$, *and*
(ii) $1/A\,|1 - \langle r\zeta_1, \zeta\rangle| \leq |1 - \langle r\zeta_1, \zeta_0\rangle| \leq A|1 - \langle r\zeta_1, \zeta\rangle|$.

Lemma 5.3. *Suppose $\alpha, \beta > 0$ and $\beta < \alpha$. Then there is C such that for all $|\lambda| < 1$ we have*

$$\int_0^1 \left(\log \frac{1}{t}\right)^{\beta-1} |1 - t\lambda|^{-\alpha}\, dt \leqslant C|1 - \lambda|^{\beta-\alpha}.$$

Lemma 5.4. *If $\zeta, \zeta_0 \in S$ and $z \in \bar{B}^n$ then*

$$|\langle z, \zeta - \zeta_0\rangle| \leqslant 2|1 - \langle \zeta, \zeta_0\rangle|^{1/2}\{|1 - \langle \zeta, \zeta_0\rangle|^{1/2} + |1 - \langle z, \zeta_0\rangle|^{1/2}\}.$$

Lemma 5.5. *If $z, w \in \bar{B}^n$, then*

$$|z - w| \leqslant \sqrt{2}\,|1 - \langle z, w\rangle|^{1/2}.$$

At first we will assume that $n/(n + \frac{1}{2}) < p \leqslant 1$.

In order to apply the results of Coifman and Weiss we need to show that:

(i) $f \mapsto \int K(z, \zeta) f(\zeta)\, d\sigma(\zeta)$ is bounded on $L^2(S)$, and

(ii) There is a constant C such that if $|1 - \langle \zeta, \zeta_0\rangle| < \delta$ and $|1 - \langle \zeta_1, \zeta_0\rangle| \geqslant C\delta$ then

$$|K(\zeta_1, \zeta) - K(\zeta_1, \zeta_0)| \leqslant C \frac{|1 - \langle \zeta, \zeta_0\rangle|^{1/2}}{|1 - \langle \zeta_1, \zeta_0\rangle|^{n+1/2}}.$$

Actually, our strategy will be to let $K_r(\zeta, \eta) = K(r\zeta, \eta)$ and show that K_r satisfies (i) and (ii) with bounds that are independent of r and then take a limit.

First we discuss (i). Fix $0 < r < 1$. Note that the kernel $\overline{K_r(\eta, \zeta)}$ is holomorphic in ζ and hence we have, for $f \in L^2$, $K_r f = K_r F$, where F is the orthogonal projection of f onto $H^2(B)$. From the previous discussion we know that $(K_r F)(\eta) = (T_1 \cdots T_k I^{k/2} F)(r\eta)$. Or, $K_r F = (T_1 \cdots T_k I^{k/2} F)_r$. It follows from the results of Section 3 that $\|(T_1 \cdots T_k I^{k/2} F)^+\|_{L^2(S)} \leqslant C\|F\|_{H^2} \leqslant C\|f\|_{L^2}$, and in particular $\int |K_r F(\eta)|^2\, d\sigma(\eta) \leqslant C\|f\|_{L^2}$, where the constant C is independent of r and of f.

We now show that (ii) holds for K_r with constants independent of r. We write

$$K(z, \zeta) = \psi(z, \zeta) L(z, \zeta) \quad \text{where} \quad \psi(z, \zeta) = C \prod_{j=1}^{k} \psi_j(z, \zeta)$$

and $L(z, \zeta)$ is the integral. We have

$$\begin{aligned} K_r(\zeta_1, \zeta) - K_r(\zeta_1, \zeta_0) &= [\psi(r\zeta_1, \zeta) - \psi(r\zeta_1, \zeta_0)] L(r\zeta_1, \zeta) \\ &\quad + \psi(r\zeta_1, \zeta_0)[L(r\zeta_1, \zeta) - L(r\zeta_1, \zeta_0)] \\ &= I + II. \end{aligned}$$

Note that

$$|L(r\zeta_1, \zeta)| \leqslant C \int_0^1 \left(\log \frac{1}{t}\right)^{\frac{k}{2}-1} |1 - tr\langle \zeta_1, \zeta\rangle|^{-n-k}\, dt$$
$$\leqslant C|1 - r\langle \zeta_1, \zeta\rangle|^{-n-\frac{k}{2}}$$
$$\leqslant C|1 - \langle \zeta_1, \zeta_0\rangle|^{-n-\frac{k}{2}}$$

because of the assumptions in (ii) concerning ζ, ζ_0, ζ_1 and Lemmas 5.2 and 5.3.

Next note that

$$|\psi(r\zeta_1, \zeta) - \psi(r\zeta_1, \zeta_0)|$$
$$\leqslant C \sum_{l=0}^{k-1} \left| \prod_{j=1}^{l} \psi_j(r\zeta_1, \zeta) \prod_{i=l+2}^{k} \psi_i(r\zeta_1, \zeta_0)[\psi_{l+1}(r\zeta_1, \zeta) - \psi_{l+1}(r\zeta_1, \zeta_0)] \right|$$
$$\leqslant C|1 - r\langle \zeta_1, \zeta\rangle|^{l/2} |1 - \langle r\zeta_1, \zeta_0\rangle|^{(k-l-1)/2} |1 - \langle \zeta, \zeta_0\rangle|^{1/2}$$
$$\leqslant C|1 - r\langle \zeta_1, \zeta_0\rangle|^{(k-1)/2} |1 - \langle \zeta, \zeta_0\rangle|^{1/2}.$$

It follows that

$$|I| \leqslant C \frac{|1 - \langle \zeta, \zeta_0\rangle|^{1/2}}{|1 - \langle \zeta_1, \zeta_0\rangle|^{n+1/2}}.$$

Now we turn to *II*. Note that

$$|L(r\zeta_1, \zeta) - L(r\zeta_1, \zeta_0)|$$
$$\leqslant \int_0^1 \left(\log \frac{1}{t}\right)^{k/2-1} t^n \left| \frac{1}{(1 - rt\langle \zeta_1, \zeta\rangle)^{n+k}} - \frac{1}{(1 - rt\langle \zeta_1, \zeta_0\rangle)^{n+k}} \right| dt.$$

Now, $|(1-\lambda)^{-(n+k)} - (1-\lambda_0)^{-(n+k)}| \leqslant C|\lambda - \lambda_0|M$ where M is the maximum of $|1 - z|^{-(n+k+1)}$ where z lies on the line segment joining λ to λ_0. A point on the segment joining $rt\langle \zeta_1, \zeta_0\rangle$ to $rt\langle \zeta_1, \zeta\rangle$ is of the form $rt\langle \zeta_1, \zeta'\rangle$ where $|1 - \langle \zeta', \zeta_0\rangle| < \delta$. It follows from Lemma 5.2 that $|1 - rt\langle z, \zeta'\rangle|$ is essentially equal to $|1 - rt\langle \zeta_1, \zeta_0\rangle|$. It follows that

$$|L(r\zeta_1, \zeta) - L(r\zeta_1, \zeta_0)|$$
$$\leqslant C|\langle \zeta_1, \zeta - \zeta_0\rangle| \int_0^1 \left(\log \frac{1}{t}\right)^{\frac{k}{2}-1} |1 - rt\langle \zeta_1, \zeta_0\rangle|^{-(n+k+1)}\, dt$$
$$\leqslant C \frac{|1 - \langle \zeta, \zeta_0\rangle|^{1/2}}{|1 - r\langle \zeta_1, \zeta_0\rangle|^{n+k/2+1/2}}.$$

We also have $|\psi(r\zeta_1, \zeta_0)| \leqslant C|1 - r\langle \zeta_1, \zeta_0\rangle|^{k/2}$ and hence

$$II \leqslant \frac{C|1 - \langle \zeta, \zeta_0\rangle|^{1/2}}{|1 - \langle \zeta_1, \zeta_0\rangle|^{n+1/2}},$$

as required.

Now we may proceed to apply the method of Coifman and Weiss [2] and the atomic decomposition of Garnett and Latter [4]. We refer to these papers for the relevant definitions. Since

$$p > \frac{n}{n + \frac{1}{2}} = \frac{2}{2 + \frac{1}{n}}$$

we may find $\epsilon, 0 < \epsilon < 1/n$ so that $p > 2/(2 + \epsilon)$. Now if we let

$$\gamma = \frac{2 - p}{p(2 + \epsilon) - 2}$$

then it follows from the above estimates for K_r that

$$\left\{\int |K_r a(\zeta)|^2 \, d\sigma(\zeta)\right\}\left\{\int |K_r a(\zeta)|^2 |1 - \langle \zeta, \zeta_0\rangle|^{n(1+\epsilon)} \, d\sigma(\zeta)\right\}^{\gamma} \leqslant C$$

whenever a is a p-atom centered at ζ_0. Here the constant C is independent of both r and a. Now by a straightforward modification of the proof of Theorem C on page 594 of [2] we see that $\|K_r a\|_{H^p_{at}} \leqslant C$, where C is independent of r and a.

If follows that we may define $K_r(f) = \sum \lambda_k K_r a_k$ whenever $f = \sum \lambda_k a_k \in H^p_{at}$ and we have $\|K_r f\|_{H^p_{at}} \leqslant C\|f\|_{H^p_{at}}$. Now if $f = \sum \lambda_k a_k \in H^p_{at}$ and $r, s < 1$ we have

$$\|K_r f - K_s f\|^p_{H^p_{at}} \leqslant \sum_{k \leq N} |\lambda_k|^p \|K_r a_k - K_s a_k\|^p_{H^p_{at}} + C \sum_{k > N} |\lambda_k|^p.$$

Now for an atom a

$$\|K_r a - K_s a\|^p_{H^p_{at}} \leqslant \|K_r a - K_s a\|_1 \to 0$$

as $r, s \to 1$. So we see that for each $f \in H^p_{at}$, $Kf = \lim_{r \to 1} K_r f$ exists in H^p_{at}.

Now suppose $f \in H^p(B^n)$, and let u be the harmonic function

$$u = T_1 \cdots T_K I^{k/2} f$$

defined in B^n. We know that $u^+ \in L^p$ from Section 2. Hence the functions u_r, where $u_r(\zeta) = u(r\zeta)$ have a distribution limit U as $r \to 1$. We will identify U with the distribution Kf. First note that if f is holomorphic in a neighborhood of $\bar{B}^n$ then

$$Kf = T_1 \cdots T_k I^{k/2} f = U \quad \text{on} \quad S.$$

Also if $f \in H^p(B^n)$ then $f_r \to f$ in H^p_{at} and so $Kf_r \to Kf$ in H^p_{at}, and hence in the sense of distributions. But $Kf_r = u_r$ and this shows that

$$U = Kf$$

as desired.

We will only say a few words about the case $p \leqslant n/(n + \frac{1}{2})$. In this case the definition of H^p atoms involves some cancellation. In this case the estimate (ii) above must be replaced by an estimate of the type

$$|K_r(\eta, \zeta) - P(\eta, \zeta_0)| \leqslant C \frac{|1 - \langle \zeta, \zeta_0 \rangle|^a}{|1 - \langle \eta, \zeta_0 \rangle|^{n+a}}$$

where a is some positive number depending on p and $P(\eta, \zeta_0)$ is a certain non-isotropic Taylor polynomial expansion of K_r about ζ_0.

6. Bergman-Sobolev Spaces

In this section we give an application of the results and techniques of the previous sections to the Bergman-Sobolev spaces $B^p_{s,\gamma}$ defined in the introduction. In some ways these spaces are much easier to deal with than the spaces $H^p_s(B^n)$. The reason is that this family of spaces $(B^p_{s,\gamma})$ is stable under differentiation, a property not shared by the Hardy-Sobolev spaces. More precisely, we have $B^p_{s,\gamma} = B^p_{t,\beta}$ as long as $sp - \gamma = tp - \beta$, (and $\gamma, \beta > -1$, of course). This is immediate in the case of the unit ball B^n since the case $n = 1$ is a well-know theorem of Hardy and Littlewood and the general case follows by slice integration.

The fact that each radial derivative controls two tangential derivatives in the context of the spaces $B^p_{s,\gamma}$ has been proved in [1], even for general domains. In the theorem below we show how this follows immediately from the results of Sections 3-4. We also show that the converse holds. This does not seem to follow directly from the statements of our earlier results but it is a simple consequence of the techniques developed in the previous sections.

Theorem 6.1. *Suppose that $k/2 \leqslant s$, k is an integer; then $f \in B^p_{s,\gamma}$ if and only if*

$$\int_{B^n} |\nabla^k_T R^{s-k/2} f(z)|^p (1 - |z|)^\gamma \, dm(z) < \infty.$$

PROOF. We may assume that $k/2 = s$. One direction follows from the inequality, (see Section 4)

$$\int |\nabla_T^k f(r\zeta)|^p \, d\sigma(\zeta) \leqslant C \int |R^{k/2} f(r\zeta)|^p \, d\sigma(\zeta),$$

upon multiplying by $(1 - r)^\gamma$ and integrating on r. For the other direction, by the remarks above it suffices to prove that

$$\int |R^k f(z)|^p (1 - |z|)^{\gamma + kp/2} \, dm(z) \leqslant \int |\nabla_T^k f(z)|^p (1 - |z|)^\gamma \, dm(z).$$

It was proved in Lemma 3.11 that

$$\sum_{j=0}^{k} d_j R^{k-j} f = \sum_{\delta \in C_k} \bar{T}_\delta T_\delta f.$$

Fix $\delta \in C_k$, then from Lemma 2.5 we have

$$|\bar{T}_\delta T_\delta f(r\zeta)|^p \leqslant \frac{C}{(1 - r)^{n+1+kp/2}} \int_{P(r\zeta, \epsilon(1-r^2))} |T_\delta f(z)|^p \, dm(z)$$

where ϵ is chosen as in Lemma 3.5 so that $P(r\zeta, \epsilon(1 - r^2)) \subset S_\alpha(r, \zeta)$ for some $\alpha > 1$. So we have

$$|\bar{T}_\delta T_\delta f(\zeta)|^p \leqslant C(1 - r)^{-n-1-kp/2} \int |T_\delta f(z)|^2 \chi(\zeta, z) \, dm(z)$$

where χ is the characteristic function of $S_\alpha(r, \zeta)$. Now we integrate this inequality over S and use Fubini's theorem to obtain

$$\int_S |\bar{T}_\delta T_\delta f(r\zeta)|^p \, d\sigma(\zeta) \leqslant C(1 - r)^{-n-1-kp/2} \int |T_\delta f(z)|^2 \int_S \chi(\zeta, z) \, d\sigma(\zeta) \, dm(z).$$

Now $\int \chi(\zeta, z) \, d\sigma(\zeta) = 0$ unless $z \in L_r = \{z: (1 - r)/2 < 1 - |z| < 2(1 - r)\}$ and in any case $\int \chi \, d\sigma \leqslant C(1 - r)^n$ so we obtain

$$\int_S |\bar{T}_\delta T_\delta f(r\zeta)| \, d\sigma(\zeta) \leqslant C(1 - r)^{-1-kp/2} \int_{L_r} |T_\delta f(z)|^2 \, dm(z).$$

Now we multiply this inequality by $(1 - r)^{\gamma + kp/2}$, integrate on r and use Fubini's theorem once more. If we now add on $\delta \in C_k$ we obtain

$$\int \left| \sum d_j R^{k-j} f(z) \right|^p (1 - |z|)^{\gamma + kp/2} \, dm(z) \leqslant C \int |\nabla_T^k f(z)|^p (1 - |z|)^\gamma \, dm(z).$$

To replace the function in the integral on the left hand side by $R^k f$ above we note that for $j > 0$ we have

$$\int |R^{k-j} f(z)|^p (1 - |z|)^{\gamma + kp/2} \, dm(z) \leqslant C \int |R^k f(z)|^p (1 - |z|)^{\gamma + kp/2 + jp} \, dm(z)$$

and proceed as in the proof of Lemma 3.12. □

References

[1] Boas, H. and Straube, E. Sobolev norms of harmonic and analytic functions, preprint.
[2] Coifman, R. and Weiss, G. Extensions of Hardy spaces and their use in analysis, *Bull. Amer. Math. Soc.* **83**(1977), 569-645.
[3] Fefferman, C. and Stein E. H^p spaces of several variables, *Acta Math.* **129**, 137-193.
[4] Garnett, J. and Latter, R. The atomic decomposition for Hardy spaces in several complex variables, *Duke Math. J.* **45**(1978), 814-845.
[5] Korányi, A. and Vagi, S. Singular integrals in homogeneous spaces and some problems of classical analysis. *Ann. Scuola Normale Superiore Pisa* **25**(1971), 575-648.
[6] Stein, E. Boundary behavior of holomorphic functions of several complex variables, *Mathematics Notes*, Princeton University Press (1972).
[7] — *Singular integrals and Differentiability Properties of Functions*, Princeton University Press (1970).
[8] — Some problems in harmonic analysis, *Proceedings of symposium in pure Mathematics* **35**(1979), 3-19.

Patrick Ahern*
Department of Mathematics
University of Wisconsin
Madison WI 53706, USA

Joaquim Bruna**
Departament de Matemàtiques
Universitat Autònoma de Barcelona
08193 Bellaterra, SPAIN

* Partially supported by a grant from the NSF.
** Partially supported by a grant from the CIRIT-Generalitat de Catalunya and CICYT grant PB-85-0374.

REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.° 1, 1988

Interpolation of Banach Spaces, Differential Geometry and Differential Equations

Stephen Semmes

In recent years the study of interpolation of Banach spaces has seen some unexpected interactions with other fields. Some of these interactions originated from work of Coifman, Cwickel, Rochberg, Sagher, and Weiss [CRSW] on the extension of the complex method of Calderón to infinite families of Banach spaces. This extension is important because some relationships with the rest of mathematics are too trivial in the classical two-space case to be seen. Examples of this include the extremal problems studied by Szegö and the theorem of Masni and Weiner on factorization of matrix-valued functions. Both of these can be viewed as special cases of the interpolation method of [CRSW].

Since then, R. Rochberg has initiated the study of interpolation from the viewpoint of curvature on vector bundles. In particular, he characterized the interpolation families in [CRSW] in terms of a vanishing curvature condition.

Also, Z. Slodkowski [S1] has shown the interpolation families of [CRSW] to be examples of analytic multifunctions, and he applied [CRSW] to problems concerning polynomial hulls in $\mathbb{C}^n$.

In this paper I shall discuss some more interactions of interpolation theory with the rest of mathematics, beginning with some joint work with Coifman

[CS]. Our basic idea was to look for the methods of interpolation that had interesting PDE's arising as examples. The Hilbert space case of the interpolation method of [CRSW], for example, gives solutions for the Dirichlet problem on the disk for the equation

$$\bar{\partial}(\Omega^{-1}\partial\Omega) = 0, \tag{0.1}$$

where $\Omega(z)$ is a positive-definite matrix-valued function.

In the course of our work we found that it is very natural to view interpolation theory as the study of the geometry of the space $\mathfrak{N}(V)$ of all norms on a vector space V. It later became clear how to make some of these formal analogies precise and thereby make sense of differential geometry on $\mathfrak{N}(V)$. This has led to some new methods of intepolation, which are defined and analyzed via differential equations.

(Let me emphasize that the reader is not assumed to have any expertise in differential geometry. The situations in this paper in which differential geometry is discussed are quite concrete, and the relevant facts are reviewed.)

These topics are discused in more detail in the sections below. At the end I shall indicate other ways that interpolation theory relates to PDE, in particular to equations of Monge-Ampére type. I would like to conclude this introductory section with some philosophical comments.

One of the main conceptual points I wish to make in this paper is that there are many ways of looking at interpolation of Banach spaces. In particular, we should not tie ourselves too tightly to interpolation estimates for operators.

An example of a natural problem in the open-minded theory of interpolation of Banach spaces is the following. Suppose we are given two convex bodies B_1 and B_2 in $\mathbb{R}^n$. What is an optimal way of connecting B_1 to B_2 by a curve of convex bodies in $\mathbb{R}^n$? What should «optimal» mean?

In the case of norming bodies —*i.e.*, convex bodies that are symmetric about the origin— we could use an interpolation method to give us a curve of norms, and hence a curve of bodies. Are there natural notions of optimality in this context?

We can also ignore classical interpolation theory altogether in our search for natural notions of optimality. We could approach this physically: what happens when we try to deform B_1 into B_2 in a way that minimizes the work done, or the total strain, or something like that? We can think of slowly squeezing one body in a maximally efficient way until we get the other body.

There are undoubtedly many natural ways to make sense of «optimal» here.

We must be more adventurous in the theory of interpolation of Banach spaces.

1. Interpolation Via Perron Processes

This section is devoted to the interpolation method in $\mathbb{C}^d$ defined in [CS].

Z. Slodkowski [S2, 3] has also considered the problem of extending the complex method of interpolation to several complex variables. His point of view is quite different than ours, and comes from his earlier work on analytic multifunctions. He works in more generality than we do, but in the case we work in, our definition is equivalent to his.

Let V be a complex vector space. We require that V be finite-dimensional to avoid technicalities, but of course we want estimates that do not depend on the dimension. Let D be a domain in $\mathbb{C}^d$. Suppose we are given some norms $\| \ \|_\zeta$ on V for each $\zeta \in \partial D$. We want to extend $\| \ \|_\zeta$ to a norm-valued function $\| \ \|_z$ in $\bar{D}$ that satisfies nice properties. For example, we may want this extension to satisfy interpolation estimates for operators.

By an interpolation method we mean, roughly speaking, some method of assigning to $\| \ \|_\zeta$ an extension $\| \ \|_z$. We would like to find interpolation methods that are related to interesting PDE's.

This motivation provides one of the basic ideas of our approach. If we want to have a PDE floating around, then we want our interpolation method to be «local». In particular, the reiteration theorem should hold: if we start with $\| \ \|_\zeta$ on ∂D, apply the interpolation method to get $\| \ \|_z$ on $\bar{D}$, restrict $\| \ \|_z$ to the boundary of a subdomain D_0 of D, and then repeat the interpolation method on D_0 again, we should get $\| \ \|_z$ back again.

This localness does not work out if we extend the ideas of Calderón and [CRSW] naively to the higher-dimensional case; their approach relies too heavily on miracles of one complex variable in the unit disk. There are problems already for multiply-connected domains in the plane.

To get the desired localness feature we must therefore take a different tack.

Let $N_z(\cdot)$ be a norm function defined in some region in $\mathbb{C}^d$. By a «norm function» we mean a function that takes values in the set of all norms on V. We say that N_z is subharmonic if $N_z(f(z))$ is a subharmonic function of z for all holomorphic V-valued functions f. (It suffices to consider only affine holomorphic functions.) This property played an important role in [CRSW]. It is clearly local.

We define our interpolation method using a Perron process based on this notion of subharmonicity. Let $\| \ \|_\zeta$, $\zeta \in \partial D$, be a given family of norms on ∂D. Set

(1.1) $\mathfrak{A} = \{N_z : N_z$ is a subharmonic norm function defined on D such that

$$\limsup_{z \to \zeta} N_z(v) \leqslant \|v\|_\zeta \quad \text{for all} \quad \zeta \in \partial D,\ v \in V\Big\}.$$

We define $\| \ \|_z$ in D by

$$\|v\|_z = \sup_{N_z \in \mathcal{A}} N_z(v). \tag{1.2}$$

There are two main issues to address for this interpolation procedure: does it have nice abstract properties, and what are the examples? I discuss first some of the abstract properties.

If D is regular for the Dirichlet problem for the usual Laplacian, then it is regular for the construction above, in the sense that if $\| \ \|_\zeta$ is continuous on ∂D, then $\| \ \|_z$ is continuous on D, and coincides with $\| \ \|_\zeta$ on ∂D. (A norm function N_z is called continuous if $N_z(v)$ is continuous in z for each $v \in V$.) One fact used in the proof of this is that if $h(z)$ is subharmonic on D and if $P(v)$ is some norm on V, then $N_z = e^{h(z)}P$ is a subharmonic norm function on D. This allows us to build barriers for the above Perron process.

If T_z is a family of linear operators on V, T_z defined and continuous on $\bar{D}$ and holomorphic in z on D, then $\log \|T_z\|_{z,z}$ is subharmonic in z, where $\| \ \|_{z,z}$ denotes the operator norm with respect to $\| \ \|_z$. Let me indicate the proof of a special case of this: if $\|T_\zeta\|_{\zeta,\zeta} \leqslant 1$ when $\zeta \in \partial D$, then $\|T_z\|_{z,z} \leqslant 1$ on D. Indeed, under these hypotheses, if $N_z \in \mathcal{A}$, then $N_z(T_z(\cdot))$ also lies in $\mathcal{A}$. The desired conclusions now follow easily from the definitions.

One can prove the reiteration theorem for the above interpolation method by a straightforward chasing of definitions. Thus, if we call a norm function $\| \ \|_z$ harmonic on D if it can be recovered from its boundary values as in (1.1) and (1.2), then any norm function that is harmonic on D is also harmonic on all subdomains of D. Moreover, being harmonic is a local property: $\| \ \|_z$ is harmonic on D if for each $z \in D$ there is a neighborhood U of z on which $\| \ \|_z$ is harmonic. There is in fact a reasonable notion of a superharmonic norm function. This notion is local, and a norm function is harmonic if and only if it is both subharmonic and superharmonic.

Let us consider some examples. If D is the unit disk, then this interpolation construction agrees with the one in [CRSW]. If $D \subseteq \mathbb{C}^d$ but $V = \mathbb{C}$, so that $\|v\|_\zeta = a(\zeta)|v|$ for some positive function $a(\zeta)$, then $\|v\|_z = a(z)|v|$, with $\log a(z)$ harmonic. (Although this example is rather trivial, it is a good test case. If we tried to extend the method of [CRSW] directly to several variables, we would have problems computing even this case.)

Suppose now that we take $V = \mathbb{C}^n$ and we let $\| \ \|_\zeta$ be the $\ell^{p(\zeta)}$ norm on V, where $p: \partial D \to [1, \infty]$ is given. Then $\| \ \|_z$ is the $\ell^{p(z)}$ norm, with $1/p$ harmonic in D.

For our last example we assume that each $\| \ \|_\zeta$ is a Hilbert space norm. Let us identify V with $\mathbb{C}^n$ and let $\langle \ , \ \rangle$ denote the standard inner product there. Thus $\| \ \|_\zeta$ is given by $\|v\|_\zeta^2 = \langle \omega(\zeta)v, v\rangle$ for some positive matrix-valued function $\omega(\zeta)$ on ∂D. Then $\|v\|_z^2 = \langle \Omega(z)v, v\rangle$, where $\Omega(z)$ is a positive matrix-valued function that satisfies $\Omega|_{\partial D} = \omega$, Ω is smooth in D, and

$$\sum_{j=1}^{d} \bar{\partial}_j(\Omega^{-1}\partial_j\Omega) = 0 \quad \text{on} \quad D. \tag{1.3}$$

This, of course, generalizes (0.1).

Equation (1.3) arises in complex differential geometry and Yang-Mills theory. In particular, when $d = 2$, this is essentially a reformulation of the anti-self-dual Yang-Mills equation.

Thus our interpolation procedure gives a Perron process for (1.3), as well as reasonable notions of sub- and supersolutions for it. This is remarkable because these things usually do not make sense for systems. This also gives a precise sense in which the potential theory for (1.3) is controlled by the potential theory for the Laplacian.

Solving the Dirichlet problem for (1.3) with smooth solutions is a serious issue. We can rewrite (1.3) as

$$\Delta\Omega - \sum_{j=1}^{d} (\bar{\partial}_j\Omega)\Omega^{-1}(\partial_j\Omega) = 0.$$

Thus there are lower-order terms that are quadratic in the gradient. This is the critical case: if the growth were less than quadratic, then there are general methods for dealing with regularity issues, while on the other hand, there are well-known simple examples of second-order elliptic systems, with the leading term given by the Laplacian an lower-order terms that are quadratic in the gradient, for which regularity fails.

It turns out that one can solve the Dirichlet problem for (1.3) with, say, continuous boundary data and get smooth solutions in D. What saves us are very strong maximum principles for (1.3) that come from the Perron process. One can solve (1.3) using the continuity method, and the natural function spaces to work with involve quadratic Carleson measure conditions on $\nabla\Omega$, exactly like the usual conditions on ∇u when u is a harmonic function that is bounded or has BMO boundary values.

This is another manifestation of how the potential theory for (1.3) is controlled by the potential theory for the Laplacian: the natural notions of Carleson measures are the same in both places. Carleson measures seem to be a natural tool for treating other second-order elliptic systems with quadratic terms in the gradient as well.

Let me emphasize that although (1.3) is defined for matrix-valued functions, to define the Perron process for (1.3) we had to fatten up the target space and work with norm-valued functions. This is unavoidable; the problem is that the maximum of two Hilbert space norms is no longer a Hilbert space norm, but just a norm.

The fact that our interpolation procedure leads to a Perron process for (1.3) is interesting by itself, quite apart from its other features. This is a good example

of comments I made in the introduction, about how we should view interpolation theory broadly, without retricting ourselves unduly to interpolation estimates for operators or other standard viewpoints.

Much of what we have done in this section can be recast very naturally in terms of curvature on holomorphic vector bundles, following Rochberg [R]. I shall not discuss this. Instead, I shall discuss a different way of bringing geometry into the picture, which is described in [CS].

2. Geometry of the Space of all Banach Spaces

Let $\mathfrak{N}(V)$ denote the space of all norms on V. The idea is to view $\mathfrak{N}(V)$ as some sort of infinite dimensional manifold, and to recast interpolation theory into the study of the geometry of $\mathfrak{N}(V)$.

We would like to think of $\mathfrak{N}(V)$ as being like a Riemannian manifold, although it is not at all clear how to put a natural Riemannian structure on $\mathfrak{N}(V)$. However, interpolation theory gives $\mathfrak{N}(V)$ some geometric structure so that it behaves in many ways like a Riemannian manifold.

For example, given any two points $\| \cdot \|_0$ and $\| \cdot \|_1$ in $\mathfrak{N}(V)$, we can use the Calderón method to join them by a curve in $\mathfrak{N}(V)$: we set $\| \cdot \|_\theta = [\| \cdot \|_0, \| \cdot \|_1]_\theta$, $0 < \theta < 1$. We call such a curve a Calderón curve, and we decide to view the Calderón curves as being the geodesics of $\mathfrak{N}(V)$.

There is a natural sense in which these geodesics minimize length. Given two norms N and M on V, define

$$\delta(N, M) = \log \sup_{\substack{v \in V \\ v \neq 0}} \frac{M(V)}{N(V)}$$

and

$$d(M, N) = \max (\delta(M, N), \delta(N, M)).$$

This defines a metric on $\mathfrak{N}(V)$, a variant of the Banach-Mazur distance. Given a curve N_t in $\mathfrak{N}(V)$, $0 \leqslant t \leqslant 1$, we define its length to be

$$\sup \sum d(N_{t_j}, N_{t_{j+1}}),$$

where the supremum is taken over all partitions $\{t_j\}$ of $[0, 1]$.

It can be shown that if N_t is a Calderón curve, then its length is equal to $d(N_0, N_1)$. (This is not hard, using (2.1) below.) Unlike the situation for a Riemannian manifold, however, the converse is not true: it is easy to cook up a curve of norms M_t, $0 < t < 1$, that is not a Calderón curve but whose length is $d(M_0, M_1)$.

Calderón curves also have the following nice property. If N_t and M_t are both Calderón curves, then

(2.1) $\qquad d(N_t, M_t)$ is a convex function of t.

In the Riemannian case, where $d(\cdot, \cdot)$ is replaced by the geodesic distance, (2.1) means that the manifold has nonpositive sectional curvature.

Not only is (2.1) natural geometrically, but it is also natural from the point of view of interpolation. For example, it is also true that $\delta(N_t, M_t)$ is convex, and this contains the interpolation of operators theorem for Calderón curves. For let T be a linear operator, and assume a priori that T is invertible. If N_t is a Calderón curve, then so is $M_t(\cdot) = N_t(T(\cdot))$, and the convexity of $\delta(N_t, M_t)$ is exactly the same as the convexity of the log of the operator norm of T relative to N_t.

Let me give another example of how interpolation theory can be recast into the study of the geometry of $\mathfrak{N}(V)$. In Section 18 of [CS] a new proof of Wolff's 4-space reiteration theorem is given (due to R. Rochberg and the authors) that uses only (2.1), the fact that any two points of $\mathfrak{N}(V)$ can be joined by a Calderón curve, and elementary geometry. This argument does require substantial a priori assumptions —dim $V < \infty$ is enough— but in cases of concrete function spaces this can usually be taken care of by suitable approximation arguments. (Yves Meyer's wavelet basis can be particularly useful in this regard.)

[The proof of Wolff's theorem given in [CS] gives strong evidence that there should be a version for quasibanach spaces, but nothing rigorous has been obtained for the Calderón method.]

These ideas from [CS] on viewing interpolation theory as the study of the geometry of $\mathfrak{N}(V)$ make sense independently of the Perron process business, but there is a relationship between the two.

There is an important generalization of geodesics in Riemannian manifolds, namely, harmonic mappings. A mapping from a domain in $\mathbb{R}^n$ into a Riemannian manifold is called harmonic if it is a critical point for the energy norm, which is just the L^2-norm of its first derivatives. When $n = 1$, harmonic mappings are geodesics. In general harmonic maps are solutions of a nonlinear second-order system whose leading term is given by the Laplacian and whose lower-order terms are quadratic in the gradient.

Consider the following example. Let X_k denote the set of all $k \times k$ positive-definite matrices over $\mathbb{C}$. There is a natural Riemannian metric on X_k that is invariant under the action $\Omega \mapsto T^*\Omega T$, $T \in GL(k, \mathbb{C})$. (In fact, X_k is a negatively curved symmetric space.)

If $\Omega(x)$ is defined on some region $R \subseteq \mathbb{R}^d$ and takes values in X_k, then Ω is harmonic if and only if

$$\sum_{j=1}^{d} \partial_j(\Omega^{-1}\partial_j\Omega) = 0. \tag{2.2}$$

This equation is of course very similar to (1.3). Matrix-valued functions on $R \subseteq \mathbb{R}^d$ can be identified with matrix functions on $D = R \times i\mathbb{R}^d \subseteq \mathbb{C}^d$ that are independent of the imaginary variables. Solutions of (2.2) correspond exactly to solutions of (1.3) which are independent of the imaginary variables.

This suggests that we define harmonic maps into $\mathfrak{N}(V)$ as follows. If N_x is a norm function defined on $R \subseteq \mathbb{R}^d$, we say that N_x is a harmonic map into $\mathfrak{N}(V)$ if the corresponding norm function on $D = R \times i\mathbb{R}^d \subseteq \mathbb{C}^d$ is harmonic in the sense of our Perron process.

If we identify X_k with the subset of $\mathfrak{N}(\mathbb{C}^k)$ consisting of all Hilbert space norms, then we see that these two notions of harmonic maps of $R \subseteq \mathbb{R}^d$ into X_k —the Riemannian and interpolation definitions— coincide. In particular, the geodesics of X_k are just the Calderón curves that lie in X_k. This corresponds to $d = 1$ and $R = (0, 1)$.

There is a version of (2.1) for harmonic mappings. If N_x and M_x are two harmonic maps of $R \subseteq \mathbb{R}^d$ into $\mathfrak{N}$, then

$$d(N_x, M_x) \text{ is subharmonic in } x. \tag{2.3}$$

The obvious analogue of (2.3) for Riemannian manifolds is true when the manifold has nonpositive sectional curvature and other conditions hold, *e.g.*, if the manifold is complete and simply-connected. For example, X_k has these properties for each k. Thus harmonic maps into X_k satisfy (2.3) when $d(\bullet, \bullet)$ is the Riemannian distance or when $d(\bullet, \bullet)$ is the distance on $\mathfrak{N}(V)$.

Norm functions defined in a domain in $\mathbb{C}^d$ that are harmonic in the sense of Section 1 also satisfy (2.3). Harmonic norm functions in the sense of Section 1 have a lot in common with harmonic maps into $\mathfrak{N}(V)$, and in particular, solutions of (1.3) have a lot in common with harmonic maps into X_k.

When I mentioned in Section 1 the importance of certain strong forms of the maximum principle for (1.3) for obtaining smooth solutions of the Dirichlet problem, what I had in mind was (2.3). The methods used in [CS] for treating the Dirichlet problem for (1.3) can also be used for harmonic maps into Riemannian manifolds which are, say, complete, simply connected, and have nonpositive sectional curvature. (These hypotheses can be weakened to those used in [H].) The analogue of (2.3) again plays an important role.

Incidentally, I have restricted myself in this section to harmonic maps defined on some domain in $\mathbb{R}^d$ —rather than on a general Riemannian manifold with boundary— only for convenience.

In this section I have outlined a number of formal analogies between Riemannian geometry and the geometry of $\mathfrak{N}(V)$ that comes from interpolation. In the next section I describe a way of making these formal analogies precise.

3. Differential Geometry on the Space of all Banach Spaces

There is a way to endow a manifold with geometric structure similar to that of a Riemannian manifold, but without having a Riemannian metric. This is done by specifying a connection on that manifold. The precise definition of a connection can be rather confusing, so instead to giving the definition I shall try to motivate and explain it in a more limited way that suits our context.

To do this we first look more closely at how the notion of a geodesic is determined by a Riemannian metric. Let us work in local coordinates, that is, we identify a piece of a given Riemannian manifold with an open subset U of $\mathbb{R}^n$, with the Riemannian metric given by some matrix-valued function (g_{ij}) on U. A curve $\gamma(t)$ in U is a geodesic relative to g if

$$\ddot{\gamma}_k(t) + \sum_{i,j} \Gamma_{ij}^k(\gamma(t))\dot{\gamma}_i(t)\dot{\gamma}_j(t) = 0,$$

where the dots denote derivatives in t. The functions $\Gamma_{ij}^k(x)$ can be expressed explicitely in terms of g and its first derivatives, but the specific formula is not relevant for the present purpose.

On a Riemannian manifold there is a globally-defined object called the Levi-Cevita connection, which in local coordinates is given in terms of the Γ_{ij}^k's above. This object encodes a great deal of the geometry of the Riemannian manifold without actually giving the metric. As in (3.1), the class of geodesics on the manifold is determined by the connection, and t his is true for harmonic maps into the manifold as well. The Riemann curvature tensor is also given in terms of the connection, without recourse to the metric.

We can also consider connections on a manifold that do not come from a Riemannian metric. A connection can be determined by specifying Γ_{ij}^k's in local coordinates in a coherent way. By choosing a connection we endow the manifold with a lot of the same geometric structure as in the Riemannian case —geodesics, harmonic maps into the manifold, and curvature— without actually having a metric.

It turns out that there is a natural choice of connection on $\mathfrak{N}(V)$, whose geodesics are the Calderón curves, and such that harmonic maps from a domain in $\mathbb{R}^n$ into $\mathfrak{N}(V)$ are the same as the harmonic maps defined in Section 2. This allows us to ask new questions about the geometry of $\mathfrak{N}(V)$, *e.g.*, what is its curvature?

The task of defining this connection is simplified by the existence of a natural choice of global coordinates on $\mathfrak{N}(V)$, so that choosing a connection reduces to choosing Γ_{ij}^k in that one coordinate system. Before describing how this choice is made, let me reformulate slightly the Γ_{ij}^k's in a way that is more convenient in infinite dimensions. Given a point x, $\Gamma_{ij}^k(x)$ defines a bilinear operator Γ_x: $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ given by $c = \Gamma_x(a, b)$,

$$c_k = \sum_{ij} \Gamma_{ij}^k(x) a_i b_j, \qquad a, b, c \in \mathbb{R}^n.$$

Thus in local-coordinates a connection is determined by a function Γ that takes values in bilinear operators.

To see how to choose our connection on $\mathfrak{N}(V)$, we must first find a differential equation for Calderón curves; then we choose the connection so that the Calderón curves are geodesics.

It is a beautiful fact due to Rochberg [R] that there is a PDE that characterizes Calderón curves. Let N_t, $0 < t < 1$, be a Calderón curve of norms on $V = \mathbb{C}^d$, and set $F(t, v) = N_t(v)^2$. Then N_t is a Calderón curve if and only if

$$\ddot{F} - \sum_{jk} F^{j\bar{k}} \dot{F}_j \dot{F}_{\bar{k}} = 0. \tag{3.2}$$

Here F_j, $F_{\bar{k}}$, and $F_{j\bar{k}}$ denote the various complex derivatives of F in $v \in \mathbb{C}^d$, *e.g.*

$$F_{j\bar{k}} = \frac{\partial^2}{\partial v_j \partial \bar{v}_k} F,$$

and $F^{j\bar{k}}$ denotes the inverse of $F_{j\bar{k}}$, so that

$$\sum_k F^{j\bar{k}} F_{l\bar{k}} = \delta_l^j.$$

The dots denote derivatives in t.

To be honest, Calderón curves may not satisfy (3.2) in the usual sense, because F may not be smooth enough, or $F_{j\bar{k}}$ may not be invertible. There are reasonable senses in which Calderón curves are generalized solutions of (3.2). I shall not discuss this here. The main point is that (3.2) tells us how to choose a connection on $\mathfrak{N}(V)$.

Let us first set some preliminaries. Let $C_2^2(V)$ denote the Banach space of real-valued functions on V that are C^2 except at the origin and also complex homogeneous of degree 2, *i.e.*, $h(\alpha v) = |\alpha|^2 h(v)$ for any $\alpha \in \mathbb{C}$, $h \in C_2^2(V)$. Let $\mathfrak{N}_0(V)$ denote the open subset of $C_2^2(V)$ of positive functions that are strictly convex. (For some purposes —such as making sense of (3.2)— strict plurisubharmonicity would suffice.) For the discussion below we work with $\mathfrak{N}_0(V)$ instead of $\mathfrak{N}(V)$ so that everything is well defined.

We view $\mathfrak{N}_0(V) \subseteq C_2^2(V)$ as playing the same role as $U \subseteq \mathbb{R}^n$ did before. In this context Γ is now a function on $\mathfrak{N}_0(V)$ that takes values in bilinear operators defined on $C_2^2 \times C_2^2$: Given $F \in \mathfrak{N}_0(V)$ and $A, B \in C_2^2(V)$, we define

$$\Gamma_F(A, B) = -Re \sum_{jk} F^{j\bar{k}} A_j B_{\bar{k}}. \tag{3.3}$$

This defines our connection on $\mathfrak{N}_0(V)$. By definition, the geodesics for Γ are precisely the solutions of (3.2), *i.e.*, Calderón curves.

Notice that $\Gamma(A, B)$ is symmetric in A and B. This property is natural and useful, and it is one of the defining properties of the Levi-Cevita connection on a Riemannian manifold. (In technical jargon, this symmetry means that the connection is torsion-free.) Notice also that the definition of Γ is forced on us once we decide that it should be symmetric and that its geodesics should be Calderón curves.

As I promised, the class of harmonic maps of a domain in $\mathbb{R}^n$ into $\mathfrak{N}_0(V)$ defined by Γ corresponds exactly to the definition given in Section 2.

Now that we have found the connection on $\mathfrak{N}_0(V)$ that corresponds to interpolation theory, let us consider its properties from the point of view of elementary differential geometry.

Let us start with a very simple question. What can we say about the initial-value problem for geodesics on $\mathfrak{N}_0(V)$? Given $F_0 \in \mathfrak{N}_0(V)$, $A_0 \in C_2^2(V)$, when can we find a solution $F(t, v)$, $0 < t < ?$, of (3.2) such that $F = F_0$ and $\dot{F} = A_0$ at $t = 0$?

If F_0 and A_0 are real-analytic except at $v = 0$, then we can solve the initial value problem for short time using the Cauchy-Kowalevsky theorem. I don't think one can do much better. I can prove the following: Given $F_0 \in \mathfrak{N}_0(V)$, if we can solve the initial value problem whenever $A_0(v)$ *is a quadratic form, and with reasonable bounds, then* F must be in the Gevrey class of order 2, *i.e.*, F is C^∞ (except at $v = 0$), and its n^{th} derivative grows in n no worse that $C^n(n!)^2$.

Let us now address a more interesting issue: what can we say about the curvature? The curvature associated to Γ can be expressed in terms of Γ and its first derivatives exactly as in the Riemannian case. I will not give the formula for it, but it can be found in most texts on differential geometry.

Let me describe three examples of how we can learn something from looking at the curvature.

The curvature of Γ is rather a mess. One reason is that the complex conjugations and the «Re» make the algebra more complicated. This suggests that we first consider a model situation without these problems.

Let $W = \mathbb{R}^d$, and let $\mathfrak{N}^r(W)$ denote the set of all norms on W. Let $C_2^2(W)$ and $\mathfrak{N}_0^r(W)$ be as before, but with real homogeneity replacing complex homogeneity. Define $\bar{\Gamma}$ for $\mathfrak{N}_0^r(W)$ by

$$\bar{\Gamma}_F(A, B) = -\sum_{j,k} F^{jk} A_j B_k, \tag{3.4}$$

$$F \in \mathfrak{N}_0^r(W), \qquad A, B \in C_2^2(W).$$

Thus $\bar{\Gamma}$ is exactly like Γ, except that we have removed all vestiges of the complex numbers.

The curvature of $\bar{\Gamma}$ is much simpler that that of Γ: it vanishes identically. General nonsense says that there is a change of variables under which $\bar{\Gamma}$ transforms to 0. (The way $\bar{\Gamma}$ transforms under a change of variables is a little complicated.) A good guess and some calculation show that the desired change of variables is given by $\delta(\|\cdot\|) = \|\cdot\|^*$, *i.e.*, the map that takes a norm to its dual norm. (We view $\|\cdot\|^*$ as a norm on W, via the standard pairing on $W = \mathbb{R}^d$.)

More calculation allows one to show that Γ is preserved by this same change of variables. This contains in particular the duality theorem for the Calderón method. Still more calculation of the same type allows one to prove a duality theorem for the interpolation method described in Section 1. This result was obtained first by Slodkowski [S2, part I], and in fact he proved a much more general theorem.

A better real-variable model for Γ on $\mathfrak{N}_0^r(W)$ is Γ', defined by

$$\Gamma'_F(A, B) = -\frac{1}{2}\sum F^{jk} A_j B_k, \qquad F \in \mathfrak{N}_0^r(W), \qquad A, B \in C_2^2(W). \tag{3.5}$$

The same calculations that show that $\tilde{\Gamma}$ is trivialized by δ also give that Γ' is invariant under δ. In particular, the class of geodesics associated to Γ' is invariant under δ, just like for the Calderón method.

Let me briefly describe a computation that arises when computing how Γ, $\tilde{\Gamma}$, and Γ' transform under δ and which is interesting in its own right. To compute these transformations, we must first compute the differential of δ. Because we are working in a somewhat unusual situation, let me spell out what I mean by this. Fix $F \in \mathfrak{N}_0^r(W)$. Given $A \in C_2^2(W)$, define

$$d\delta_F(A) = \lim_{t \to 0} \frac{\delta(F + tA) - \delta(F)}{t}, \tag{3.10}$$

i.e., the directional derivative of δ at F in the direction of A.

It turns out that there is a nice formula for $d\delta_F(A)$. Let J be the duality map associated to F, that is, if N is the norm such that $N^2 = F$, then $N^*(J(v)) = N(v)$ and $\langle v, J(v)\rangle = N(v) \cdot N^*(J(v))$. Here $\langle \cdot, \cdot \rangle$ denotes the standard pairing on $\mathbb{R}^d$, which is used to define the dual norm. Our formula for $d\delta$ is

$$d\delta_F(A) = -A \circ J^{-1}. \tag{3.11}$$

The proof of (3.11) is not infinitely difficult. First we recall the well-known fact that

$$J(V) = \frac{1}{2}\nabla F(v).$$

Indeed, by definition the linear functional $w \mapsto \langle w, J(v)\rangle$ attains its maximum on the F-sphere $S = \{w: F(w) = F(v)\}$ at v, so that the hyperplane $\{w: \langle w, J(v)\rangle = \langle v, J(v)\rangle\}$ must be tangent to S at v, which implies that $J(v)$ is proportional to ∇F at v. The constant $1/2$ is determined by the relation $F(v) = F^*(J(v))$.

Next, we differentiate the relation

$$F = F^* \circ J, \tag{3.12}$$

viewing J as a function of F, and apply the chain rule. When we differentiate both sides of (3.12) in the direction of A, just as in the right side of (3.10), we get

$$A = (d\delta_F(A)) \circ J + \sum_l \left(\left(\frac{\partial}{\partial v_l} F^* \right) \circ J \right) \cdot (dJ_F(A))_l. \tag{3.13}$$

Let me explain the right side of (3.13). By definitions, $d\delta_F(A)$ is a function on W (homogeneous of degree 2), and when we compose that function with J, we get the first term on the right. The second term on the right comes from the chain rule. Because J is a map from W to W for each F, so is $dJ_F(A)$, and $(dJ_F(A))_l$ denotes the l^{th} component of this vector-valued function.

We can simplify this considerably. Because

$$J = \frac{1}{2}\nabla F, \qquad dJ_F(A) = \frac{1}{2}\nabla A.$$

Also,

$$K = \frac{1}{2}\nabla F^*$$

is just the duality map for F^*, so that $K = J^{-1}$, and thus

$$\frac{1}{2}(\nabla F^*) \circ J$$

is the identity. If we write out the last term in (3.13) as a function of v using these remarks, we get

$$\sum_l 2v_l \cdot \frac{1}{2} \frac{\partial}{\partial v_l} A(v).$$

Because A is homogeneous of degree 2, one can show that this equals $2A(v)$. Plugging this into (3.13) gives (3.11).

Once we have computed $d\delta$, it is not difficult to compute the effect of the change of variables δ on $\tilde{\Gamma}$ and Γ'. The case of Γ is more complicated, but in the same spirit.

This concludes the discussion of the first example of how we can learn something from computing the curvature of these connections. We have seen that it has led us to important formulae, which helps us compute the action of duality on Γ. It also led to an interesting real-variable model for Γ, to wit, Γ'.

The second example deals with the following question: Is there some natural sense in which $\mathfrak{N}_0(V)$ is negatively curved? Usually «negatively curved» is defined to mean that the sectional curvatures are negative, and the definition of sectional curvature involves both the curvature tensor and the Riemannian metric. In the case of $\mathfrak{N}_0(V)$, it is not clear how «negatively curved» should be interpreted.

Let me give two pieces of evidence that suggest that $\mathfrak{N}_0(V)$ is negatively curved in some reasonable sense. First, we have seen that if M_t, N_t are two Calderón curves, then $d(N_t, M_t)$ is convex. In the Riemannian case this means that the sectional curvature is nonpositive.

Second, $\mathfrak{N}_0(V)$ contains the space of all Hilbert space norms on V as a submanifold, which we identify with the manifold of all $d \times d$ positive-definite matrices. This manifold has a natural Riemannian structure with nonpositive sectional curvature, and the corresponding connection is the same as what you get by restricting Γ to it.

There are in fact natural senses in which Γ and Γ' have nonpositive curvature. Understanding how this works leads to a better knowledge of the maximum principles for geodesics and harmonic maps of Γ and Γ', and the relationship between the curvature and these maximum principles.

For the third example we start with the following question: is there a Riemannian structure on $\mathfrak{N}_0(V)$ or $\mathfrak{N}_0^r(W)$ compatible with Γ or Γ'?

Let me begin by explaining how this relates to curvature. I shall restrict the discussion to the specific example of $\mathfrak{N}_0^r(W)$, equipped with Γ'.

A Riemannian metric on $\mathfrak{N}_0^r(W)$ is a bilinear-form valued function g on $\mathfrak{N}_0^r(W)$; more precisely, for each $F \in \mathfrak{N}_0^r(W)$, $g_F(\cdot, \cdot)$ should be a positive bilinear form on $C_2^2(W)$. (I am also willing to consider g's that are only defined on subclasses of C_2^2 with more smoothness, even real-analyticity.)

There are a number of ways of thinking of the curvature tensor R associated to Γ'. I prefer to think of it as follows. For each $F \in \mathfrak{N}_0^r(W)$ and every $A, B \in C_2^2(W)$, $R_F(A, B)$ defines a linear operator on homogeneous functions of degree 2 on W. Thus for each $C \in C_2^2(W)$, $R_F(A, B)C$ is a homogeneous function of degree 2 on W. Also, $R_F(A, B)$ is linear in A and B, and $R_F(A, B) = -R_F(B, A)$.

In the case of Γ', the formula for the curvature is not so bad:

$$R_F(A, B)C = \frac{1}{4} \sum_{ijkl} \{F^{ki} B_{ij} F^{jl} A_k C_l - F^{ki} A_{ij} F^{jl} B_k C_l\}. \tag{3.14}$$

As usual, subscripts denote derivatives, and F^{ki} denotes the inverse of F_{lm}. (The curvature for Γ is more complicated, but in the same spirit.)

If Γ' is the canonical connection associated to a Riemannian metric g, then the curvature tensor must satisfy some extra conditions. For example, if $F \in \mathfrak{N}_0^r(W)$ is given, and $A, B, C, D \in C_2^2(W)$ are arbitrary, then we would have to have

$$g_F(R_F(A, B)C, D) = -g_F(C, R_F(A, B)D). \tag{3.15}$$

In other words, $R_F(A, B)$ should define an antisymmetric operator relative to $g_F(\cdot, \cdot)$.

Let us choose F to be the square of the ordinary Euclidean norm, and let us denote by g_0 and R_0 the values of g_F and R_F at this particular point. In this case (3.14) reduces to

$$R_0(A, B)C = \frac{1}{16} \sum_{jk} \{B_{jk}A_jC_k - A_{jk}B_jC_k\}. \tag{3.16}$$

Suppose now that $\dim W = 2$. We can identify homogeneous functions of degree 2 with functions on the circle. One can easily show that every tangent vector field X on the circle arises as the scum of two operators of the form $R_0(A, B)$. Hence the compatibility condition (3.15) tells us that such X must be antisymmetric with respect to $g_0(\cdot, \cdot)$. It is not hard to show that this implies that

$$g_0(A, B) = a \int_0^{2\pi} A(t)B'(t)\, dt$$

for some $a \in \mathbb{R}$, where I have now identified functions on the circle with periodic functions on $[0, 2\pi]$. Because this bilinear form is never symmetric and positive-definite, we conclude that Γ' is not compatible with any Riemannian metric when $\dim W = 2$.

It does not seem reasonable to expect that in higher dimensions the linear span of the operators $R_0(A, B)$ should give all tangent vector fields on the sphere. It seems much more reasonable to hope that the Lie algebra generated by these operators would give all such tangent vector fields. For the purposes above, it would suffice if the Lie algebra generated by $R_0(A, B)$ and covariant derivatives of the curvature operator gave everything. In any case, it is very likely still true in higher dimensions that Γ' is never compatible with a Riemannian metric, and that one can prove this using (3.15) and its variants.

The preceeding discussion can and should be phrased in terms of the holonomy group of the connection Γ'. Let me briefly indicate some of the main points.

The holonomy group of Γ' can be naturally realized as a subgroup of the group of all orientation-preserving diffeomorphisms of the unit sphere of

$W = \mathbb{R}^d$, no matter what dim W is. Calculations like those above show that when dim $W = 2$ the holonomy group is in some moral sense the group of all the orientation-preserving diffeomorphisms on the circle. To be specific, the Lie algebra of the holonomy group contains all real-analytic vector fields on the circle. I do not know whether the analogue of this last fact is true in higher-dimensions.

Although there is more to say about the behavior of Γ and Γ' from the viewpoint of differential geometry, and on the relationship of this to interpolation theory, it is probable better to move on to the next topic.

4. Interpolation Theory Via Differential Equations

Since Γ' seems to be a good real-variable model for Γ, it is natural to ask whether the geodesics for Γ' —*i.e.*, solutions $F(t, v)$ of

$$\ddot{F} - \frac{1}{2} \sum_{jk} F^{jk} \dot{F}_j \dot{F}_k = 0 \tag{4.1}$$

— share many properties with Calderón curves. For example, does (4.1) define a good interpolation method?

I think that this is a fun question. We are starting with nothing but the equation, and we want to analyze its interpolation-theoretic properties directly from the PDE.

One of the first questions is whether we can solve the boundary value problem for (4.1). We will not be able to get smooth solutions in general, but we can look for generalized solutions via a Perron process.

Let me warn the reader from the outset that in this discussion I intend to ignore all issues related to making sense of generalized solutions, and instead do the computations as though everything is smooth. This will help me to convey the essential ideas without getting bogged down in less interesting technicalities.

To define a Perron process we first need to have a good notion of suharmonicity. It can be shown that $F(t, v)$ is a subsolution of (4.1) (*i.e.*, the left side is $\geqslant 0$) if and only if for each affine W-valued function $f(t)$, we have

$$F(t, f(t))^{\cdot\cdot} \geqslant - \sum_{j,k} F_{jk}(t, f(t)) \dot{f}_j \dot{f}_k . \tag{4.2}$$

This should be compared with the notion of subharmonicity discussed in Section 1. [Incidentally, subsolutions for the geodesic equation for $\tilde{\Gamma}$ (which amounts to replacing the 1/2 in (4.1) with a 1) are characterized by the property that $F(t, f(t))^{\cdot\cdot} \geqslant 0$ for all affine functions $f(t)$. Also, $F(t, v)$ is a geodesic for $\tilde{\Gamma}$ if and only if F^* is affine. This is because $\tilde{\Gamma}$ is trivialized by the change of variables $\delta(\| \ \|) = \| \ \|^*$.]

Notice that the subharmonicity condition (4.2) is equivalent to requiring that $F(t, f(s+t))$ be a subharmonic function of (s, t) for all affine functions f. This helps make clear certain basic properties of this notion of a subharmonic norm function, for instance, that the maximum of two such things is also subharmonic.

Using this notion of subharmonicity we can define a Perron process that gives us generalized solutions of the boundary value problem for (4.1).

Just as in the case of Calderón curves, if $F(t, v)$ and $G(t, v)$ are two solutions of (4.1), then $\delta(F, G)$ and $d(F, G)$ are convex functions of t, where δ and d are as defined in Section 2. One can derive from this an interpolation of operators theorem: if $\| \ \|_t$ is a curve of norms such that $F(t, v) = \|v\|_t^2$ satisfies (4.1), and if $T: W \to W$ is linear, then $\log \|T\|_{t,t}$ is convex, where $\| \ \|_{t,t}$ denotes the operator norm of T relative to $\| \ \|_t$.

Thus (4.1) does define an interpolation method. There is a duality theorem for this method because Γ' is invariant under the duality mapping on $\mathfrak{N}^r(W)$, as we saw in Section 3.

Let us now address the issue of computing examples for this interpolation method. The Hilbert space case works out in essentially the same way as for the Calderón method. The only difference is that now the Hilbert spaces are real, and one computes with real symmetric positive-definite matrices instead of their complex cousins.

The situation is quite different for L^p spaces. If $\| \ \|_t$ is taken to be the $\ell^{p(t)}$ norm on $W = \mathbb{R}^d$, then $F(t, v) = \|v\|_t^2$ gives a solution of (4.1) if and only if $p(t)$ is constant.

This turns out to be not as bad as it may seem. If $1/p(t)$ is affine, then F gives a subsolution of (4.1) when $p(t) \geqslant 2$, and it gives a supersolution when $p(t) \leqslant 2$.

More generally, given any function $p(t)$, $1 < p(t) < \infty$, with $1/p$ affine, we can define another interpolation method using the equation

$$\ddot{F} - \frac{p-1}{p} \sum F^{jk} \dot{F}_j \dot{F}_k + \left(\frac{\dot{p}}{p}\right)^2 F = 0. \tag{4.3}$$

Here F corresponds to a norm function $\| \ \|_t$ by $F(t, v) = \|v\|_t^{p(t)}$. As with (4.1), we can find generalized solutions for the boundary-value problem for (4.3) using a Perron process, based on a notion of subharmonicity that is similar to (4.2), but more complicated. There is an interpolation of operators theorem for (4.3), as well as a duality theorem: the dual of a solution of (4.3) is again a solution of (4.3), but with p replaced by its conjugate exponent.

If $1/q(t)$ is affine, then the $\ell^{q(t)}$ norm yields a solution of (4.3) if $q(t) = p(t)$, it gives a subsolution when $q(t) \geqslant p(t)$, and it gives a supersolution when $q(t) \leqslant p(t)$.

Using these interpolation methods we recover the theorem of Marcel Riesz, which states that if $1/p(t)$ and $1/q(t)$ are affine, $p(t) \leqslant q(t)$, and if T is a linear operator, then the log of the L^p-L^q operator norm of T is convex. Here L^p denotes the L^p-space of real functions. I should emphasize that for real L^p-spaces this theorem is false if we do not assume that $p \leqslant q$.

These new interpolation methods seem to be natural versions of the complex method for real Banach spaces. Although the method of analysis of (4.1) and (4.3) is based on differential equations, and hence is quite different from the usual approaches for the complex method, a very similar analysis can be used for the complex method.

These new methods hold open the alluring prospect of being able to build interpolation methods to suit your needs, by writting down a differential equation.

Let me mention one other property of the interpolation methods defined by (4.3), concerning the relationship between these methods and the classical real method. Suppose we are given two norms A and B on W, and let $[A, B]_{\theta, q}$ denote the norms obtained from A and B by the real method. Let $p(t)$ be a function such that $1/p(t)$ is affine. Then given any $\epsilon > 0$ there is a constant $C < \infty$ and a norm function N_t defined for $\epsilon \leqslant t \leqslant 1 - \epsilon$, such that $F(t, v) = N_t(v)^{p(t)}$ is a generalized solution of (4.3), and N_t is equivalent to $[A, B]_{t, p(t)}$ on $\epsilon \leqslant t \leqslant 1 - \epsilon$, with constant C. (Of course, C does not depend on dim W.)

In other words, the classical real method generates interpolation families for (4.3).

Let me sketch the proof of this. Let $K(t, v)$ denote the K-functional associated to the couple (A, B). Then the $[A, B]_{t, p(t)}$ norm is equivalent

$$H(t, v) = \left(\sum_j (2^{-jt} K(2^j, V))^{p(t)} \right)^{1/p(t)}.$$

I claim that $H(t, v)^{p(t)}$ is a subsolution of (4.3). Indeed, for any sequence of norms $\| \ \|_j$,

$$G(t, v) = \sum (2^{-jt} \| v \|_j)^{p(t)} \tag{4.4}$$

is a subsolution of (4.3). This is because a sum of subsolutions is a subsolution, and one can simply compute that each of the individual terms is a subsolution.

Thus the K-functional definition of $[A, B]_{t, p(t)}$ gives a subsolution for (4.3). One can show that the J-functional definition gives a supersolution in much the same way (or you could use duality). From here it is not difficult to get N_t.

From the point of view of nonlinear PDE, this is quite remarkable: we are getting some sort of superposition principle for (4.3).

Note that although $G(t, v)$ as in (4.4) always gives a subsolution, in general it need not be anywhere near a supersolution. In this particular case, where all the $\| \ \|_j$'s come from the K-functional as above, what we get is also equivalent to a supersolution.

Before leaving this section I should mention that we can also consider higher-dimensional versions of the interpolation methods defined by (4.1) and (4.3), just as in Section 1.

5. A Few Last Comments

It seems natural at this stage to reiterate my opinion (indicated in the introduction) that we should look for new ways of viewing interpolation theory and new connections with other fields. To give an example of this I shall discuss some observations of Coifman and myself.

Much of what we did in Section 3 still makes sense if we work with spaces of convex functions instead of norms. For example, if we let $H(W)$ denote the space of strictly convex C^2 functions on $W = \mathbb{R}^d$, then we can still define connections $\tilde{\Gamma}$ and Γ' on $H(W)$ exactly as before, and we can define geodesics associated to these connections, just as before. Similarly, we can also define Γ on $H(V)$, $V = \mathbb{C}^d$.

The equation for geodesics for $\tilde{\Gamma}$ is

$$\ddot{F} - \sum F^{jk}\dot{F}_j\dot{F}_k = 0, \qquad F = F(t, v), \qquad v \in W. \tag{5.1}$$

A little computation shows that (5.1) is equivalent to the real Monge-Ampére equation, *i.e.*, $\det(\operatorname{Hess} F) = 0$, where Hess F denotes the Hessian of F, that is, the matrix of second derivatives of F in all variables (t and v).

In the context of the Monge-Ampére equation it is natural not to require our functions to be defined on all of W, and even to allow their domains to change with t. Although this is a bit awkward, the formalism with connections and such still makes sense, at least locally.

Even in this larger setting, the curvature of $\tilde{\Gamma}$ vanishes. There is also a nice change of variables that trivializes $\tilde{\Gamma}$, analogous to duality in the context of norms. Given a convex function F defined on some domain in W, its conjugate F^* is defined by

$$F^*(w) = \sup_{v \in W} \{\langle v, w\rangle - F(v)\}. \tag{5.2}$$

Here $\langle\ ,\ \rangle$ is the standard pairing on $W = \mathbb{R}^d$. To be precise, the domain of F^* is defined to be those $w \in W$ such that the above supremum is finite. The change of variables $F \mapsto F^*$ can be shown to trivialize $\tilde{\Gamma}$, just as the duality function did in the norm case.

The preceeding contains in particular the following well-known fact. If u is a solution of Monge-Ampére in a domain in $\mathbb{R}^n$, with Rank (Hess u) $= n - 1$, then u determines a foliation of its domain by lines, with Du constant on each line. Note that this is a fact about individual solutions of Monge-Ampére, while the above stuff about trivializing $\tilde{\Gamma}$ is a property of the Monge-Ampére operator.

[To be honest, to get this business about foliations by lines in full generality, we have to have a version of the conjugate function F^* when F is not convex but satisfies det (Hess F) $\neq 0$. This can be done, but only locally, which is good enough for the foliation story.]

Calderón curves, and more generally interpolation families in the sense of [CRSW], give rise to solutions of the complex Monge-Ampére equation. If $\|\cdot\|_z$ is a family of norms on the unit disk, and $F(z, v) = \|v\|_z^2$, then $\|\cdot\|_z$ is an interpolation family in the sense of [CRSW] if and only if

$$F_{z\bar{z}} - \sum F^{j\bar{k}} F_{jz} F_{\bar{k}\bar{z}} = 0. \tag{5.3}$$

Here the z and $\bar{z}$ subscripts denote $\partial/\partial z$ and $\partial/\partial \bar{z}$ derivatives.

This characterization is due to Rochberg [R]. (He does not write it in quite this form, however.) Equation (3.2) is a special case of this, corresponding to the case where $F(z, v)$ is defined on the strip, instead of the disk, and where $F(z, v)$ does not depend on Im (y).

A little calculation shows that (5.3) is equivalent to requiring that $F(z, v)$ satisfies the complex Monge-Ampére equation, as a function of both z and v. The complex Monge-Ampére equation is defined by

$$\det\left(\frac{\partial^2}{\partial \zeta_i \partial \bar{\zeta}_j} u\right) = 0, \tag{5.4}$$

for $u(\zeta)$ defined in some domain in $\mathbb{C}^n$.

The relationship between complex Monge-Ampére and the interpolation construction of [CRSW] is more fundamental than one sees at first glance. Let me give some examples of this. If u satisfies (5.4) and also the rank of the complex Hessian $\partial_i \bar{\partial}_j u$ is $n - 1$, then it is well known that the domain of u is foliated by a family of Riemann surfaces, such that u restricts to a harmonic function on each of these surfaces. This should be compared with the foliation by lines in the real case. In the case where u is given by $\|v\|_z^2$, these Riemann surfaces are just the graphs of the extremal functions given in [CRSW]. (Recall that these extremal fuctions are the holomorphic V-valued functions $f(z)$ such that $\|f(z)\|_z$ in constant.)

L. Lempert [L1, 2] has solved certain boundary-value problems for the homogeneous complex Monge-Ampére equation on domains in $\mathbb{C}^n$ using a set-up that is strikingly similar to the interpolation method in [CRSW]. Lempert

builds his solutions by first solving extremal problems for holomorphic functions defined on the unit disc and taking values in the given domain. These extremal problems are much like those in [CRSW]. The way that the solution of Monge-Ampére is built out of the solutions of these extremal problems is also very similar to the interpolation construction in [CRSW]. Convexity assumptions on the given domain (in $\mathbb{C}^n$) play an important role in [L1, 2], just as convexity is important in [CRSW] (in obtaining the reiteration theorem, for example).

It is not hard to see why extremal problems like those in [CRSW] and [L1, 2] should be related to finding solutions of the homogeneous complex Monge-Ampére equation. As I have mentioned, such a solution often has associated to it a foliation of Riemann surfaces on which the solution is harmonic. If you can find these Riemann surfaces, and if their boundaries lie inside the boundary of the domain on which you are working, then you can obtain the solution from its boundary values by solving the Dirichlet problem for harmonic functions on these Riemann surfaces. In [L1, 2] and [CRSW], these Riemann surfaces are analytic disks obtained by solving an extremal problem.

These Riemann surfaces will always be associated to extremal problems. If u is a plurisubharmonic solution to (5.4), then its restriction to any Riemann surface is subharmonic, and the ones in the foliation are exactly those for which the restriction is harmonic. If these good Riemann surfaces have their boundaries contained in the boundary of the domain you are working on, then the preceeding remark can be used to write down an extremal problem in terms of the boundary values of u for which the good Riemann surfaces are the solutions.

I should emphasize that the details of the analysis in [L] and [CRSW] are quite different, despite the similarities in the set-up. The techniques Lempert introduces are very interesting and should prove useful in interpolation theory.

The preceeding discussion gives another example of how interpolation theory interacts with PDE in nontraditional ways.

Many of the questions that we asked at the end of the introduction in the case of norms also make sense in the more general context indicated at the beginning of the section. For example, is there a reasonable sense in which the geodesics for $\Gamma, \tilde{\Gamma}$, or Γ' give «optimal» curves of convex functions joining a given pair of convex functions? What are natural notions of «optimality» in this context, not necessarily related to these connections? Are there other natural connections on the space of convex functions?

In the norm case we saw that Γ' has more in common with Γ than $\tilde{\Gamma}$ does. In what ways is this manifested in this more general setting? Perhaps the geodesic equation for Γ' is interesting in the Monge-Ampére scheme of things.

References

[CRSW1] Coifman, R., Cwickel, M., Rochberg, R., Sagher, Y. and Weiss, G. The complex method for interpolation of operators acting on families of Banach spaces, in «Euclidean Harmonic Analysis», edited by J. Benedetto, *Lecture Notes in Math.* **779**(1980), Springer-Verlag.

[CRSW2] Coifman, R., Cwickel, M., Rochberg, R., Sagher, Y. and Weiss, G. A theory of complex interpolation for families of Banach spaces, *Adv. in Math.* **33**(1982), 203-229.

[CS] Coifman, R. and Semmes, S. Interpolation of Banach spaces, Perron processes, and Yang-Mills, to appear, in *Amer. J. Math.*

[H] Hildebrandt, S. Harmonic mappings of Riemannian manifolds, in «Harmonic Mappings and Minimal Immersions», edited by E. Guisti, *Lecture Notes in Math.* **1161**(1986).

[L1] Lempert, L. La métrique de Kobayashi et la représentation des domaines sur la boule, *Bull. Soc. Math. France* **109**(1981), 427-474.

[L2] —, Solving the degenerate complex Monge-Ampére equation with one concentrated singularity, *Math. Ann.* **263**(1983), 515-532.

[R] Rochberg, R. Interpolation of Banach spaces and negatively curved vector bundles, *Pac. J. Math.* **110**(1984), 355-376.

[Sl1] Slodkowski, Z. Polynomial hulls with convex sections and interpolation, *Proc. Amer. Math. Soc.* **96**(1986), 255-260.

[Sl2] Slodkowski, Z. Complex interpolation of normed and quasinormed spaces in several dimensions I, to appear in *Trans. Amer. Math. Soc.*; II Properties of harmonic interpolation, preprint; III Regularity results for harmonic inteprolation, preprint.

[Sl3] Slodkowski, Z. Pseudoconvex classes of functions, I, Pseudoconcave and pseudoconvex sets, to appear in *Pacific J. Math.*; II Affine pseudoconvex classes on $\mathbb{R}^n$, preprint; III Characterization of dual pseudoconvex classes on complex homogeneous spaces, to appear in *Trans. Amer. Math. Soc.*

Stephen Semmes
Department of Mathematics
Rice University
Houston, TX 77251, U.S.A.

REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.° 1, 1988

L^p Estimates for Degenerate Elliptic Equations

Antonio Sánchez-Calle

Dedicated to the memory of J. L. Rubio de Francia

Introduction

In this note we are going to address the question of when is a second order differential operator «controlled» by a subelliptic second order differential operator.

By a second order subelliptic operator on a compact manifold M we mean an operator L that in local coordinates is of the form

$$L = \Sigma a^{ij}(x)\partial x_i\,\partial x_j + \Sigma b^j(x)\,\partial x_j + c(x),$$

with a^{ij}, b^j, c all real and C^∞, $(a^{ij})^T = (a^{ij})$ positive semidefinite, and L satisfies a subelliptic estimate: for some $\epsilon > 0$

$$\|u\|_{H^\epsilon} \leqslant c(\|Lu\|_2 + \|u\|_2), \qquad u \in C^\infty(M)$$

where H^s denote the classical Sobolev spaces and $\|\ \|_p$ stands for the norm in $L^p = L^p(M, d\mu)$ for some smooth positive measure μ on M (fixed from now on).

Let now P be a second order differential operator on M. We want to know under which conditions an a priori inequality

$$(*) \qquad \|Pu\|_p \leqslant C_p(\|Lu\|_p + \|u\|_p), \qquad u \in C^\infty(M), \qquad 1 < p < \infty$$

holds.

If $P = \Sigma b^{ij}(x)\,\partial x_i\,\partial x_j + \cdots$ in local coordinates, then testing with $u(x) = e^{itx\cdot\xi}\phi(x)$ with $\phi \in C_0^\infty$ and letting $t \to +\infty$, we see that if the inequality $(*)$ holds for some p, $1 < p < \infty$, then

$$|\Sigma b^{ij}(x)\xi_i\xi_j| \leqslant C\,\Sigma a^{ij}(x)\xi_i\xi_j,$$

i.e. the principal symbol for P is bounded by a constant times the principal symbol of L.

We will show in this article that if P is selfadjoint (with respect to μ), then this condition is also sufficient.

We will also state a more technical necessary and sufficient condition for first order operators, much on the flavor of the results of Fefferman an Phong [1], where some sufficient conditions for the L^2 estimates are given.

For prior results see [3, 4].

2. Background

We are going to review some facts about subelliptic operators that will be needed, specially the results of [3] on which this paper relies.

Assume first that L is self-adjoint, $L = \Sigma a^{ij}(x)\,\partial x_i\,\partial x_j + \cdots$ in local coordinates. We say a tangent vector at x, $\Sigma\alpha^i\,\partial x_i$, is subunit if $(\Sigma\alpha^i\xi_i)^2 \leqslant \Sigma a^{ij}(x)\xi_i\xi_j$ for all ξ.

With this, one can associate a distance d to L whose corresponding balls are given by [2]:

$$B_L(x;\lambda) = \{y \in M\colon \text{there exists } \phi\colon [0,\lambda] \to M \text{ Lipschitz with } \phi(0) = x,\ \phi(\lambda) = y \text{ and } \phi'(t) \text{ subunit at } \phi(t) \text{ for a.e. } t\}.$$

Observe that $B_L(x;\lambda) = B_{\lambda^2 L}(x;1)$.

In the case of a general L, write $L = L^{sa} + X_0$, with L^{sa} selfadjoint and X_0 a vector field we set

$$B_L(x;\lambda) = B_{\lambda^2 L^{sa} + \lambda^4 X_0^* X_0}(x;1) \qquad (= B_{\lambda^2 L}(x;1)).$$

(*a*) *Standard balls.* We are going to recall the construction of the «standard» balls $B_\lambda(x) = B_\lambda^L(x)$ as in [1], [3]. These balls are equivalent to the metric balls in the sense that

$$B_L(x;c\lambda) \subseteq B_\lambda^L(x) \subseteq B_L(x;C\lambda).$$

Assume for simplicity that $x = 0$ and L is selfadjoint. Let $L = \Sigma a^{ij}(x)\,\partial x_i\,\partial x_j + \cdots$ in local coordinates and consider

$$\lambda^2 L + \lambda^{2N}\Delta = \lambda^2\Sigma a_\lambda^{ij}(x)\,\partial x_i\,\partial x_j + \cdots,$$

where Δ is the Laplacian $\partial_{x_1}^2 + \cdots + \partial_{x_n}^2$ and $N >> 1/\epsilon$ (as in [2] the $\lambda^{2N}\Delta$ term is introduced for technical reasons, mainly to assume we are dealing with polynomials when convenient).

From all the cubes Q_δ with center 0 and side $\delta = 2^{-k}$ take the biggest one for which

$$\max_i \max_{Q_\delta} \lambda^2 a_\lambda^{ii}(x) \geqslant K\delta^2$$

where K is a constant much bigger than a bound on the a^{ij}'s and their derivatives up to order 2. It is no restriction to assume the maximum is reached for $i = 1$.

Since $Q_{2\delta}$ was not chosen, $\lambda^2 a_\lambda^{ii}(x) \leqslant 4K\delta^2$ for $x \in Q_{2\delta}$, *i.e.* $a_\lambda^{ii}(x) \leqslant 4K(\delta/\lambda)^2$ in $Q_{2\delta}$. Since $|\partial^\alpha a_\lambda^{ij}| << K$ if $|\alpha| \leqslant 2$ and $a_\lambda^{11} \geqslant 0$ then

$$\begin{aligned} a_\lambda^{11}(x) &\geqslant c(\delta/\lambda)^2 \quad \text{in} \quad Q_{\delta/\lambda} \\ a_\lambda^{ii}(x) &\leqslant 10K(\delta/\lambda)^2 \quad \text{in} \quad Q_{\delta/\lambda} \\ |a_\lambda^{ij}(x)| \leqslant [a_\lambda^{ij}(x) a_\lambda^{ii}(x)]^{1/2} &\leqslant 10K(\delta/\lambda)^2 \quad \text{in} \quad Q_{\delta/\lambda}. \end{aligned}$$

Also $|\partial^\alpha a_\lambda^{ij}(x)| \leqslant C_\alpha(\delta/\lambda)^{2-|\alpha|}$ in $Q_{\delta/\lambda}$ since that is clearly true for $|\alpha| = 0$, $|\alpha| \geqslant 2$ and so for $|\alpha| = 1$ by interpolation. After scaling by δ/λ and a change of variables (with bounds independent of δ/λ) we can assume

$$\lambda^2 L + \lambda^{2N}\Delta \approx \lambda^2\Big(\partial_{u_1}^2 + \sum_{i,j\geq 2} r^{ij}(u_1, \bar{u})\,\partial u_i\,\partial u_j + \cdots\Big),$$

$\bar{u} = (u_2, \ldots, u_n)$. (Here $\Sigma c^{ij}\partial_i\partial_j + \cdots \approx \Sigma d^{ij}\partial_i\partial_j + \cdots$ means $s(c^{ij}) \leqslant (d^{ij}) \leqslant (1/s)(c^{ij})$ as matrices, s independent of the parameters). Also one has bounds for r^{ij} and its derivatives independent of δ, λ and $\sum_{i,j\geq 2} r^{ij}(u_1, \bar{u})\eta_i\eta_j \geqslant \lambda^{2N'}|\bar{\eta}|^2$ so a Taylor expansion around $u_1 = 0$ allows us to assume r^{ij} is a polynomial in u_1 (if $|u_1| \leqslant C\lambda$, λ small).

In these coordinates

$$B_\lambda^L(0) = (-\lambda, \lambda) \times B_\lambda^{\bar{L}}(0)$$

where

$$\bar{L} = \sum_{i,j\geq 2} \frac{1}{2\lambda} \int_{-\lambda}^{\lambda} r^{ij}(u_1, \bar{u})\,du_1\,\partial u_i\,\partial u_j + \cdots$$

The process is completed by using induction on the dimension.

Composing all the changes of variables and scaling to the unit cube gives a map $\Phi: Q_1 \to B_\lambda(0)$. This map is of the form $\Phi = \Phi_1 \circ \cdots \circ \Phi_n$, where

$$\Phi_j(u) = \left(u_1, \ldots, u_{j-1}, \frac{\delta_i}{\lambda}\phi_j(\lambda u_j, u_{j+1}, \ldots, u_n)\right)$$

with ϕ_j and its inverse ψ_j having bounds for its derivatives independent of λ. As a consequence $|\partial^\alpha \Phi(u)| \leqslant C_\alpha \lambda^{|\alpha|}$ and the Jacobian $|\Phi'|$ satisfies $|\Phi'(z)| = \mu(B_\lambda(0))g(z)$, with $c \leqslant g(z) \leqslant C$, $|\partial^\alpha g(z)| \leqslant C_\alpha \lambda^{|\alpha|}$. Also the scaling Φ is defined in a much larger cube and the estimates above hold in $Q_s = \{z: |z_i| \leqslant s\}$ with constants depending on s too, of course (see [3] for more properties).

(*b*) *Fundamental solution for L.* In [3] an approximate solution for $\boldsymbol{L}$ was constructed, that is, an operator $\boldsymbol{K}$ such that $\boldsymbol{LKu} = u + \boldsymbol{Eu} + \boldsymbol{Su}$ with $\boldsymbol{S}$ smoothing and $\boldsymbol{E}$ a singular integral with respect to $d(x, y)$ whose L^p operator norm can be made arbitrarily small (for p fixed). In particular, it is not difficult to see that to prove an estimate of the form

$$\|\boldsymbol{Pu}\|_p \leqslant C(\|\boldsymbol{Lu}\|_p + \|u\|_p)$$

it suffices to show that $\boldsymbol{PK}$ is bounded in L^p.

The operator $\boldsymbol{K}$ is of the form $\Sigma \boldsymbol{K}_j$ with

$$\boldsymbol{K}_j f(x) = \int K_j(x, y) f(y)\, d\mu(y);$$

here $K_j(x, y)$ is smooth and satisfies

(1) $\operatorname{supp} K_j \subseteq \{(x, y): d(x, y) \leqslant CR^{-j}\}$.

(2) $\int K_j(x, y)\, d\mu(y) = 0$.

(3) If $\Phi: Q_1 \to B_{R^{-j}}(x)$ is one of the scalings then

$$|\partial_\omega^\alpha K_j(\Phi(w), y)| \leqslant C\alpha \frac{R^{-2j}}{\mu(B_L(x; R^{-j}))}, \tag{3.1}$$

$$|\partial_\omega^\alpha K_j(\Phi(w), y) - \partial_\omega^\alpha K_j(\Phi(w), y')| \leqslant C\alpha \frac{R^{-j}}{\mu(B_L(x; R^{-j}))} d(y, y') \tag{3.2}$$

if $d(y, y') \leqslant cR^{-j}$.

(3.3) Similarly for $K_j(y, \Phi(w))$.

(see [3] for these properties; (3) is not stated explicitely but it follows easily from the results there).

3. L^p Bounds for the Self-Adjoint Case

Assume $\boldsymbol{P}$ is a smooth, selfadjoint second order differential operator in M. Assume also that $\boldsymbol{P}$ has no zeroth order term. In local coordinates $d\mu = h(x)\, dx$, $\boldsymbol{P} = (1/h)\Sigma \partial_i (h b^{ij} \partial_j)$ and $\boldsymbol{L} = \Sigma a^{ij}(x) \partial_i \partial_j + \cdots$, where $b(x, \xi) = \Sigma b^{ij}(x) \xi_i \xi_j$ and $a(x, \xi) = \Sigma a^{ij}(x) \xi_i \xi_j$ are the principal symbols of $\boldsymbol{P}$ and $\boldsymbol{L}$ respectively.

The basic ingredient in the proof of the estimates in the following

Lemma. *Assume* $|b(x, \xi)| \leqslant Ca(x, \xi)$ *and that* $\Phi: Q_1 \to B_\lambda(y)$, λ *small, is one of the scalings described above. Then the pullback of* $\boldsymbol{P}$ *by* Φ satisfies

$$\Phi^*(\lambda^2 \boldsymbol{P}) = \Sigma d^{ij}\partial_i\partial_j + \Sigma d^j\partial_j$$

where the d's and their derivatives have bounds independent of λ, Φ.

PROOF. Assume for simplicity of notation that $y = 0$. Under Φ, h is transformed into $h(\Phi(w))|\Phi'(w)|$, and this is of the form $\mu(B_\lambda(0))f(w)$, with $0 < c \leqslant f(w) \leqslant C$ and $|\partial_w^\alpha f| \leqslant C_\alpha$. Since the constant $\mu(B_\lambda(0))$ cancels out, we only need to check how the b^{ij}'s transform. We will do that by induction on the dimension, so assume the lemma holds in dimension $n - 1$ (the initial case of dimension 1 is done as the induction step).

Recall that in the construction of Φ we have first a change of variables $x = \delta\phi(u)/\lambda$ followed by a scalling by λ in u_1 (with bounds for ϕ, $\psi = \phi^{-1}$ and their derivatives independent of λ). The change

$$x = \frac{\delta}{\lambda}\phi(u)$$

sends $(\lambda^2 b^{ij}(x))$ to

$$(\lambda^2 \hat{b}^{ij}(u)) = \lambda^2 \frac{\lambda^2}{\delta^2}\psi'(\phi(u))\left(b^{ij}\left(\frac{\delta}{\lambda}\phi(u)\right)\psi'^t(\phi(u))\right).$$

Claim. $|\partial^\alpha \hat{b}^{ij}| \leqslant C_\alpha$.

To see this, it suffices to prove that $|\partial_x^\alpha b^{ij}| \leqslant C_\alpha(\delta/\lambda)^{2-|\alpha|}$ in $Q_{\delta/\lambda}$. Since this is clearly true for $|\alpha| \geqslant 2$, it suffices to show $|b^{ij}| \leqslant C(\delta/\lambda)^2$ (the derivatives of order one follow by an easy interpolation argument). Now

$$|\Sigma b^{ij}(x)\xi_i\xi_j| \leqslant C\,\Sigma a^{ij}(x)\xi_i\xi_j,$$

so

$$|b^{ii}(x)| \leqslant Ca^{ii}(x) \quad \text{and} \quad |b^{ij}(x)| \leqslant C(a^{ii}(x) + |a^{ij}(x)| + a^{jj}(x)).$$

Since by construction $|a^{ij}(x)| \leqslant C(\delta/\lambda)^2$ in $Q_{\delta/\lambda}$, that proves the claim.

Now, in the u-coordinates

$$\left|\sum_{i,j\geq 2} \hat{b}^{ij}(u_1, \bar{u})\eta_i\eta_j\right| \leqslant C \sum_{i,j\geq 2} r^{ij}(u_1, \bar{u})\eta_i\eta_j$$

$$\leqslant C \sum_{i,j\geq 2} \frac{1}{2\lambda}\int_{-\lambda}^{\lambda} r^{ij}(u_1, \bar{u})\,du_1\eta_i\eta_j$$

(the principal symbol of $\bar{L}$) since

$$\sum_{i,j\geq 2} r^{ij}(u_1, \bar{u})\eta_i\eta_j \geqslant 0$$

is a polynomial in u_1.

If we now change variables $u_1 = \lambda w_1$, $\bar{u} = \bar{\Phi}(\bar{w})$, where $\bar{\Phi}: \bar{Q}_1 \to B_\lambda^{\bar{L}}(0)$ is the map corresponding to $\bar{L}$ and we call $(\bar{b}^{ij}(w))$ the matrix that $(\lambda^2\hat{b}^{ij}(u))$ is transformed into, we want to show that $|\partial_w^\alpha b^{ij}| \leqslant C_\alpha$.

By the induction hypothesis $|\partial_{\bar{w}}^{\bar{\alpha}}\bar{b}^{ij}| \leqslant C_\alpha$ if $i,j \geqslant 2$, so $|\partial_w^\alpha \bar{b}^{ij}| \leqslant C_\alpha$ for $i,j \geqslant 2$. Also $\bar{b}^{11}(w) = \lambda^2\hat{b}^{11}(\lambda w_1, \bar{\Phi}(\bar{w}))$, so $|\partial_w^\alpha \bar{b}^{11}| \leqslant C_\alpha$. We are left with $\bar{b}^{1j}$, $2 \leqslant j \leqslant n$. To deal with them, recall that $\bar{\Phi}$ is a composition of maps of the form

$$u \to (u_1, \ldots, u_j, (\delta_i/\lambda)\phi_j(\lambda u_{j+1}, u_{j+2}, \ldots, u_n)).$$

It is not difficult then to see that $|\partial^\alpha \bar{b}^{1j}| \leqslant C_\alpha(\lambda^{|\alpha|+2}/(\delta_1 \cdots \delta_n)) \leqslant C_\alpha^*$ if $|\alpha|$ is large enough ($\delta_1 \cdots \delta_n \geqslant c\lambda^{n+1/\epsilon}$ as a consequence of [2]). Again it suffices then to get the estimates for $|\alpha| = 0$, *i.e.* to prove the

Claim. $|\bar{b}^{1j}| \leqslant C$.

To see this, observe that if $(\bar{r}^{ij}(w))$ is the matrix for

$$\sum_{i,j\leq 2} \lambda^2 r^{ij}(\lambda w_1, \bar{u})\,\partial u_i\,\partial u_j + \cdots$$

after the change of variables $\bar{u} = \bar{\Phi}(\bar{w})$ then we know that

$$\left|\sum_{i,j=1}^{n} \bar{b}^{ij}(w)\zeta_i\zeta_j\right| \leqslant C\left(\zeta_1^2 + \sum_{i,j\leq 2} \bar{r}^{ij}(w)\zeta_i\zeta_j\right)$$

so

$$|\bar{b}^{1j}(w)| \leqslant C(1 + \bar{r}^{jj}(w) + |\bar{b}^{11}(w)| + |\bar{b}^{jj}(w)|)$$

From those terms we only have to worry about checking that $\bar{r}^{jj}(w)$ is bounded. However that is a consequence of the induction hypothesis applied to $\sum_{i,j\geq 2} r^{ij}(\lambda w_1, \bar{u})\,\partial u_i\,\partial u_j + \cdots$, since as it was mentioned

$$\left|\sum_{i,j\geq 2} r^{ij}(\lambda w_1, \bar{u})\eta_i\eta_j\right| \leqslant C \sum_{i,j\geq 2} \frac{1}{2\lambda}\int_{-\lambda}^{\lambda} r^{ij}(u_1, \bar{u})\,du_1\eta_i\eta_j.$$

Thus finishes the proof of the Lemma.

With the notation used above we can now state

Theorem. *Let L be a subelliptic second order operator, P a smooth self-adjoint second order differential operator and $1 < p < \infty$. Then the estimate*

$$\|Pu\|_p \leqslant C(\|Lu\|_p + \|u\|_p), \qquad u \in C^\infty(M)$$

holds if and only if $|b(x, \xi)| \leqslant Ca(x, \xi)$ where $b(x, \xi)$, $a(x, \xi)$ are the principal symbols of P and L respectively.

PROOF. To show that $|b(x, \xi)| \leqslant Ca(x, \xi)$ is a necessary condition, take, in local coordinates, u of the form $u(x) = e^{itx\cdot\xi}\phi(x)$, $\phi \in C_0^\infty$, real. If the estimate $\|Pu\|_p \leqslant C(\|Lu\|_p + \|u\|_p)$ holds, we have

$$t^{2p}\int |\Sigma b^{ij}(x)\xi_i\xi_j|^p\phi(x)^p h(x)\,dx \leqslant C\Big(t^{2p}\int (\Sigma a^{ij}(x)\xi_i\xi_j)^p\phi(x)^p h(x)\,dx + O(t^p)\Big)$$

for all t, so

$$\int (|b(x, \xi)|^p - Ca(x, \xi)^p)\phi^p(x)h(x)\,dx \leqslant 0$$

for all $\phi \in C_0^\infty$ and hence $|b(x, \xi)|^p \leqslant Ca(x, \xi)^p$.

To prove the converse we can clearly assume that $\boldsymbol{P}$ has no zeroth order therm and, as it was mentioned, we only need to show that $\boldsymbol{PK}$ is bounded in L^p. Now $\boldsymbol{PK} = \Sigma \boldsymbol{PK}_j$, and $\boldsymbol{PK}_j$ is given by integration against a kernel $F_j(x, y) = \boldsymbol{P}^x K_j(x, y)$.

The L^p estimate follows by classical arguments if we show that $\boldsymbol{PK}$ is a singular integral, *i.e.* if we prove that

(s.i.1) $F_j(x, y)$ is supported in $\{(x, y): d(x, y) \leqslant cR^{-j}\}$ and

$$\int F_j(x, y)\,d\mu(x) = \int F_j(x, y)\,d\mu(y) = 0.$$

(s.i.2) $$|F_j(x, y)| \leqslant c/\mu(B_L(x; R^{-j}))$$

and

$$|F_j(x, y) - F_j(x, y')| + |F_j(y, x) - F_j(y', x)| \leqslant C\frac{R^j}{\mu(B_L(x; R^{-j}))}d(y, y').$$

But (s.i.1) is a consequence of the properties (1), (2) of $K_j(x, y)$ and the fact that $\boldsymbol{P}(1) = 0$. The estimates (s.i.2) follow from the property (3) of $K_j(x, y)$ and the lemma applied to the scaling $\Phi: Q_1 \to B_{R^{-j}}(x)$.

4. First Order Operators

Consider now the case of a smooth vector field Y on M. In local coordinates $Y = \Sigma y^j(x)\,\partial x_j$ and its symbol is $y(x, \xi) = i\Sigma y^j(x)\xi_j$. If $\Phi\colon Q_1 \to B_\lambda(m)$ is a scaling map then the symbol of $\Phi^*(Y)$, the pullback of Y by Φ, is given by $y(\Phi^c(z, \eta))$, where $\Phi^c(z, \eta) = (\Phi(z), \eta(\Phi'(z))^{-1})$ is the induced map on the cotangent space. We can now state

Theorem. *Let $1 < p < \infty$. The estimate*

$$\|Yu\|_p \leqslant C(\|Lu\|_p + \|u\|_p) \qquad u \in C^\infty(M)$$

holds if and only if there is a constant C_0 such that

$$\max_{\eta \in Q_1} \max_{z \in Q_1} |y(\Phi^c(z, \eta))| \leqslant C_0 \lambda^{-2}$$

for all scalings $\Phi\colon Q_1 \to B_\lambda(m)$, λ small.

PROOF. To show that the L^p estimate holds under the symbol condition we can argue as we did for P. We need to prove then that $\Phi^*(\lambda^2 Y) = \Sigma d^j(z)\,\partial z_j$ with $|\partial^\alpha d^j| \leqslant C_\alpha$, the C_α's independent of the scaling $\Phi\colon Q_1 \to B_\lambda(m)$. Assume $m = 0$. Recalling that Φ is a composition of maps

$$u \to (u_1, \ldots, u_{k-1}, (\delta_k/\lambda)\phi_k(\lambda u_k, \ldots, u_n))$$

it is not difficult to check that

$$|\partial_z^\alpha d^j(z)| \leqslant C_\alpha \frac{\lambda^{|\alpha|+2}}{\delta_1 \cdots \delta_n}$$

so if $|\alpha|$ is large $\partial^\alpha d^j$ is bounded. Since by assumption $d^j(z)$ is bounded, it follows that $\partial^\alpha d^j$ is bounded for all α.

To prove the converse, observe that applying the L^p inequalities to $v \circ \Phi^{-1}$ we get

$$\|\tilde{Y}v\|_p \leqslant C(\|\tilde{L}v\|_p + \lambda^2\|v\|_p), \qquad v \in C_0^\infty(Q_2)$$

where $\tilde{Y}, \tilde{L}$ are the pullbacks of $\lambda^2 Y$, $\lambda^2 L$ by the scaling of $B_\lambda(m)$ and $\|\ \|_p$ now denotes L^p norm with respect to Lebesgue measure (we can do that since the Jacobian $|\Phi'|$ is of the order of magnitude of $\mu(B_\lambda(m))$ in Q_2, so we can divide by it).

Taking now $v(z) = e^{z\cdot\eta}\phi(z)$ with $\eta \in Q_1$ and $\phi \in C_0^\infty(Q_2)$ a function with $\phi \equiv 1$ on Q_1 and using the fact that the coefficients of $\tilde{Y}$ have bounds independent of Φ and λ we get

$$\int_{Q_1} |\Sigma d^j(z)\eta_j|^p \, dz \leqslant C$$

where $\tilde{Y} = \Sigma d^j(z)\partial_{z_j}$. This, in turn, implies

$$|\lambda^2 y(\Phi^c(z,\eta))| = |\Sigma d^j(z)\eta_j| \leqslant C_0 \quad \text{for} \quad z \in Q_1.$$

In fact, using that $|\partial^\alpha(\Sigma d^j(z)\eta_j)| \leqslant C_\alpha$ if $|\alpha|$ is large and that

$$\max_{Q_1} |f(z)| \leqslant C\Big(\int_{Q_1} |f(z)|^p \, dz\Big)^{1/p}$$

for polynomials of some fixed degree, we get

$$\max_{Q_1} |\Sigma d^j(z)\eta_j| \leqslant C_1\Big(\int_{Q_1} |\Sigma d^j(z)\eta_j|^p\Big)^{1/p} + C_2 \leqslant C_0.$$

This finishes the proof.

5. Final Remarks

The same results hold if L^p is replaced by the Hölder spaces

$$\Gamma^\alpha(M) = \{f \text{ continuous in } M: |f(x) - f(y)| \leqslant Cd(x,y)^\alpha\}, \qquad 0 < \alpha < 1.$$

Also, by reducing it to the case of a compact manifold, one can get similar results for bounded open sets in $\mathbb{R}^n$.

References

[1] Fefferman, C. The uncertainty principle, *Bull. Amer. Math. Soc.* **9**(1983), 129-206.
[2] Fefferman, C. L. and Phong, P. H. Subelliptic eigenvalue problems, Proceedings of the Conference on Harmonic Analysis in Honor of Antoni Zygmund, *Wadsworth Math. Series* (1981), 590-606.
[3] Fefferman, C. L. and Sánchez-Calle, A. Fundamental solutions for second order subelliptic operators, *Annals of Math.* **124**(1986), 247-272.
[4] Rothschild, L. and Stein, E. M. Hypoelliptic differential operators and nilpotent Lie groups, *Acta Math.* **137**(1977), 247-320.

Antonio Sánchez-Calle*
Department of Mathematics
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139, U.S.A.

* Research supported in part by a NSF grant.

REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.° 2, 1988

A Regularity Lemma for Functions of Several Variables

J. L. Journé

Dedicated to the memory of José Luis Rubio de Francia

We shall prove the following

Theorem. *Let F_s and F_u be two continuous transverse foliations with uniformly smooth leaves, of some manifold. If f is uniformly smooth along the leaves of F_s and F_u, then f is smooth.*

In the case where the foliations are the stable and unstable foliations of an Anosov diffeomorphism, this theorem is proved in [1] using elliptic theory and in [2] using Fourier transform. Both proofs make use of the fact that the foliations are absolutely continuous. For the applications of [3] one does not even have foliations: only two transverse submanifolds through each point with continuity of the tangent spaces, and f is assumed smooth along these submanifolds. Our proof is the only one which can handle these assumptions and moreover, is totally elementary.

In order to avoid notational complications we present the proof in dimension 2. The modifications needed in higher dimensions are very minor and will be omitted.

Let $n \in \mathbb{N}$ and $\alpha \in]0, 1[$. We shall prove the $C^{n,\alpha}$ version of the theorem, which is sufficient, and this will show that there is no loss in regularity. The problem being local we may work in some open set $U \subseteq \mathbb{R}^2$. Let f be uniformly $C^{n,\alpha}$ along the leaves of F_s and F_u. To show that f is $C^{n,\alpha}$, by a lemma of Campanato [4], it is enough to show that at each point M in U we can find a $C^{n,\alpha}$ function f_M such that $|f_M(P) - f(P)| \leqslant O(\|PM\|^{n+\alpha})$ for all P in U,

all constants being locally uniform. After a $C^{n,\alpha}$ change of variables we can assume that the leaves L_u and L_s passing through M are the coordinate axes, and that f vanishes on these two leaves by replacing $f(x,y)$ by $f(x,y) - f(x,0) - f(0,y) + f(0,0)$. In these new coordinates we shall look for f_M to be a polynomial of total degree n. Then it will have to be the Taylor polynomial of order n of f at M. This polynomial will be obtained as the limit of polynomials of degree n in each variable, determined by interpolating f on sets of points converging to M.

For $r < 1$, let $x_k = y_k = r^k$ for $k \in \mathbb{N}$. For every point P, let $L^u(P)$ and $L^s(P)$ be the two leaves of F_s and F_u containing P. Let $M_{k,l}$ be $L^u((x_k,0)) \cap L^s((0,y_l))$. Let $S(k,l) = \{(0,0)\} \cup \{(x_{k'},0): k \leqslant k' < k+n\} \cup \{(0,y_{l'}): l \leqslant l' < l+n\} \cup \{(M_{k,l}): k \leqslant k' < k+n, l \leqslant l' < l+n\}$. We want for k and l large enough, with $l = k$ or $k+1$, a polynomial of degree n in x and in y which interpolates f on $S(k,l)$. In one variable, in order to interpolate a function g by a polynomial of degree n one needs $(n+1)$ distinct reals $z_0, \ldots, z_n$. If R denotes $\sup_{0 \leqslant i \leqslant n} |z_i|$ and η denotes $\inf_{i \neq l} |z_i - z_l|$, and c_p, $0 \leqslant p \leqslant n$ are the coefficients of the interpolating polynomial, then

$$\sum_{0 \leqslant p \leqslant n} |c_p| R^p \leqslant C \sup_{0 \leqslant i \leqslant n} |g(z_i)|, \tag{1}$$

where C remains bounded if R/η remains bounded. In dimension 2 one can interpolate a function by a polynomial of degree n in each variable on the cartesian product of two sets of $(n+1)$ real numbers. However, we will need to interpolate on sets of the type $S(k,l)$, which are not products, but close enough, because the foliations are continuous. The following lemma, which is probably well known, will allow us to do so.

Lemma 1. *For each $B > 1$, there exist $\epsilon > 0$ and $C > 0$ with the following properties. There exists $\epsilon > 0$ with the following property. Let $(z_{k,l})_{0 \leqslant k \leqslant n, 0 \leqslant l \leqslant n}$ be a collection of points in $\mathbb{R}^2$ for which there exists $2n+2$ real numbers $(x_k)_{0 \leqslant k \leqslant n}$ and $(y_l)_{0 \leqslant l \leqslant n}$, such that*

$$|z_{k,l} - (x_k, y_l)| \leqslant \epsilon\eta,$$

where

$$\eta = \inf_{(k,l) \neq (k',l')} |z_{k,l} - z_{k',l'}|.$$

Let

$$R = \sup_{k,l} |z_{k,l}|.$$

and suppose $R/\eta < k$.

Then given $(n+1)^2$ *real numbers* $(b_{k,l})_{0\le k\le n,\,0\le l\le n}$, *there exists a unique polynomial*

$$p(x,y)=\sum_{\substack{0\le p\le n\\ 0\le q\le n}} c_{pq}x^p y^q$$

such that $p(z_{k,l})=b_{k,l}$. *Also*

$$\sum_{p,q}|c_{p,q}|R^{p+q}\leqslant C\sup_{(k,l)}|b_{k,l}|. \tag{2}$$

PROOF. We fix $B>1$. By iterating one-dimensional interpolation one can construct a polynomial P_1, of degree n in x and in y, such that $P_1(x_k,y_l)=b_{k,l}$ for $0\leqslant k,l\leqslant n$. Then it is easy to see that for $\epsilon>0$ small enough

$$\sup_{k,l}|b_{k,l}-P_1(z_{k,l})|\leqslant\frac{1}{2}\sup_{k,l}|b_{k,l}|.$$

Then one replaces $b_{k,l}$ by $b_{k,l}-P_1(z_{k,l})$ and constructs similarly P_2. Then

$$\begin{aligned}\sup_{k,l}|b_{k,l}-P_1(z_{k,l})-P_2(z_{z,l})|&\leqslant\frac{1}{2}\sup_{k,l}|b_{k,l}-P_1(z_{k,l})|\\ &\leqslant\frac{1}{4}\sup_{k,l}|b_{k,l}|.\end{aligned}$$

It is clear that repeating the process will yield a convergent series whose sum will be P. The inequality (2) will follow from (1) and uniqueness follows from (2). We omit the details.

The continuity of the foliations implies that $|M_{k,l}-(x_k,y_l)|=O(|(x_k,y_l)|)$. It follows that for k_0 large enough and $k\geqslant k_0$, $S(k,k)$ and $S(k,k+1)$ satisfy uniformly the assumptions of the lemma. Let $k\geqslant k_0$ be fixed and P be the polynomial of degree n in x and in y which interpolates f on $S(k,k)$ and let P' be the same for $S(k,k+1)$ and let $c_{p,q}$ and $c'_{p,q}$ denote their coefficients.

We want to show

$$|c'_{p,q}-c_{p,q}|=O(r^{k(n+\alpha-p-q)}). \tag{3}$$

By Lemma 1 it is enough to show that if $(x,y)\in S_{k,k+1}$

$$|P'(x,y)-P(x,y)|=O(r^{k(n+\alpha)}).$$

Observe that P and P' agree on $S_{k,k+1}$, except on the n points $M_{k',k+n}$, $k\leqslant k'<k+n$. On these points P' agrees with f. Hence we just need to estimate $f-P$ at those n points. Now we fix k' and look at $L^u((x_{k'},0))$,

which we parametrize by the y-coordinate. A point on $L^u((x_{k'}, 0))$ is denoted $(x_{k'}(y), y) = M_y$. Since f and P are both $C^{n,\alpha}$ along $L^u((x_{k'}, 0))$, and coincide on $(n+1)$ points whose minimum and maximum distance are of the order of r^k, it follows that if $|y| \leqslant Cr^k$

$$|(f-P)(M_y)| \leqslant Cr^{k(n+\alpha)} \left\| \frac{d^n}{dy^n}(f-P) \right\|_\alpha . \tag{4}$$

By the assumption on f,

$$\left\| \frac{d^n}{dy^n} f \right\|_\alpha \leqslant c.$$

So we must estimate

$$\left\| \frac{d^n}{dy^n} P(x_{k'}(\cdot), \cdot) \right\|_\alpha ,$$

where we are only interested in values of y such that $|y| \leqslant Cr^k$. If $x_{k'}(y)$ was constant, this would be zero. The foliations being continuous, $x_{k'}$ is close to be a constant, in that for any $\delta > 0$ and for k large enough $|x'_{k'}(y)| < \delta$ if $|y| \leqslant Cr^k$.

Lemma 2. *Let $\delta > 0$. Then for k large enough*

$$\left\| \frac{d^n}{dy^n} x_k^p(y) y^q \right\|_{C^\alpha[-Cr^k, Cr^k]} \leqslant C\delta (r^k)^{p+q-n-\alpha} \tag{5}$$

if p and q are non-zero and $p+q > n$.

This follows trivially from Leibniz' rule. Also one has to use that x_k is $C^{n,\alpha}$ and that if y and y' lie in $[-Cr^k, Cr^k]$, then

$$\sup_{y \neq y'} \frac{|x_{k'}(y) - x_{k'}(y')|}{|y-y'|^\alpha} \leqslant \delta C (r^k)^{1-\alpha}.$$

We omit the details.

Notice that because we have set f to be zero on the x and y axes, no term of the form $c_{p,0}x^p$ or $c_{0,q}y^q$ appears in P. In the case where $p+q \leqslant n$, instead of (5) we use

$$\left\| \frac{d^n}{dy^n} x_k^p(y) y^q \right\|_{C^\alpha} \leqslant C. \tag{6}$$

From (4), (5) and (6) we obtain, if $|y| \leqslant Cr^k$

$$|(P' - P)(M_{k,k'})| \leqslant Cr^{k(n+\alpha)} + \delta \sum_{p+q>n} |c_{pq}| r^{k(p+q)} + \sum_{p+q \leq n} |c_{pq}| r^{k(n+\alpha)}.$$

By Lemma 1, this implies

$$(7)\quad \sum |c'_{p,q} - c_{p,q}| (r^k)^{p+q} \leqslant C\left\{ r^{k(n+\alpha)} + \delta \sum_{p+q>n} |c_{p,q}| r^{k(p+q)} + \sum_{p+q \leq n} |c_{p,q}| r^{k(n+\alpha)} \right\}.$$

We denote by P_{2k} for all k the interpolation polynomial corresponding to $S(k, k)$ and P_{2k+1} the one corresponding to $S(k, k+1)$. Then (7) relates the coefficients of P_{2k+1} to those of P_{2k}. The same proof shows that the coefficients of P_{2k+2} are related to those of P_{2k+1} in the same way. Then, we want to show by induction on m, that if $c_{p,q}$ is the coefficient of $x^p y^q$ in P_m, and $m > m_0$,

$$(8)\qquad |c_{p,q}| \leqslant K \sum_{j=m_0}^{m-1} (r^{j/2})^{n+\alpha-p-q},$$

for K and m_0 large enough.

Assume that this is true. Then it will be true for $m+1$, by (7), if

$$(9)\quad C\left\{ r^{k(n+\alpha)} + \delta K \sum_{\substack{p+q>n \\ m_0 \leq j \leq m-1}} r^{k(p+q)} (r^{j/2})^{n+\alpha-p-q} + K \sum_{\substack{p+q \leq n \\ m_0 \leq j \leq m-1}} (r^{j/2})^{n+\alpha-p-q} r^{k(n+\alpha)} \right\} \leqslant K r^{m(n+\alpha)/2}$$

where $k = [m/2]$.

The second term is less than $C\delta K r^{m(n+\alpha)/2}$ and the third one is less than $CKr^{(n+\alpha)(m+m_0)/2}$. So by choosing m_0 large enough, δ small enough, and K large enough so that (8) is satisfied for $m = m_0 + 1$, it follows by induction that (8) is satisfied for all $m > m_0$. Then the same is true for (9), (7) and (3).

We are going to see that this implies the existence of a polynomial Q of degree n, independent of m, such that

$$(10)\qquad |Q - P_m| \leqslant Cr^{m(n+\alpha)/2} \quad \text{for} \quad |x|, |y| \leqslant Cr^{m/2}.$$

The coefficients $l_{p,q}$ of Q are obtained in the following way. If $p+q > n$, let $l_{p,q} = 0$. If $p+q \leqslant n$ let

$$l_{p,q} = \lim_{m \to +\infty} c_{p,q}.$$

From (3) we know that if $p+q>n$, $|c_{p,q}| = O(r^{m(n+\alpha-p-q)/2})$ and if $p+q \leqslant n$, $|c_{p,q} - l_{p,q}| = O(r^{m(n+\alpha-p-q)/2})$. This implies (10).

From (3), (4), (5) and (6), we know that

$$|(f - P_m)(M_y)| \leqslant Cr^{m(n+\alpha)/2} \quad \text{if} \quad |y| \leqslant Cr^k.$$

Using the triangle inequality and (10) we see that

$$|(Q - f)(x_k(y), y)| \leqslant Cr^{m(n+\alpha)/2} \quad \text{if} \quad |y| \leqslant Cr^k.$$

Therefore if $|y| \leqslant 2|x|$ and $(x, y) \in \lim \bigcup_{k>0} L^u(r^k, 0)$ then:

$$|Q(x, y) - f(x, y)| \leqslant C|(x, y)|^{n+\alpha}. \tag{11}$$

Let $\epsilon > 0$ be a fixed small number. If we replace in the preceding statement the sequence r^k by another sequence α_k such that $r^{k+1} \leqslant |\alpha_k| \leqslant r^{k-1}$ for all k and $|\alpha_k - \alpha_{k-1}| \geqslant \epsilon r^k$ then one obtains the same conclusion with a possibly different polynomial $\tilde{Q}$. However by mixing two such sequences one sees that the polynomial $\tilde{Q}$ is actually independent of the sequence and that (11) holds with the only assumption $|y| < 2|x|$. Also (11) holds with a possibly different polynomial Q' of degree n if $|x| < 2|y|$. But then $|(Q - Q')(x, y)| = O(|(x, y)|^{n+\alpha})$ if $|y|/2 < |x| < 2|y|$, so that $Q = Q'$. Therefore (11) holds with no extra assumption on (x, y). According to the lemma of Campanato already mentioned, this proves the theorem.

Remark. It is easy to see that if one of the two foliations is locally trivial, —that is, the leaves are, in some coordinates neighborhood, open subsets of parallel affine subspaces— then one can remove the hypothesis of continuity on the other foliation. Indeed if it is merely measurable, one can make its local oscillation arbitrarily small by a rescaling along the direction of the trivial foliation and this is enough to carry out the proof. It is to be noted that in this slightly different context the analytic version of the theorem is false. The counterexample suggests that analytic versions of such results require much more transverse regularity than their smooth counterparts. For example in dimension 2, one can show the analytic version of the theorem if the foliations have the following property: suppose that, in some coordinates, the leaves are the graphs of functions f_y, $y \in]-R, R[$, with $f_y(0) = y$. Then it is enough to assume that $y \to f_y$, which maps $]-R, R[$ into $H^\infty(\Omega)$, for some Ω open set of $\mathbb{C}$ containing the origin, is a lipschitzian function.

Acknowledgement

The author wishes to thank Rafael de la Llave and the referee for several comments which improved the exposition.

References

[1] Llave, R. de la, Marco, J. M. y Moriyón, R. Canonical perturbation theory of Anosov systems and regularity results for the Livsic cohomology equation, *Ann. of Math.* **123**(1986), 537-611.

[2] Hurder, S. y Katok, A. Differentiability, Rigidity and Godbillon-Vey Classes for Anosov Flows. To appear in *Pub. Mat. I.H.E.S.*

[3] Llave, R. de la. New invariant manifold theorems and applications, Preprint, 1988.

[4] Campanato, S. Proprietá di una famiglia di spazi funzionali. *Ann. Sc. Norm. Super. Pisa*, **18**(1969), 137-160.

J. L. Journé
Princeton University
Mathematics Department
Princeton NJ 08544
U.S.A.

Revista Matemática Iberoamericana
Vol. 4, n.º 2, 1988

Sur la Distribution des Fonctions Dérivées

Jean-Pierre Kahane

en souvenir de José-Luis Rubio de Francia

On considère une fonction F à valeurs réelles définie et dérivable sur un intervalle réel I, et sa dérivée $f = F'$. Un théorème de Denjoy, trop peu connu, dit que, quels que soient les nombres réels A et B, l'ensemble

$$\{x: A < f(x) < B\}$$

et soit vide, soit de mesure (de Lebesgue) positive.

Prenons $I = [0, 1]$, F dérivable sur $]0, 1[$, dérivable à droite en 0, à gauche en 1, $f = F'$ sur $]0, 1[$, $f(0) = F'(0 + 0)$, $f(1) = F'(1 - 0)$. J'appellerai f une fonction dérivée définie sur I. La distribution de f (image par f de la mesure de Lebesgue sur I) est une mesure de probabilité sur $\mathbb{R}$, qu'on désigne par μ_f. Je me propose de caractériser les mesures de probabilités sur $\mathbb{R}$ qui s'obtiennent de cette façon.

Théorème. *Pour qu'une mesure de probabilité μ sur $\mathbb{R}$ soit la distribution d'une fonction dérivée f définie sur I, il faut et il suffit que*

(*a*) *le support de μ soit connexe*
(*b*) *on ait soit*

$$\int_{\mathbb{R}} |y| \, d\mu(y) < \infty \qquad (b_1)$$

soit

$$\int_{\mathbb{R}_+} y \, d\mu(y) = \int_{\mathbb{R}_-} -y \, d\mu(y) = \infty. \qquad (b_2)$$

Démonstration. Soit f une fonction dérivée définie sur I. La condition (*a*) vient du théorème de Denjoy ci-dessus. La condition (*b*) est nécessaire pour

que l'intégrale de Denjoy de f (la «totale») soit finie. Montrons maintenant que, si μ vérifie (a) et (b), il existe f telle que $\mu = \mu_f$. Voici les étapes.

1. Soit μ une mesure de probabilité sur $\mathbb{R}$, à support connexe et borné, et de valeur moyenne nulle $\left(\int y\, d\mu(y) = 0\right)$. Elle est la distribution d'une fonction φ continue et croissante sur I, telle que $\varphi(\alpha) = 0$ pour au moins un $\alpha \in I$. Montrons qu'elle est aussi la distribution d'une fonction g continue sur I et telle que $g(0) = g(1) = 0$. Il suffit de poser

$$\begin{aligned} \varphi(-x) &= \varphi(x) && (0 \leqslant x \leqslant 1) \\ \varphi(x+2) &= \varphi(x) && (-1 \leqslant x \leqslant 1) \\ g(x) &= \varphi(a + 2x) && (0 \leqslant x \leqslant 1) \end{aligned}$$

La vérification est évidente en faisant un dessin.

2. Soit μ comme ci-dessus, et $\epsilon > 0$. Montrons que μ est la distribution d'une fonction h continue sur I, telle que $h(0) = h(1) = 0$ et

$$\sup_x \left| \int_0^x h(t)\, dt \right| < \epsilon.$$

If suffit de prolonger la fonction g ci-dessus par 1-périodicité, et de poser

$$h(x) = \chi_I(x) g(Nx)$$

pour un $N = N(\epsilon)$ assez grand.

3. Soit μ une mesure de probabilité sur $\mathbb{R}$, à support connexe, vérifiant (b_1) avec $\int y\, d\mu(y) = 0$, ou (b_2). Dans les deux cas on peut écrire

$$\mu = \sum_1^\infty 2^{-n} \mu_n,$$

où, pour chaque n, μ_n est une mesure de probabilité à support connexe borné et à valeur moyenne nulle (le choix de μ_n est très arbitraire; il s'agit seulement que, pour chaque n, la mesure

$$2^n \left(\mu - \sum_{m=1}^n 2^{-m} \mu_m \right)$$

ait son support connexe). Posons $\epsilon_n = 2^{-n}$, $I_n = [2^{-n}, 2^{-n+1}]$, considérons les fonctions h_n associées aux μ_n et ϵ_n (étape 2), prolongées par 0 hors de I, et posons

$$f(x) = \sum_1^\infty h_n(2^n(x - 2^{-n})) \qquad (x \in]0, 1[).$$

Le n-ième terme de la somme est porté par I_n et sa distribution est $2^{-n}\mu_n$.

Donc la distribution de f est μ. La fonction

$$F(x) = -\int_x^1 f(t)\,dt$$

admet f pour dérivée sur $]0, 1]$ (au point 1, cela signifie $F'(1-0) = f(1) = 0$). Elle est nulle aux points 2^{-n} $(n = 1, 2, \ldots)$, et majorée en module par $2^{-n}\epsilon_n$ sur I_n. Donc, en posant $F(0) = 0$ et $f(0) = 0$, F est dérivable à droite en 0 et $F'(0+0) = f(0) = 0$. Comme la valeur de $\int y\,d\mu(y)$, quand (b_1) a lieu, ne joue aucun rôle, cela achève la démonstration.

Ajoutons quelques remarques.

1. Le théorème de Denjoy ci-dessus date de 1916 [3]. Il a été réénoncé et redémontré par Clarkson en 1946 [4]. Il apparaît comme corollaire (p. 24) dans la grande étude faite par Zahorski en 1950 [5] (avec les références à Denjoy et à Clarkson). Il est énoncé dans le livre de Bruckner de 1978 [7], p. 87, dans le chapitre consacré aux résultats de Zahorski (je dois cette dernière référence à G. Choquet). La démonstration donnée par Denjoy me paraît la plus complète et la plus simple; elle repose sur les trois propriétés suivantes des fonctions dérivées.

- (α) La propriété de Darboux (f prend entre a et b toute valeur comprise entre $f(a)$ et $f(b)$; autrement dit, l'image par f d'un intervalle est connexe).
- (β) La propriété de Baire (f est limite d'une suite de fonctions continues donc, pour tout sous-ensemble P parfait de I et tout $\epsilon > 0$, il existe une portion P' de P où l'oscillation de f est inférieure à ϵ).
- (γ) Le fait que, si sur un intervalle $J \subset I$, on a $f \geqslant 0$ partout et $f \geqslant a$ (resp. $f \leqslant b$) presque partout, alors $f \geqslant a$ (resp. $f \leqslant b$) partout sur J.

Par conséquent, le théorème de Denjoy est applicable à toute fonction vérifiant (α), (β), (γ). [3] en donne des exemples.

2. Dès son mémoire de 1875 [1], Darboux avait donné un exemple de fonction dérivée discontinue sur un ensemble dense. Plusieurs auteurs, ensuite, ont donné ou cru donner des exemples de fonctions dérivées prenant sur tout intervalle des valeurs positives et des valeurs négatives. Des exemples concluants sont dus à Denjoy [2]. Des exemples plus simples se trouvent en [6] et [7] p. 32.

3. La caractérisation des distributions des fonctions dérivées sur l'intervalle *ouvert* $]0, 1[$ est évidente dès qu'on a le théorème de Denjoy: ce sont les mesures de probabilité sur $\mathbb{R}$ dont le support est connexe. En effet, une telle mesure est la distribution d'une (unique) fonction continue croissante (au sens large) sur $]0, 1[$.

4. Les fonctions dérivées sur l'intervalle *fermé* $[0, 1]$ peuvent être considérées comme limites de martingales dyadiques, définies sur $\Omega = \{0, 1\}^{\mathbb{N}}$, et

partout convergentes sur Ω. En effet, F désignant la primitive et $\omega = (\omega_1, \omega_2, \dots)$ un point de Ω, la suite f_n définie par

$$f_n(\omega) = 2^{-n}\left(F\left(\frac{\omega_1}{2} + \cdots + \frac{\omega_n}{2^n} + \frac{1}{2^n}\right) - F\left(\frac{\omega_1}{2} + \cdots + \frac{\omega_n}{2^n}\right)\right)$$

est une martingale qui converge vers la dérivée $f(x)$ au point $x = \sum_1^\infty 2^{-n}\omega_n$. Les distributions des limites de martingales partout convergentes se caractérisent ainsi: ce sont les mesures de probabilité sur $\mathbb{R}$ qui vérifient la condition (b), cf. [8].

Références

[1] Darboux, G. Mémoire sur les fonctions discontinues, 1875. [Pour un résumé, voir *Bull. Sc. Math. France* **10**(1876), 76-82].
[2] Denjoy, A. Sur les fonctions dérivées sommables. *Bull. Soc. Math. France* **43**(1915).
[3] — Sur une propriété des fonctions dérivées. *L'Enseignement Mathématique.* **18**(1916), 320-328.
[4] Clarkson, J. A. A property of derivatives. *Bull. Amer. Math. Soc.* **53**(1947), 124-125.
[5] Zahorski, Z. Sur la première dérivée. *Trans. Amer. Math. Soc.* **69**(1950), 1-54.
[6] Katznelson Y. and Stromberg, O. Everywhere differentiable, nowhere monotone functions. *Amer. Math. Monthly* **81**(1974), 349-354.
[7] Bruckner, A. M. Differentiation of real functions. *Lecture Notes in Math.* **659**(1978).
[8] Kahane, J. P. Une théorie de Denjoy des martingales dyadiques. *L'Enseignement Mathématique* (1988), volume à paraître.

Jean-Pierre Kahane
Université de Paris-Sud
Mathématique (Bât. 425)
91405 Orsay Cedex,
FRANCE

REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.° 2, 1988

An Extremal Problem for Certain Subharmonic Functions in the Plane

Albert Baernstein II

Dedicated to the memory of José Luis Rubio de Francia

1. Introduction

Let $\delta \in (0, 1)$, and suppose that $K \subset \mathbb{R}$ is a closed set which is δ-dense in the sense that

$$|K \cap I| \geqslant \delta \tag{1.1}$$

for every interval $I \subset \mathbb{R}$ with $|I| \equiv$ length $I = 1$. Suppose also that every point of K is regular for the Dirichlet problem on the domain $\mathbb{C} \setminus K$. Then there exists a function $u = u(\cdot, K): \mathbb{C} \to [0, \infty)$ with the following properties:

(i) u is harmonic in $\mathbb{C} \setminus K$ and continuous on $\mathbb{C}$.
(i) $u = 0$ on K.
(iii) $u(z) = u(\bar{z}), \quad z \in \mathbb{C}$.
(iv) $\lim_{y \to +\infty} \dfrac{u(x + iy)}{y} = 1, \quad x \in \mathbb{R}$.

Such functions were constructed by A. E. Schaeffer [S], who applied them to extend Bernstein's theorem about sup $|f'(x)|$ for entire functions of exponential type. More recently, Benedicks [Be] has studied a class of functions which includes $u(\cdot, K)$, and has proved, among other things, a conjecture of Kjellberg. From results in [Be] it follows that $u(\cdot, K)$ is unique.

B. Ya. Levin [L] posed the extremal problem of finding a δ-dense K which maximizes $\sup_{\mathbb{R}} u(x, K)$. Define

$$K_\delta = \bigcup_{n=-\infty}^{\infty} \left[n - \frac{1}{2}\delta, n + \frac{1}{2}\delta \right].$$

Levin's conjecture. For each $x \in \mathbb{R}$,

$$u(x, K) \leqslant u\left(\frac{1}{2}, K_\delta\right) = \frac{1}{\pi} \log\left(\cot\frac{\pi\delta}{4}\right). \tag{1.2}$$

We show in Section 3 how to calculate this value of $u(1/2, K_\delta)$.

In this paper we shall prove Levin's conjecture in stronger form, by showing that for every convex nondecreasing function Φ on $[0, \infty)$, $y \in \mathbb{R}$, and interval $I \subset \mathbb{R}$ with $|I| = 1$,

$$\int_I \Phi(u(x + iy, K))\, dx \leqslant \int_I \Phi(u(x + iy, K_\delta))\, dx. \tag{1.3}$$

Then (1.2) follows by taking $y = 0$, $\Phi(t) = t^p$, and letting $p \to \infty$.

The functions $u(\cdot, K)$ may be regarded as limiting cases of certain harmonic measures. For $B > 0$ define

$$\Omega(B, K) = \{z: |\operatorname{Im} z| < B, z \notin K\}.$$

Let $u(\cdot, B, K)$ be the function which is harmonic in $\Omega(B, K)$ and takes the boundary value zero on K, one on $\operatorname{Im} z = \pm B$. In Section 3 we will show that locally uniformly in $\mathbb{C}$,

$$\lim_{B \to \infty} Bu(z, B, K) = u(z, K). \tag{1.4}$$

Thus, (1.3) is a consequence of the following theorem.

Theorem 1. *If K satisfies* (1.1), *then for every interval $I \subset \mathbb{R}$ with $|I| = 1$, every $y \in [-B, B]$, and every convex non-decreasing function Φ on* $[0, \infty)$,

$$\int_I \Phi(u(x + iy, B, K))\, dx \leqslant \int_I \Phi(u(x + iy, B, K_\delta))\, dx. \tag{1.5}$$

We can also prove a strong uniqueness theorem.

Theorem 2. *Suppose that, with the situation of Theorem* 1, *equality holds in* (1.5) *for some Φ with $\Phi(0) = 0$, some $y \in (-B, B)$ and some I. Then either K is translation of K_δ or both sides are zero. Moreover, if K is not a translation of K_δ then for every $x \in \mathbb{R}$ and $y \in (-B, B)$,*

$$u(x+iy, B, K) < u\left(\frac{1}{2}+iy, B, K_\delta\right). \tag{1.6}$$

The corresponding uniqueness theorem also holds for $u(\cdot, K)$ and $u(\cdot, K_\delta)$. It may be proved by the same argument used to prove Theorem 2.

In [L] Levin offers the uniqueness conjecture that

$$\max u(x, K) = u\left(\frac{1}{2}, K_\delta\right)$$

implies K is a translation of K_δ. Our inequality (1.6), which holds for $u(\cdot, K)$ and $u(\cdot, K_\delta)$, confirms this when K is a set of period 1. When K is a nonperiodic set, for example, if $K = K_\delta \cup [0, 1]$, then it is easy to see that the maximum need not exist, and that

$$\sup u(x, K) = u\left(\frac{1}{2}, K_\delta\right)$$

can occur. In any case, if K is not a translation of K_δ then (1.6) holds for $u(\cdot, K)$ and $u(\cdot, K_\delta)$.

Theorem 1 is a close relative of a theorem in [Ba2, p. 167] about the behavior of harmonic measure under circular symmetrization. Suppose that Ω is a subdomain of an annulus $R_1 < |\zeta| < R_2$, $0 < R_1 < R_2 < \infty$. Let Ω^* be the domain obtained from Ω by circular symmetrization, and u be the function harmonic in Ω with boundary value zero on $\partial\Omega \cap \{R_1 < |\zeta| < R_2\}$, and one on $\{|\zeta| = R_1\} \cup \{|\zeta| = R_2\}$. Let v be the corresponding function for Ω^*.

Theorem A. *For $r \in [R_1, R_2]$ and convex nondecreasing Φ on $[0, \infty)$,*

$$\int_{-\pi}^{\pi} \Phi(u(re^{i\theta}))\, d\theta \leqslant \int_{-\pi}^{\pi} \Phi(v(re^{i\theta}))\, d\theta.$$

In [Ba2] the theorem is stated for subdomains of a disk. The annulus case is proved exactly the same way. When K has period 1 Theorem 1 follows directly from Theorem A via conformal mapping. The chief novelty of Theorem 1 is that we are able to compare a nonperiodic K with the periodic K_δ. The methods used here to do this can be applied to solve various other extremal problems in which the data are essentially one-dimensional.

It would be interesting to prove results of this type involving higher dimensional periodicities. Suppose, for example, that $K \subset \mathbb{R}^2$ is δ-dense in the sense that area $(K \cap \Delta) \geqslant \delta$ for every disk Δ with area $(\Delta) = 1$. For $B > 0$ define

$$\Omega(K, B) = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : (x_1, x_2) \notin K, |x_3| < B\}$$

and let $u(\cdot, K, B)$ be the harmonic measure, for the Laplacian in $\mathbb{R}^3$, of

$x_3 = \pm B$ with respect to $\Omega(K, B)$. Lemma 8 of [Be, p. 63] implies existence of a constant $C(\delta, B)$ such that

$$\sup_{(x_1, x_2) \in \mathbb{R}^2} u(x_1, x_2, 0, K, B) \leqslant C(\delta, B).$$

One can pose the problem of finding the extremal configuration K. Theorems about packing and covering in the plane, see, *e.g.* [FT, Chap. 4], suggest that the extremal K should be associated with the decomposition of $\mathbb{R}^2$ into regular hexagons. The conjecture of Ahlfors and Grunsky, see, *e.g.* [M], according to which the extremal case for the theorem of Bloch involves the plane punctured at the centers of these hexagons, provides another question of this type.

The main tool used to prove Theorem A was the subharmonicity of a certain maximal function u^*. See Theorem B in Section 2. In order to adapt this proof we introduce in Section 2 certain variants u^*_λ, whose subharmonicity under appropriate hypotheses is established in Theorem 3.

Essén and Shea [ES] proved a uniqueness theorem associated with Theorem A. For annuli, it states that if equality holds for some $r \in (R_1, R_2)$ and some strictly convex Φ then Ω^* is a rotation of Ω. Our Theorem 2 has a stronger conclusion, in that «strictly» is not assumed. The method also works for sufficiently regular domains Ω, and gives a slight extension of the Essén-Shea theorem for those domains. Probably with more effort it could be made to work for all domains.

It is a pleasure to acknowledge discussions of Levin's problem I had with J. L. Fernández in Zaragoza while visiting Spain under the auspices of the Hispano-American accord.

In sadness I dedicate this paper to the memory of José Luis Rubio de Francia. I met him on only a few occasions, but his penetrating mind, cheerful demeanor, and strength of character made a vivid impression that I will carry with me always.

2. A Variant of the *-function

Suppose that $-\infty < A_1 < A_2 < \infty$ and that f is a real valued integrable function on $[A_1, A_2]$. Let

$$L = \frac{1}{2}(A_2 - A_1).$$

Define the symmetric decreasing rearrangement $\tilde{f}: [-L, L] \to \mathbb{R}$ and the $*$-function $f^*: [0, L] \to \mathbb{R}$ by

$$\tilde{f}(x) = \inf\{t: |\{y: f(y) > t\}| \leqslant 2x\}, \qquad |x| < L,$$

$$\tilde{f}(\pm L) = \lim_{x \to L-} \tilde{f}(x) = \operatorname{ess\,inf} f,$$

$$f^*(x) = \sup \int_E f(t)\, dt, \tag{2.1}$$

where the supremum is over all $E \subset [A_1, A_2]$ with $|E| = 2x$. These functions were studied in [Ba2, 3] where the following properties were proved.

(2.2) For each $x \in [0, L]$ there exists $E \subset [A_1, A_2]$ for which the supremum in (2.1) is attained.

$$f^*(x) = \int_{-x}^{x} \tilde{f}(t)\, dt, \qquad 0 \leqslant x \leqslant L. \tag{2.3}$$

(2.4) For $f, g \in L^1[A_1, A_2]$ the following two statements are equivalent:

(*a*) $$f^*(x) \leqslant g^*(x), \quad \text{for every} \quad x \in [0, L],$$

(*b*) $$\int_{A_1}^{A_2} \Phi(f(x))\, dx \leqslant \int_{A_1}^{A_2} \Phi(g(x))\, dx,$$

for any convex non-decreasing function Φ on $(-\infty, \infty)$.

These results were stated for $A_1 = -\pi$, $A_2 = \pi$ in [Ba2]. The more general versions considered here can be reduced to that one. The argument which proves (2.4) gives also a statement about strict inequality.

(2.5) For non-negative $f, g \in L^1[A_1, A_2]$, with g not the zero function, the following two statements are equivalent:

(*a*) $f^*(x) < g^*(x)$, for every $x \in (0, L]$,

(*b*) for each convex non-decreasing Φ on $(0, \infty)$ with $\Phi(0) = 0$, either

$$\int_{A_1}^{A_2} \Phi(f(x))\, dx < \int_{A_1}^{A_2} \Phi(g(x))\, dx,$$

or both integrals are zero.

Suppose next that u is a real function defined in the annulus $R_1 < |\zeta| < R_2$ for which $u(re^{i\theta}) \in L^1[-\pi, \pi]$, $R_1 < r < R_2$. Then u^* is defined in the semi-annulus $S = \{re^{i\theta}: R_1 < r < R_2, 0 \leqslant \theta \leqslant \pi\}$.

Theorem B. *Suppose that u is subharmonic in $R_1 < |\zeta| < R_2$. Then*

(*a*) $u^* \in C(S)$.

(*b*) *u^* is subharmonic in the interior of S.*

(*c*) *For each $\theta \in [0, \pi]$ the function $r \to u^*(re^{i\theta})$ is a convex function of $\log r$ on (R_1, R_2).*

For (*a*) and (*b*) see [Ba2, pp. 141, 147] and for (*c*) [Ba3, p. 85].

Let us return now to $f \in L^1[A_1, A_2]$. For $E \subset \mathbb{R}$ write $\operatorname{diam} E = \sup E - \inf E$. For $0 < \lambda \leqslant L$ define $f_\lambda^*: [0, \lambda] \to \mathbb{R}$ by

$$f_\lambda^*(x) = \sup \int_E f(t)\, dt \tag{2.6}$$

where the supremum is over all $E \subset [A_1, A_2]$ with $|E| = 2x$ and $\operatorname{diam} E \leqslant 2\lambda$. Relatives of f_λ^* appear in [Ba1]. It is easy to see that again a set $E = E(x, \lambda)$ exists for which the supremum in (3.6) is attained.

We adopt the following notation for open rectangles:

$$R(A_1, A_2, B_1, B_2) = \{x + iy: A_1 < x < A_2, B_1 < y < B_2\}.$$

The closure of such an R will be denoted by $\bar{R}$.

Suppose that u is defined in $R(A_1, A_2, B_1, B_2)$ and that $u(x + iy) \in L^1[A_1, A_2]$ for each $y \in (B_1, B_2)$. Then u_λ^* is defined in $R(0, \lambda, B_1, B_2)$. In general, subharmonicity of u does not imply subharmonicity of u_λ^*. Consider, for example, $u(x, y) = x^2 - y^2$ in $R(-A, A, -\infty, \infty)$, with $A \in (1, \infty)$. Take $\lambda = 1/2$, then for $x + iy \in R(0, 1/2, -\infty, \infty)$,

$$u_\lambda^*(x + iy) = \int_{A-2x}^{A} (t^2 - y^2)\, dt, \qquad \Delta u_\lambda^* = 12x - 8A < 0.$$

To get a positive result we need to impose a condition which prevents the maximal sets E from hitting the left or right hand boundary of R. This is the role of (2.7) in the theorem below. Recall that $L = (A_2 - A_1)/2$.

Theorem 3. *Suppose that u is subharmonic in $R(A_1, A_2, B_1, B_2)$ and belongs to $L^1[A_1, A_2]$ for each $y \in (B_1, B_2)$, that $0 < \lambda < L/2$, and that, as $z \to \zeta$ from inside $R(A_1, A_2, B_1, B_2)$,*

$$\overline{\lim_{z \to \zeta}}\, u(z) < \varliminf_{z \to \zeta} u(z + 2\lambda), \qquad \overline{\lim_{z \to \zeta}}\, u(z) < \varliminf_{z \to \zeta} u(z - 2\lambda), \tag{2.7}$$

where in the first inequality $\zeta = A_1 + iy$ and in the second $\zeta = A_2 + iy$, $B_1 < y < B_2$. Then

(*a*) $u_\lambda^* \in C[R(0, \lambda, B_1, B_2) \cup \{\lambda + iy: B_1 < y < B_2\}]$.

(*b*) *u_λ^* is subharmonic in $R(0, \lambda, B_1, B_2)$.*

(*c*) *For each $x \in [0, \lambda]$, the function $y \to u_\lambda^*(x + iy)$ is convex on (B_1, B_2).*

For the proof we need two lemmas about sets $E \subset \mathbb{R}$. For $\epsilon \in \mathbb{R}$ write $E_\epsilon = E + \epsilon = \{x + \epsilon: x \in E\}$. Equivalence of two sets means that they differ by a set of measure zero. Also, we will write $\sup E$, $\inf E$, and $\operatorname{diam} E$ instead of ess sup E, ess inf E, ess diam E to denote the essential supremum, infimum and diameter of E.

Lemma 1. *Suppose that E is bounded, $|E| > 0$, and that E is not equivalent to an interval. Then for all sufficiently small $\epsilon > 0$, $|E_\epsilon \cap E_{-\epsilon}| \leqslant |E| - 4\epsilon$.*

PROOF. If E is equivalent to a pair of disjoint closed intervals the lemma follows by direct verification. Suppose that E is equivalent to neither a single interval nor a pair. Then there exist points $a_1 < b_1 < a_2 < b_2 < a_3 < b_3$ such that the a_i are density points of $\mathbb{R}\setminus E$ and the b_i are density points of E. Take $c_1 \in (b_1, a_2)$, $c_2 \in (b_2, a_3)$. Then for $\epsilon > 0$ and $\chi = \chi_E$,

$$|E_\epsilon \cap E_{-\epsilon}| = \int_{-\infty}^{\infty} \chi(t+\epsilon)\chi(t-\epsilon)\,dt = \int_{-\infty}^{c_1} + \int_{c_1}^{c_2} + \int_{c_2}^{\infty}.$$

Take $\varkappa > 0$. The proof on page 146 of [Ba2] shows that for ϵ sufficiently small

$$\int_{-\infty}^{c_1} \chi(t+\epsilon)\chi(t-\epsilon)\,dt \leqslant \int_{-\infty}^{c_1} \chi(t+\epsilon)\,dt - (2-\varkappa)\epsilon.$$

The same inequality holds when $(-\infty, c_1)$ is replaced by (c_1, c_2) and (c_2, ∞). Hence

$$|E_\epsilon \cap E_{-\epsilon}| \leqslant \int_{-\infty}^{\infty} \chi(t+\epsilon)\,dt - (6-3\varkappa)\epsilon < |E| - 4\epsilon,$$

provided $\varkappa < 2/3$.

Lemma 2. *Let E satisfy the hypotheses of Lemma 1. Then for each sufficiently small $\epsilon > 0$ there exist sets $E', E'' \subset \mathbb{R}$ such that*

(*a*) $E_\epsilon \cup E_{-\epsilon} = E' \cup E''$ *and* $E_\epsilon \cap E_{-\epsilon} = E' \cap E''$.
(*b*) $|E'| = |E| - 2\epsilon$ *and* $|E''| = |E| + 2\epsilon$.
(*c*) $\operatorname{diam} E' \leqslant \operatorname{diam} E$ *and* $\operatorname{diam} E'' \leqslant \operatorname{diam} E$.

PROOF. Let $F = (E_\epsilon \cup E_{-\epsilon})\setminus(E_\epsilon \cap E_{-\epsilon})$. Then $|F| \geqslant 4\epsilon$, by Lemma 1. Let G be a set of the form $G = F \cap [a, \infty)$ for which

$$|G| = |E| - |E_\epsilon \cap E_{-\epsilon}| - 2\epsilon.$$

Then, by Lemma 1, G exists and $|G| \geqslant 2\epsilon$. Define $E' = (E_\epsilon \cap E_{-\epsilon}) \cup G$, $E'' = (E_\epsilon \cap E_{-\epsilon}) \cup (F\setminus G)$. Then (*a*) is true. Also, $|E'| = |E| - 2\epsilon$. Now (*a*) can be restated as

$$\chi_{E_\epsilon} + \chi_{E_{-\epsilon}} = \chi_{E'} + \chi_{E''}.$$

Hence $2|E| = |E'| + |E''|$, so that $|E''| = |E| + 2\epsilon$ and (*b*) is proved.

Suppose that $\inf G < \epsilon + \inf E$. (Recall that inf means essential infimum.) Then, since $(F\setminus G) \subset [-\epsilon + \inf E, \inf G]$, we would have $|F\setminus G| < 2\epsilon$, so that

$|E''| < |E| + 2\epsilon$, a contradiction, Similarly, $\sup(F\setminus G) > -\epsilon + \sup E$ would imply $|G| < 2\epsilon$, contradicting an earlier inequality. Hence

$$\inf G \geqslant \epsilon + \inf E, \qquad \sup(F\setminus G) \leqslant -\epsilon + \sup E.$$

Thus $E' \subset [\epsilon + \inf E, \epsilon + \sup E]$, $E'' \subset [-\epsilon + \inf E, -\epsilon + \sup E]$, and (*c*) is proved.

PROOF OF THEOREM 3. Suppose first that $u \in C^2(R(A_1, A_2, B_1, B_2))$ and that $u \in C(\bar{R}(A_1, A_2, B_1, B_2)$. Then continuity of u_λ^* follows from a straightforward argument, and we omit it. Let us prove subharmonicity. Take $z_0 = x_0 + iy_0 \in R(0, \lambda, B_1, B_2)$ and $E_0 \subset [A_1, A_2]$ with $\operatorname{diam} E_0 \leqslant 2\lambda$, $|E_0| = 2x_0$ and

$$u_\lambda^*(z_0) = \int_{E_0} u(t + iy_0)\, dt. \tag{2.8}$$

From (2.7) it follows that there exists $\varkappa > 0$ such that

$$u(t + iy_0) < u(t + 2\lambda + iy_0), \qquad u(t + iy_0) < u(t - 2\lambda + iy_0),$$

when, respectively,

$$A_1 \leqslant t \leqslant A_1 + \varkappa \quad \text{or} \quad A_2 - \varkappa \leqslant t \leqslant A_2.$$

I claim that E_0 is essentially contained in $[A_1 + \varkappa, A_2 - \varkappa]$. Let

$$b = \sup E_0 = (\operatorname{ess\,sup} E_0), \qquad a = \inf E_0, \quad \text{and} \quad F = E_0 \cap [A_2 - \varkappa, A_2].$$

If $|F| > 0$ then $b \in (A_2 - \varkappa, A_2]$ and hence $a \geqslant b - 2\lambda > A_2 - \varkappa - 2\lambda$. Let

$$E_1 = (E_0\setminus F) \cup (-2\lambda + F).$$

Then $|E_1| = |E_0|$, $\operatorname{diam} E_1 \leqslant 2\lambda$, and the integral in (2.8) strictly increases when E_0 is replaced by E_1, contradicting the definition of u_λ^*. Similarly, $|E_0 \cap [A_1, A_1 + \varkappa]| = 0$. Thus, after possible deletion of a null set,

$$E_0 \subset [A_1 + \varkappa, A_2 - \varkappa]. \tag{2.9}$$

For $0 < \rho < \varkappa$, $-\pi/2 \leqslant \phi \leqslant \pi/2$, define a function Q by

$$Q(\rho, \phi) = \int_{E_0} [u(t + \rho\cos\phi + i(y_0 + \rho\sin\phi)) + u(t - \rho\cos\phi + i(y_0 + \rho\sin\phi))]\, dt.$$

Using (2.8) and subharmonicity of u, we have

$$2\pi u_\lambda^*(z_0) \leqslant \int_{E_0} dt \int_{-\pi}^{\pi} u(t + iy_0 + \rho e^{i\phi})\, d\phi = \int_{-\pi/2}^{\pi/2} Q(\rho, \phi)\, d\phi. \tag{2.10}$$

Fix ρ, ϕ and write $y = y_0 + \rho \sin\phi$, $\epsilon = \rho \cos\phi$. Then

$$Q(\rho, \phi) = \int_{(E_0)_\epsilon} u(t + iy)\, dt + \int_{(E_0)_{-\epsilon}} u(t + iy)\, dt.$$

If E_0 is not equivalent to an interval, then, for small enough ρ, we can find sets E', E'' for E_0 satisfying the conclusions of Lemma 2, and (2.9) guarantees that these sets are contained in $[A_1, A_2]$. From Lemma 2(a) it follows that

$$Q(\rho, \phi) = \int_{E'} u(t + iy)\, dt + \int_{E''} u(t + iy)\, dt,$$

and from Lemma 2(b), (c) we obtain

$$Q(\rho, \phi) \leqslant u_\lambda^*(z_0 - \rho e^{i\phi}) + u_\lambda^*(z_0 + \rho e^{i\phi}). \tag{2.11}$$

If $E_0 = [a, b]$, then $b - a = 2x_0 < 2\lambda$ and we achieve (2.11) by taking $E' = [a - \epsilon, b - \epsilon]$, $E'' = [a + \epsilon, b + \epsilon]$ (assumming $\rho < \lambda - x_0$). Inserting (2.11) in (2.10), we arrive at

$$2\pi u_\lambda^*(z_0) \leqslant \int_{-\pi}^{\pi} u_\lambda^*(z_0 + \rho e^{i\phi})\, d\phi, \tag{2.12}$$

so that u_λ^* is subharmonic in $R(0, \lambda, B_1, B_2)$.

Turning now to the proof of convexity in the C^2 case, let $u_\epsilon(z) = u(z) + \epsilon y^2$. It will suffice to prove that $(u_\epsilon)_\lambda^*(x + iy)$ is convex in y for each $\epsilon > 0$.

Take $z_0 = x_0 + iy_0$ with $0 < x_0 \leqslant 2\lambda$, $y_0 \in (B_1, B_2)$, and let E_0 be a maximal set for u_λ^* as in (2.8). Then (2.9) holds. Since u_λ^* and $(u_\epsilon)_\lambda^*$ have the same maximal sets, (2.8) holds also with u replaced by u_ϵ. Define

$$J_1(x) = \int_{E_0} u_\epsilon(x + t + iy_0)\, dt, \qquad J_2(y) = \int_{E_0} u_\epsilon(t + iy)\, dt.$$

Then J_1 is C^2 in a neighborhood of 0, J_2 is C^2 in a neighborhood of y_0, and J_1 has a local maximum at $x = 0$. Hence

$$\begin{aligned} 0 \geqslant J_1''(0) &= \int_{E_0} (u_\epsilon)_{xx}(t + iy_0)\, dt \\ &= \int_{E_0} [(\Delta u_\epsilon)(t + iy_0) - (u_\epsilon)_{yy}(t + iy_0)]\, dt \\ &\geqslant 2\epsilon |E_0| - J_2''(y_0). \end{aligned}$$

Thus $J_2''(y_0) > 0$, and J_2 is convex in a neighborhood of y_0. Hence, for small $\tau > 0$,

$$\begin{aligned} (u_\epsilon)_\lambda^*(z_0) = J_2(y_0) &\leqslant \frac{1}{2}[J_2(y_0 + \tau) + J_2(y_0 - \tau)] \\ &\leqslant \frac{1}{2}[(u_\epsilon)_\lambda^*(y_0 + \tau) + (u_\epsilon)_\lambda^*(y_0 - \tau)], \end{aligned}$$

which proves (c) in the C^2 case.

It remains to prove Theorem 3 in the nonsmooth case. Take B_3, B_4 with $B_1 < B_3 < B_4 < B_2$. By (2.7), there exists $A_3 \in (A_1, A_1 + \lambda)$, $A_4 \in (A_2 - \lambda, A_2)$ such that $u(z) < u(z + 2\lambda)$, $u(z) < (z - 2\lambda)$ when $z = x + iy$ with $y \in [B_3, B_4]$ and, respectively, $x \in (A_1, A_3)$ or $x \in (A_4, A_2)$. As in the C^2 case any maximal set for $u_\lambda^*(x + iy)$ will be essentially contained in $[A_3, A_4]$ when $y \in (B_3, B_4)$.

Let $K \geqslant 0$ be a radial C^2 function on $\mathbb{C}$ such that $K(|z|)$ decreases as $|z|$ increases, $K(z) = 0$ for $|z| \geqslant 1/2$, and $\int_{\mathbb{C}} K(z)\, dx\, dy = 1$. Let u_ϵ be the $\mathbb{R}^2$-convolution of u with $\epsilon^{-2} K(z\epsilon^{-1})$. Then u_ϵ is defined on a slightly smaller rectangle than u, and satisfies a boundary condition like (2.7) when ϵ is small. For each $z \in R\,(A_1, A_2, B_1, B_2)$ we have $u_\epsilon(z) \downarrow u(z)$. Moreover, using continuity of the mean value

$$y \to \int_{A_1 + \tau}^{A_2 - \tau} u(x + iy)\, dx, \qquad \tau > 0,$$

as in the proof of (11) on page 145 of [Ba2], one can show that for each $\tau > 0$,

$$\lim_{\epsilon \to 0} \sup_{y \in [B_3, B_4]} \int_{A_1 + \tau}^{A_2 - \tau} |u_\epsilon(x + iy) - u(x + iy)|\, dx = 0.$$

Using this, it is not hard to prove that $(u_\epsilon)_\lambda^* \to u_\lambda^*$ uniformly on $\bar{R}[0, \lambda, B_3, B_4]$. Then the general case of Theorem 3 follows from the C^2 case.

3. Other Preliminaries

There are also versions of u_λ^* and of Theorem 3 in the periodic case. Suppose that f is an integrable real valued function on the unit circle T and that $0 < \lambda \leqslant \pi$. Define $f_\lambda^*: [0, \lambda] \to \mathbb{R}$ by

$$f_\lambda^*(\theta) = \sup \int_E f(e^{it})\, dt, \tag{3.1}$$

where the supremum is over all $E \subset T$ with $|E| = 2\theta$ and E contained in some arc I of T with $|I| = 2\lambda$. This f_λ^* differs from the one defined in Section 2 for the function $\theta \mapsto f(e^{i\theta})$ on $[-\pi, \pi]$, since in (3.1) sets like $E = \{e^{i\theta}: \pi - \epsilon < \theta < \pi + \epsilon\}$ are permitted in the competition, when $2\epsilon \leqslant 2\lambda$.

Suppose that u is defined in the annulus $R_1 < |z| < R_2$ and is integrable on each r circle, $R_1 < r < R_2$. Then u_λ^* is defined in the annular sector

$$S = \{re^{i\theta}: R_1 < r < R_2, 0 \leqslant \theta \leqslant \lambda\}.$$

Theorem 4. *Suppose that u is subharmonic in $R_1 < |z| < R_2$. Then*

(a) $u_\lambda^* \in C(S)$.

(*b*) *u_λ^* is subharmonic in the interior of S.*

(*c*) *For each $\theta \in [0, \lambda]$, $r \mapsto u^*(re^{i\theta})$ is a convex function of* $\log r$ *on* (R_1, R_2).

When $\lambda = \pi$ this theorem reduces to Theorem B. Note that in the periodic case we do not need any extra conditions such as (2.7). The proof of Theorem 4 follows the same pattern as that of Theorem 3, with a few simplifications. We will not use Theorem 4 in this paper, except for the previously known case $\lambda = \pi$. It is recorded here for possible use in other contexts.

Next, we examine the extremal function $u(z, B, K_\delta)$. For fixed $\delta \in (0, 1)$ and $0 < B < \infty$ define

$$v(z) = u(z, B, K_\delta). \tag{3.2}$$

Then

$$v(z) = v(z + 1) \quad \text{for} \quad z \in R\ (-\infty, \infty, -B, B).$$

Thus there is a well defined function V in the annulus $e^{-2\pi B} \leqslant |\zeta| \leqslant e^{2\pi B}$ given by

$$v(z) = V(-e^{2\pi i z}).$$

This function is harmonic in the annulus with the slit $\{e^{i\phi}: \pi - \pi\delta \leqslant \phi \leqslant \pi + \pi\delta\}$ deleted, takes the value zero on the slit and the value one when $|\zeta| = e^{\pm 2\pi B}$. The argument in Proposition 5 of [Ba2, p. 153] is applicable, and one finds that $V(re^{i\phi})$ is a symmetric decreasing function of ϕ, for $|\log r| \leqslant 2\pi B$. Equivalently, for $z = re^{i\theta}$ in the upper half of the annulus,

$$V^*(re^{i\theta}) = \int_{-\theta}^{\theta} V(re^{i\phi})\, d\phi$$

where V^* is the usual periodic *-function (with $\lambda = \pi$). Moreover, V^* is continuous in the closed half-annulus, harmonic in the open half-annulus with $\{e^{i\theta}: \pi - \pi\delta \leqslant \phi \leqslant \pi\}$ deleted and, since $V^*(re^{i\pi})$ is the mean value over the circle $|\zeta| = r$, it is a linear function of $\log r$ for $1 < r < e^{2\pi B}$.

Returning to v, define, for $z \in \bar{R}(0, 1/2, -B, B)$,

$$v^*(z) = \int_{-x}^{x} v\left(\frac{1}{2} + t + iy\right) dt. \tag{3.3}$$

Define also

$$\alpha = \frac{1}{2}(1 - \delta). \tag{3.4}$$

Then $v^*(z) = V^*(e^{2\pi i z})$ and the properties of V^* translate to the following properties of v^*.

Proposition 1. *For v defined by* (3.2),

(*a*) $v^* \in C(\bar{R}(0, 1/2, -B, B))$,
(*b*) v^* *is harmonic in* $R(0, 1/2, -B, B) \setminus [\alpha, 1/2)$,
(*c*) $y \to v^*(1/2 + iy)$ *is linear on* $[0, B]$.
(*d*) $\sup_K v(x) = v\left(\begin{matrix}1\\2\end{matrix}\right)$.

Next, we shall show that

$$\lim_{B\to\infty} Bu\left(\begin{matrix}1\\2\end{matrix}, B, K_\delta\right) = \frac{1}{\pi}\log\left(\cot\frac{\pi\delta}{4}\right). \tag{3.5}$$

Let

$$F_1(z) = \sin \pi z, \qquad F_2(w) = \frac{1}{2}\left(\sin\frac{\pi\delta}{2}\right)(w + w^{-1}).$$

Then F_1 conformally maps the half strip $R(-1/2, 1/2, 0, \infty)$ onto the upper half plane, and F_2 conformally maps the exterior of the unit disk onto

$$\mathbb{C}\setminus\left[-\sin\frac{\pi\delta}{2}, \sin\frac{\pi\delta}{2}\right].$$

Let

$$\rho = \frac{1}{4}\left(\csc\frac{\pi\delta}{2}\right)e^{\pi B}.$$

Then, for large B, $F_1^{-1} \circ F_2$ maps the half disk $1 < |w| < \rho$, $\operatorname{Im} w > 0$, into $R(-1/2, 1/2, 0, B)$. Let $h_B(w) = u(F_1^{-1}(F_2(w)), B, K_\delta)$. Then h_B is harmonic in the half disk, $h_B = 0$ on $|w| = 1$, $h_B \leqslant 1$ on $|w| = \rho$ and $\partial h_B/\partial n = 0$ on $\operatorname{Im} w = 0$, $1 < |\operatorname{Im} w| < \rho$, except perhaps at the points $F_2^{-1} \circ F_1(\pm 1/2)$, where h_B remains bounded. Extend h_B to the annulus $1 < |w| < e^{\pi B}$ by $h_B(\bar{w}) = h_B(w)$. Then h_B is harmonic in the annulus and

$$h_B(w) \leqslant \frac{\log|w|}{\log\rho}, \qquad 1 < |w| < \rho.$$

Now

$$F_2^{-1} \circ F_1\left(\begin{matrix}1\\2\end{matrix}\right) = \cot\frac{\pi\delta}{4}.$$

Hence, as $B \to \infty$,

$$u\left(\begin{matrix}1\\2\end{matrix}, B, K_\delta\right) \leqslant \frac{\log\left(\cot\frac{\pi\delta}{4}\right)}{\pi B + O(1)}. \tag{3.6}$$

Choosing instead

$$\rho = 4\left(\csc\frac{\pi\delta}{2}\right)e^{\pi B},$$

then $F_1^{-1} \circ F_2$ maps the half disk onto a super-domain of $R(-1/2, 1/2, 0, B)$. One obtains the inequality opposite to (3.6), and (3.5) follows.

PROOF OF (1.4). This can easily be derived by the methods of [Be], but is particularly simple if one uses Theorem 1 and (3.5). These results imply existence of $C = C(\delta)$ such that $Bu(x, B, K) \leqslant C$ for all $B > 0$, $x \in \mathbb{R}$. Hence, by the maximum principle, for $0 \leqslant y \leqslant B$, and $z = x + iy$,

$$y \leqslant Bu(z, B, K) \leqslant y + C. \tag{3.7}$$

Thus, the symmetric functions $Bu(\bullet, B, K)$ possess a subsquence which converges locally uniformly on $\mathbb{C}\backslash K$ to a harmonic function u, which satisfies $|y| \leqslant u(z) \leqslant |y| + C$. So, for $x \in \mathbb{R}$,

$$\lim_{y\to\infty} \frac{u(x+iy)}{y} = 1.$$

It remains to show that for each $\zeta \in K$,

$$\lim u(z) = 0, \tag{3.8}$$

when $z \to \zeta$, $z \in K^c$.

Fix $\zeta \in K$ and let $S = \{z: A_1 < \operatorname{Re} z < A_2\}$ be a strip with $\zeta \in S$. Let H be the solution of the Dirichlet problem in $S\backslash K$ with boundary data $H = 0$ on $K \cap S$, $H(z) = C + |y|$ on ∂S. Since the boundary function is integrable with respect to the harmonic measure on $S\backslash K$ and each point of K is Dirichlet regular, potential theory shows that such an H exists, and that $\lim_{z\to\zeta} H(z) = 0$. From (3.7) we deduce $u(z) \leqslant H(z)$ in $S\backslash K$, and (3.8) follows.

For the proof of Theorems 1 and 2 it will be convenient to have a special form of the maximum principle. Suppose that $0 < \alpha < a$ and $b > 0$. Write $R = R(0, a, -b, b)$, and let w be a function defined on $\bar{R}$.

Proposition 2. *Suppose that*

(3.9) *w is subharmonic in $R\backslash[\alpha, a)$ and continuous on $\bar{R}$,*

(3.10) *$y \mapsto w(a + iy)$ is convex on $[0, b]$,*

(3.11) *$w(z) = w(\bar{z})$, $z \in \bar{R}$,*

(3.12) *$w(iy) = 0$ for $|y| \leqslant b$, and $w(x + ib) \leqslant 0$ for $0 \leqslant x \leqslant a$,*

(3.13) $w(x) = w(\alpha)$ *for* $\alpha \leqslant x \leqslant a$, *and*

$$\lim_{x \to \alpha -} \frac{w(\alpha) - w(x)}{\alpha - x} = 0.$$

Then

(3.14) $w \leqslant 0$ *in* $\bar{R}$.

Moreover, if either

(3.15) w *is not harmonic in* $R \setminus [\alpha, a)$

or

(3.16) $w(a + iy) < 0$ *for* $0 < y < b$,

then $w(z) < 0$ *for* $z \in R \cup \{a + iy: |y| < b\}$.

PROOF. Let $M = \sup_R w$. Then $M \geqslant 0$, by (3.12). The hypotheses imply that $M = w(z_0)$ for some z_0 with $\operatorname{Re} z_0 = 0$, $\operatorname{Im} z_0 = b$ or $z_0 = \alpha$. In the first two cases we have $M = 0$, which is (3.14). If $M > 0$ then $z_0 = \alpha$, w is non-constant, and hence $w < M$ everywhere in $R(0, \alpha, -b, b)$. From Hopf's lemma (see, *e.g.* [GT, p. 34]), it follows that $\partial w(\alpha)/\partial x > 0$ (derivative from the left). This contradicts (3.13), and therefore (3.14) holds.

If either (3.15) or (3.16) holds, then $w < 0$ in $R \setminus [\alpha, a)$. Hopf's lemma shows again that $w(\alpha) < 0$, and thus $w(x) < 0$ for $x \in [\alpha, a]$, by (3.13). Thus $w(a) < 0$, while $w(a + ib) \leqslant 0$, by (3.12). Hence $w(a + iy) < 0$ for $|y| < b$, by (3.10), and the proof is complete.

4. Proof of Theorem 1

We may assume that K contains no interval (a, ∞) or $(-\infty, a)$, since otherwise we can replace K by a closed subset with this property which still satisfies the δ-density condition. The corresponding $u(\cdot, B, K)$ pointwise dominates the original one.

Choose sequences $A_n' \downarrow -\infty$, $A_n'' \uparrow +\infty$ such that $1 + A_n' \notin K$, $-1 + A_n'' \notin K$. Let u_n be the function harmonic in $R(A_n', A_n'', -B, B) \setminus K$ with boundary values zero on $K \cup (\operatorname{Re} z = A_n' \text{ or } A_n'')$, one on $\operatorname{Im} z = \pm B$. Then $u_n \uparrow u(\cdot, B, K)$ locally uniformly in $|y| \leqslant B$, so it suffices to prove Theorem 1 with u_n in place of $u(\cdot, B, K)$, and $I \subset [A_n', A_n'']$.

When $A_n'' - A_n' > 1$, Theorem 3, with $\lambda = 1/2$, is applicable to u_n. Moreover, it is easy to see that $(u_n)_\lambda^* \in C[\bar{R}(0, 1/2, -B, B)]$. Let $w = (u_n)_\lambda^* - v^*$, where v^*

is defined by (3.3) and v by (3.2). Take a set E_0 with $|E_0| = 2\alpha = 1 - \delta$ for which

$$(u_n)^*_\lambda(\alpha) = \int_{E_0} u_n(t)\,dt$$

and $E_0 \subset I$, where I is an interval of length 1. Define

$$u^*_I(x) = \sup \int_E u_n(t)\,dt$$

where the supremum is over $E \subset I$ with $|E| = 2x$. By the δ-density condition and (2.3),

$$\lim_{x \to \alpha -} \frac{u^*_I(\alpha) - u^*_I(x)}{\alpha - x} = 0. \tag{4.1}$$

Now $u^*_I(\alpha) = (u_n)^*_\lambda(\alpha)$, and $u^*_I(x) \leqslant (u_n)^*_\lambda(x)$ for $0 \leqslant x < \alpha$. Thus (4.1) holds with $(u_n)^*_\lambda$ in place of u^*_I, and by (3.3) it holds for v. Hence w satisfies (3.13). From Theorem 3 and Proposition 1 it follows that w satisfies the other hypotheses of Proposition 2 in $R(0, 1/2, -B, B)$, so that $w \leqslant 0$ there. By (2.4), this proves Theorem 1.

5. Proof of Theorem 2. Periodic case

Assume in this section that there exists a closed set $K' \subsetneq [0, 1]$ such that

$$K = \bigcup_{n=-\infty}^{\infty} (n + K').$$

Fix $B > 0$ and write $u(z) = u(z, B, K)$. Then $u(z + 1) = u(z)$.

Define U in the annulus $e^{-2\pi B} \leqslant |\zeta| \leqslant e^{2\pi B}$ by the formula $U(e^{2\pi i z}) = u(z)$. Then u is the function harmonic in the annulus with the set $e^{2\pi i K'}$ deleted, which takes the values one on $|\zeta| = e^{\pm 2\pi B}$ and zero on $e^{2\pi i K'}$. Define

$$U^*(re^{i\theta}) = \sup_E \int_E U(re^{it})\,dt, \qquad u^*(z) = \sup_E \int_E u(t + iy)\,dt, \tag{5.1}$$

where in the first case the supremum is over subsets of the unit circle with $|E| = 2\theta$ and in the second over $E \subset \mathbb{R}$ with $|E| = 2x$ and $\operatorname{diam} E \leqslant 1$. Then $u^*(z) = U^*(e^{2\pi i z})$. By Theorem B, u^* is subharmonic in the rectangle

$$R = R(0, 1/2, -B, B),$$

continuous on its closure, and the function $y \to u^*(1/2 + iy)$ is convex on $[-B, B]$.

Let $w = u^* - v^*$, where v is defined in (3.2) and v^* by (3.3). Then $w \leqslant 0$ in R, by Theorem 1 and (2.4). By (2.5), the strict Φ-inequality statement of Theorem 2 is equivalent to $w(z) < 0$ for $z \in R \cup \{1/2 + iy: |y| < B\}$. By propositions 1 and 2, it is sufficient to show that w fails to be harmonic in $R \setminus [\alpha, 1/2)$, unless $e^{2\pi i K'}$ is a single arc on the unit circle with length exactly $2\pi\delta$.

Let $\alpha' = (1 - |K'|)/2$. Then $0 < \alpha' \leqslant \alpha = (1 - \delta)/2$. Let $H = e^{2\pi i K'}$. If H is a single arc then $\alpha' < \alpha$, and consideration of the function U shows that u^* is harmonic in $R \setminus [\alpha', 1/2)$ but not in $R \setminus [\alpha, 1/2)$. Hence w fails to be harmonic in $R \setminus [\alpha, 1/2)$, and we are done.

Assume, then that H is not a single arc. We are going to show that for x_0 slightly smaller than α' and for all small enough ρ,

$$u^*(x_0) < \frac{1}{2\pi} \int_{-\pi}^{\pi} u^*(x_0 + \rho e^{i\phi})\, d\phi. \tag{5.2}$$

Since v^* is harmonic at x_0 this will show that w is not harmonic in $R \setminus [\alpha, 1/2)$, as required.

Suppose that f is a continuous real function on the unit circle for which $|\{\phi \in [-\pi, \pi]: f(e^{i\phi}) = t\}| = 0$ for each $t > 0$. Let $E(t) = \{\phi: f(e^{i\phi}) > t\}$. Then the function $t \to |E(t)|$ is strictly decreasing and continuous for $0 < t < \sup f$, and for each $\theta \in (0, |E(0)|/2)$ there is an essentially unique set E for which the supremum

$$\sup_{|E| = 2\theta} \int_E f(e^{i\phi})\, d\phi, \qquad E \subset |\zeta| = 1,$$

is attained. This set is $E = E(t)$ with t chosen so that $|E(t)| = 2\theta$.

The function U is harmonic in $e^{-2\pi B} < |\zeta| < e^{2\pi B}$ except on H and is not constant on $|\zeta| = 1$. It follows that for $t > 0$ the sets $E(t) = \{\phi: U(e^{i\phi}) > t\}$ have only finitely many components. Thus, when $r = 1$ and $0 < \theta < 2\pi\alpha'$ each maximal set E in (5.1) can be taken to consist of finitely many open arcs, or of finitely many disjoint closed arcs. We work with the latter, and note that U is constant on ∂E. Since H is not a single arc, the equation $U(e^{i\phi}) = t$ will have at least four solutions when t is close to zero, and hence the closed maximal set E will consist of at least two disjoint closed arcs, when θ is close to $2\pi\alpha'$.

Return now to the function u. For $x_0 \in (0, \alpha')$ and sufficiently close to α' there is, by the above discussion, a set

$$E_0 = \bigcup_{i=1}^{N} [a_i, b_i]$$

with

$$a_i < b_i < a_{i+1}, \quad i = 1, \ldots, N,$$

$b_N - a_1 < 1$ (note the strict inequality), $|E_0| = 2x_0$, and $2 \leqslant N < \infty$, for which

$$u^*(x) = \int_{E_0} u(t)\,dt.$$

Moreover, $u(a_i) = u(b_i) = t_0$ for some t_0, $u(t) > t_0$ for $t \in E_0$, and $u(t) < t_0$ for t slightly to the left of an a_i or to the right of a b_i.

For $\epsilon > 0$ define sets $E' = E'(\epsilon)$ and $E'' = E''(\epsilon)$ by

$$E' = [a_1 + \epsilon, b_1 - \epsilon] \cup \left(\bigcup_{i=2}^{N} [a_i + \epsilon, b_i + \epsilon] \right),$$

$$E'' = [a_1 - \epsilon, b_1 + \epsilon] \cup \left(\bigcup_{i=2}^{N} [a_i - \epsilon, b_i - \epsilon] \right).$$

When ϵ is small enough these sets satisfy the conclusion of Lemma 2 in Section 2, with $\operatorname{diam} E' \leqslant 1$, $\operatorname{diam} E'' \leqslant 1$. Furthermore,

$$\int_{E''} u(t)\,dt < u^*(x_0 + \epsilon),$$

since the integral strictly increases when E'' is replaced by

$$[a_1 - \epsilon, b_1 + \epsilon] \cup \left(\bigcup_{i=2}^{N} [a_i, b_i] \right).$$

Let $Q(\rho, \phi)$ be defined as in (2.10). The analysis which led to (2.11) shows that, for small enough ρ, $Q(\rho, \phi)$ will be strictly smaller than

$$u^*(x_0 - \rho e^{i\phi}) + u^*(x_0 + \rho e^{i\phi}) \quad \text{for} \quad \phi = 0.$$

By continuity this is true in a neighborhood of $\phi = 0$, and the argument used to establish (2.12) proves (5.2). The proof of the Φ inequality in Theorem 2 is complete.

To prove (1.6), suppose again that K is not translation of K_δ. Fix $x_0 \in \mathbb{R}$ and define for $z \in R(0, 1/2, -B, B) = R$,

$$u_1(z) = \int_{-x}^{x} u(x_0 + t + iy)\,dt, \qquad z = x + iy.$$

Then u_1 is subharmonic in r, and we proved above that $u_1(z) \leqslant u^*(z) < v^*(z)$ there. Applying Hopf's lemma, as in the proof of Proposition 2, Section 3, in $R(0, \alpha, -B, B)$ to $w_1 = u_1 - v^*$ at $z = iy$, we obtain for $|y| < B$,

$$2\left(u(x_0 + iy) - v\left(\frac{1}{2} + iy \right) \right) = \lim_{x \to 0+} \frac{w_1(x + iy)}{x} < 0,$$

which gives (1.6).

6. Proof of Theorem 2. Nonperiodic case

Suppose that $u(z) = u(z, B, K)$ is not a 1-periodic function. For $z \in R(-\infty, \infty, -B, B)$ define, for $z = x + iy$,

$$h(z) = \int_0^1 u(x + t + iy)\, dt, \qquad H(z) = \int_0^1 v(x + t + iy)\, dt,$$

where v is defined by (3.2).

Then H is actually a function of the form $a|y| + b$, since v is 1-periodic, while h is subharmonic in $R(-\infty, \infty, -B, B)$. By Theorem 1 we have $h \leqslant H$ in $R = R(-\infty, \infty, -B, B)$. Thus, either $h < H$ in R or $h \equiv H$. If $h \equiv H$ in R then $h(x) = H(x)$ is constant for $x \in \mathbb{R}$ and hence

$$0 = h'(x) = u(x + 1) - u(x).$$

This violates our non-periodicity assumption. Hence, for all

$$z \in R(-\infty, \infty, 0, B),\ z = x + iy,$$

$$\int_0^1 u(x + t + iy)\, dt < \int_0^1 v(x + t + iy)\, dt. \tag{6.1}$$

Let us suppose that the Φ-inequality statement in Theorem 2 is false. Then from Theorem 1 and (2.5) it follows that there exists $z_0 \in R(0, 1/2, -B, B) \cup \{1/2 + iy\colon -B < y < B\}$ and $E_0 \subset \mathbb{R}$ with $\operatorname{diam} E_0 \leqslant 1$, $|E_0| = 2x_0$, such that

$$\int_{E_0} u(t + iy_0)\, dt = v^*(z_0), \qquad z_0 = x_0 + iy_0. \tag{6.2}$$

The rest of the proof will consist of showing that no such E_0 can exist. Assuming that it does, let $I \subset \mathbb{R}$ be an interval with $E_0 \subset I$ and $|I| = 1$. Define

$$f(t) = u(t + 1 + iy_0) - u(t + iy_0).$$

I claim that $f(t) < 0$ for some $t \in E_0$. Suppose not. Then

$$\int_{E_0} u(t + iy_0 + 1)\, dt \geqslant \int_{E_0} u(t + iy_0)\, dt = v^*(z_0),$$

with strict inequality unless $f \equiv 0$ on E_0. The strict inequality is ruled out by (2.4) and Theorem 1. Hence

$$u(t + iy_0 + 1) = u(t + iy_0) \quad \text{for all} \quad t \in E_0.$$

Now u is harmonic in $(|\operatorname{Im} z| < B) \setminus K$, constant on $y = \pm B$, and satisfies $u(z) = u(\bar{z})$. Moreover, E_0 has positive measure, and it cannot be contained in K if (6.2) holds. One can easily show that under these conditions we must have in fact $u(z) = u(z + 1)$ for every z with $|\operatorname{Im} z| \leqslant B$. Since this contradicts

non-periodicity our claim is established, and thus there exists $t_1 \in E_0 \subset I$ with $u(t_1 + iy_0 + 1) < u(t_1 + iy_0)$.

Similarly, there exists $t_2 \in E_0$ with $u(t_2 + iy_0 - 1) < u(t_2 + iy_0)$. For $\epsilon > 0$ let $R = R(t_2 - 1, t_1 + 1, -\epsilon, \epsilon)$. By continuity, when ϵ is small enough the function $u(z + iy_0)$ satisfies the hypotheses of Theorem 3 in R with $\lambda = 1/2$. Write $u^* = u^*_{1/2}$ for the corresponding *-function, let v, v^* be as in Sections 4, 5, and define

$$w(z) = u^*(z + iy_0) - v^*(z + iy_0).$$

Suppose first that $y_0 \neq 0$. We may assume that $y_0 > 0$, $\epsilon < y_0$, and then, by Theorem 3 and Proposition 1, w is subharmonic in $R_1 = R(0, 1/2, -\epsilon, \epsilon)$ and continuous on the closure. It follows from Theorem 1 that $w \leqslant 0$ in R_1 and, by (6.1), $w(1/2 + iy) < 0$ for $|y| < \epsilon$. Hence $w < 0$ in R. But since $E_0 \subset I \subset [t_2 - 1, t_1 + 1]$, (6.2) implies $w(x_0) = 0$, and (6.1) implies $x_0 < 1/2$, so that $x_0 \in R$. These contradictions show that $y_0 \neq 0$ is impossible.

If $y_0 = 0$ then w is subharmonic in $R \setminus [\alpha, 1/2)$, where α is defined by (3.4). The argument of the preceding paragraph rules out the possibility that $0 < x_0 < \alpha$. Suppose that $x_0 = \alpha$, so that $w(\alpha) = 0$. For our interval I define

$$u^*_I(x) = \sup \int_E u(t)\, dt$$

where the supremum is over E with $|E| = 2x$ and $E \subset I$. Then

$$u^*_I(\alpha) = u^*(\alpha) = v^*(\alpha).$$

Hence

$$0 \leqslant u^*(\alpha) - u^*(x) \leqslant u^*_I(\alpha) - u^*_I(x).$$

Using (2.3) with $I = [A_1, A_2]$ and the δ-density condition (1.1), we deduce that

$$\lim_{x \to \alpha -} \frac{u^*(\alpha) - u^*(x)}{\alpha - x} = 0.$$

It follows that w satisfies all the conditions of Proposition 2 in $R(0, 1/2, -\epsilon, \epsilon)$. By (6.1) we have $w(1/2 + iy) < 0$ for $0 < y < \epsilon$, and Proposition 2 implies that $w(\alpha) < 0$, a contradiction.

The last remaining possibility is that $\alpha < x_0 \leqslant 1/2$, $y_0 = 0$. But the δ-density condition and (2.3) imply that if (6.2) holds for such a z_0 and E_0 then it also holds for some E_1 with $|E_1| = 2\alpha$ and $y_0 = 0$. By the previous paragraph this is impossible. Thus, there are no circumstances under which (6.2) can hold. The Φ inequality for Theorem 2 is now established. The inequality $u(x + iy) < v(1/2 + iy)$ is proved exactly as in the periodic case.

References

[Ba1] Baernstein, A. A generalization of the $\cos \pi\rho$ Theorem, *Trans. Amer. Math. Soc.* **193**(1974), 181-197.

[Ba2] —, Integral means, univalent functions, and circular symmetrization, *Acta Math.* **133**(1974), 139-169.

[Ba3] —, Regularity theorems associated with the spread relation, *J. Analyse Math.* **31**(1977), 76-111.

[Be] Benedicks, M. Positive harmonic functions vanishing on the boundary of certain domains in $\mathbb{R}^n$, *Ark. Mat.* **18**(1980), 53-72.

[ES] Essén, M. and Shea, D. On some questions of uniqueness in the theory of symmetrization, *Ann. Acad. Sci. Fenn. Ser. A.I.* **4**(1978/9), 311-340.

[FT] Fejes-Toth, L. *Regular Figures*, Macmillan, New York, 1964.

[GT] Gilbarg, D. and Trudinger, N. *Elliptic Partial Differential Equations of Second Order*, Second Edition, Springer-Verlag, 1983.

[L] Ya Levin, B. Problem 11.7. in *Linear and Complex Analysis Problem Book*, V. P. Havin, et al., eds., Lecture Notes in Mathematics 1043, Springer-Verlag, 1983.

[M] Minda, D. Bloch constants, *J. Analyse Math.* **43**(1982), 53-84.

[S] Schaeffer, A. C. Entire functions and trigonometric polynomials, *Duke Math. J.* **20**(1953), 77-88.

Albert Baernstein II*
Department of Mathematics
Washington University
St. Louis, MO 63130,
U.S.A.

* This research was supported by grant from the National Science Foundation.

Note

After this paper was typeset I learned that Theorem 1 had already been obtained by A. E. Fryntov, (Dokl. Akad. Nauk USSR, Tom 300 (1988), No. 4, English Translation in Soviet Math. Dokl. **37**(1988), 754-755). His proof is similar to the one given here. Fryntov's article makes no mention of uniqueness questions. I learned also about a result related to Levin's conjecture which had been proved by E. V. Gleizer and A. A. Gol'dberg (Analysis Matematica **11**(1985), 23-28). I thank Professor Havin for telling me about the Gleizer-Gol'dberg result, and Professor Gol'dberg for telling me about Fryntov's.

REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.º 2, 1988

Inversion in Some Algebras of Singular Integral Operators

Michael Christ

To the memory of José Luis Rubio de Francia

Let $\boldsymbol{g}$ be a finite dimensional nilpotent Lie algebra. Suppose that $\boldsymbol{g}$ decomposes as a direct sum $\boldsymbol{g} = \oplus_{\alpha \in (0,\infty)} \boldsymbol{g}_\alpha$ where the Lie brackets of the summands satisfy $[\boldsymbol{g}_\alpha, \boldsymbol{g}_\beta] \subset \boldsymbol{g}_{\alpha+\beta}$ for all $\alpha, \beta \in (0, \infty)$. Let G denote the associated connected, simply connected nilpotent Lie group. Identify elements of $\boldsymbol{g}$ with left-invariant vector fields on G, so that the exponential map $\exp : \boldsymbol{g} \mapsto G$ identifies G with $\mathbb{R}^N$ for some N. Define dilations $\{\delta_r : r > 0\}$ on G by $\delta_r \exp\left(\sum X_\alpha\right) = \exp\left(\sum r^\alpha X_\alpha\right)$ for all $X_\alpha \in \boldsymbol{g}_\alpha$. These are group automorphisms. Denote the homogeneous dimension of G by $d = \sum \alpha \cdot \text{dimension}\,(\boldsymbol{g}_\alpha)$; of course all but finitely many of the $\boldsymbol{g}_\alpha$ have dimension zero. Denote by $\mathcal{S}$ the Schwartz class of functions and by S the unit sphere in G, identified with $\mathbb{R}^N$. Any $x \in G \setminus \{0\}$ may be expressed uniquely as $\delta_r \theta$ for some $r > 0$, $\theta \in S$. Define $|x| = r$, and of course $|0| = 0$. In the exponential coordinates 0 is the group identity element, and $x^{-1} = -x$ for all $x \in G$. All integration over G will be with respect to Haar measure, which is unimodular and agrees with Lebesgue measure in exponential coordinates. Let I denote the identity operator.

A tempered distribution $K \in \mathcal{S}'$ is said to be homogeneous of degree $-d$ if

$$\langle K, f^r \rangle = \langle K, f \rangle$$

for all $f \in \mathcal{S}$, $r > 0$, where $f^r(x)$ denotes $f(\delta_r x)$. To any distribution ψ acting on test functions on S and annihilating constants we may assign a distribution $K \in \mathcal{S}'(G)$ as follows: For $r > 0$ and $f \in \mathcal{S}(G)$ write $f_r(\theta) = f(\delta_r \theta)$. Then define

$$\langle K, f \rangle = \int_0^\infty \langle \psi, f_r \rangle \frac{dr}{r}.$$

The integral converges absolutely because $f \in \mathcal{S}$ and $\langle \psi, 1 \rangle = 0$. We denote by PV the class of all such distributions on G; they are homogeneous of degree $-d$. In fact any $K \in \mathcal{S}'$ homogeneous of degree $-d$ is necessarily of the form $K = a\delta + K'$ for some $K' \in PV$ and $a \in \mathbb{C}$, where δ denotes the Dirac mass at $0 \in G$. See [C2, Lemma 2.4] for a proof. The decomposition is unique.

For $q \in (1, \infty)$ let A_q denote the set of all operators $T: \mathcal{S}(G) \mapsto \mathcal{S}'(G)$ of the form $Tf = f * K = af + f * K'$ where $K' \in PV$ and K' restricted to S belongs to L^q. Define $|T|_q = |a| + \|K'\|_{L^q(S)}$.

Proposition 1. *For all $q \in (1, \infty)$, any $T \in A_q$ extends to an operator bounded on $L^p(G)$, for all $p \in (1, \infty)$.*

Proposition 2. *A_q is a Banach algebra for all $q \in (1, \infty)$.*

By virtue of Proposition 1, $S \circ T$ is defined for all $S, T \in A_q$, as a bounded operator on L^2. What must be proved here is that

$$S \circ T \in A_q \quad \text{and} \quad |S \circ T|_q \leqslant C_q |S|_q |T|_q.$$

(Consequently there exists an equivalent norm $|\cdot|'_q$ on A_q, such that $|I|'_q = 1$ and $|S \circ T|'_q \leqslant |S|'_q |T|'_q$ for all $S, T \in A_q$, cf. [K, p. 197].)

Let $\mathcal{B}$ denote the algebra of all bounded linear operators on $L^2(G)$, and let $|\cdot|_{\mathcal{B}}$ be the operator norm. For any algebra A and element $T \in A$, let

$$\operatorname{spec}_A(T) = \{\lambda \in \mathbb{C}: T - \lambda I \text{ is not invertible in } A\}.$$

The point of this article is

Theorem 3. *For all $q \in (1, \infty)$, for all $T \in A_q$, $\operatorname{spec}_{A_q}(T) = \operatorname{spec}_{\mathcal{B}}(T)$.*

Proposition 1 implies one inclusion, $\operatorname{spec}_{\mathcal{B}}(T) \subset \operatorname{spec}_{A_q}(T)$.

In the Euclidean case, that is when $G = \mathbb{R}^N$ with the Euclidean group law and $\{\delta_r\}$ are the usual dilations, these results were obtained by Calderón and Zygmund [CZ1], [CZ2]. This class of operators was further studied by Duoandikoetxea and Rubio de Francia [DR]. In the nilpotent case Proposition 1 is a corollary of the main theorem of [C1] and the Calderón-Zygmund method of rotations if K is odd, and in general is a direct consequence of the method of proof in that paper. The question was asked of us by J. Duoandikoetxea. Proposition 2 is a corollary of the proof of Proposition 1.

Our main result, Theorem 3, answers a question posed by A. Carbery. In the Euclidean case A_q is a commutative Banach algebra, the Fourier transform sets up a natural identification of its maximal ideal space with the unit sphere in $\mathbb{R}^N$, and the result follows from Plancherel's theorem [CZ2]. However in the nilpotent case commutativity is lost.

We were led to this issue in the nilpotent case in studying the regularity of solutions of $\bar{\partial}_b$ [C4], a partial differential operator which arises in complex analysis in several variables. A principal step in that work was an analogue of Theorem 3, in which the condition $K' \in L^q(S)$ was replaced by the stronger condition $K' \in L^q_\beta$ in an annular neighborhood of S, where L^q_β was a certain type of Sobolev space with exponent q and order of differentiability $\beta > 0$. The set of all associated operators T is still an algebra, and we showed that the spectrum in that algebra agrees with the spectrum in $\mathcal{B}$. Thus Theorem 3 is the limiting case $\beta = 0$ of the earlier result, and our first proof of it was by the same technique. The purpose of the present article is instead to present a considerably simpler proof which emphasizes the Banach algebraic point of view. This argument could also be used to give a simpler proof in the case $\beta > 0$, but we shall not give the details here.

The proof relies on estimation of the spectral radius in A_q. L. Carleson has pointed out that it is related to Beurling's proof [B] of the Wiener Tauberian Theorem. I am indebted to A. Carbery, L. Carleson, and A. McIntosh for helpful conversations concerning this work.

To begin let $A \subset B$ be two Banach algebras with different norms. Suppose $|x|_B \leqslant C|x|_A$ for all $x \in A$, and that $|xy|_B \leqslant |x|_B |y|_B$ for all $x, y \in B$. Define the spectral radius of an element of A by

$$\rho_A(x) = \lim_{n \to \infty} |x^n|_A^{1/n}.$$

The limit always exists [K, p. 208] and is finite.

Lemma 4. *Suppose that there exist $C < \infty$ and $\theta \in (0, 1]$ such that for all $x, y \in A$,*

$$|xy|_A \leqslant C(|x|_A |y|_A^{1-\theta} |y|_B^{\theta} + |y|_A |x|_A^{1-\theta} |x|_B^{\theta}).$$

Then $\rho_A(x) \leqslant |x|_B$ for all $x \in A$.

PROOF. The hypothesis gives

$$|x^{2k}|_A^{1/2k} \leqslant (2C)^{1/2k}(|x^k|_A^{(2-\theta)/2k} |x^k|_B^{\theta/2k}).$$

for all $x \in A, k > 0$. Passing to the limit gives

$$\rho_A(x) \leqslant \rho_A(x)^{1-\theta/2} |x|_B^{\theta/2}.$$

This simplification due to Carleson replaces the author's slightly longer proof. Now suppose that $B = \mathcal{B}$, and denote by x^* the adjoint of $x \in \mathcal{B}$.

Lemma 5. *Suppose that $A \subset \mathcal{B}$ satisfy the hypothesis of Lemma 4 and that A is closed under adjoints. Then*

$$\operatorname{spec}_A(x) = \operatorname{spec}_{\mathcal{B}}(x)$$

for all $x \in A$.

PROOF. What must be demonstrated is that if $x \in A$ is invertible in $\mathcal{B}$, then it is also invertible in A. It suffices to invert both x^*x and xx^* in A, for then x will have both a left and a right inverse. If $\epsilon > 0$ is chosen sufficiently small, then $|I - \epsilon xx^*|_{\mathcal{B}} < 1$ because x is invertible in $\mathcal{B}$. Set $y = I - \epsilon xx^* \in A$. Then $\rho_A(y) \leqslant |y|_{\mathcal{B}} < 1$ by Lemma 4, so $I - y$ may be inverted in A by summing the series $I + y + y^2 + \cdots$. But $I - y = \epsilon xx^*$. The reasoning for x^*x is the same. This argument was pointed out to us by J. L. Journé.

Introduce an auxiliary function $\zeta \in C_0^\infty(G)$, with $\zeta(x)$ a function of $|x|$ alone, identically one in a neighborhood of 0 and supported in a small neighborhood of 0. Denote by $\|\cdot\|_{q,q}$ the operator norm on $L^q(G)$.

Lemma 6. *Suppose that $T \in \mathcal{B}$ is of the form $Tf = f * K$ where $K = a\delta + K'$, $K' \in PV$. Then $|a| \leqslant C\|T\|_{\mathcal{B}}$.*

PROOF. Fix $\phi \in C_0^\infty(G)$, identically one in a neighborhood of 0. If the support of ζ is chosen to be sufficiently small then $\phi * (\zeta K') \equiv 0$ in some neighborhood of 0, since $\zeta K'$ annihilates constants by the definition of PV and the fact that ζ is itself radial. But $(1 - \zeta)K' = (1 - \zeta)K$ and the L^2 operator norm of $f \mapsto f * (1 - \zeta)K$ is majorized by $C_\zeta \|T\|_{\mathcal{B}}$ [C2, Lemma 2.10]. In a neighborhood of 0, $T\phi = a\phi + \phi * ((1 - \zeta)K)$. So $\|T(\phi)\|_{L^2} \geqslant C_1|a| - C_2|T|_{\mathcal{B}}$.

Assume for now the validity of Proposition 1. To complete the proof of Theorem 3 we show that $A = A_q$, $B = \mathcal{B}$ satisfy the hypothesis of Lemma 4. Suppose $q \neq 2$. Let $S, T \in A_q$ be given by convolution on the right with $K = a\delta + K'$ and by $L = b\delta + L'$ respectively, where $a, b \in \mathbb{C}$ and $K', L' \in PV$. Set $K_\infty = (1 - \zeta)K = (1 - \zeta)K'$ and similarly define L_∞. Since $S \circ T$ is well-defined as a bounded operator on $L^2(G)$, $L * K$ is a well-defined element of $\mathcal{S}'$, manifestly homogeneous of degree $-d$. Furthermore

$$L * K = (\zeta L) * (\zeta K) + [(1 - \zeta)L] * K + L * [(1 - \zeta)K] + [(1 - \zeta)L] * [(1 - \zeta)K]. \tag{1}$$

Since $L * K$ is homogeneous of degree $-d$, its restriction to S will be in L^q if and only if its restriction to $\{x \in G: |x| \geqslant 1\}$ is in L^q with respect to Haar measure on G. If we choose ζ to be supported sufficiently near 0, then $(\zeta L) * (\zeta K)$ will be supported in $\{|x| < 1\}$, so may be neglected.

For the second term of (1), $\|(1-\zeta)L\|_{L^q(G)} \leqslant C|T|_q$ so that

$$\begin{aligned}\|[(1-\zeta)L]*K\|_{L^q} &\leqslant \|(1-\zeta)L\|_q \|S\|_{q,q} \\ &\leqslant C|T|_q \|S\|_{q,q} \\ &\leqslant C|T|_q \|S\|_{r,r}^{1-\theta} \|S\|_{2,2}^{\theta} \\ &\leqslant C|T|_q |S|_q^{1-\theta} |S|_{\mathcal{B}}^{\theta}\end{aligned}$$

where $r \in (1, \infty)$ is chosen so that q lies in the open interval with endpoints $r, 2$. The second-to-last inequality then follows from the Riesz-Thörin interpolation theorem, and the last from Proposition 1.

The fourth term in (1) may be estimated in the same way since again by Lemma 2.10 of [C2], the L^q operator norm of $f \mapsto f*[(1-\zeta)K]$ is majorized by $C_\zeta \|S\|_{q,q}$. As for the third, it will be no loss of generality to assume that K' is real, and ζ may be chosen to be real also. Then $L*[(1-\zeta)K](-x)$ is the complex conjugate of $T^*(h)(x)$ where $h(y) = (1-\zeta(-y))K'(-y)$ and T^* denotes the adjoint of T. Inversion about 0 leaves the $L^q(G)$ norm invariant, so the reasoning of the last paragraph applies.

For any $S, T \in A_q$, $S \circ T$ is given by convolution with a distribution homogeneous of degree $-d$, which must have the form $a\delta + K'$ for some $a \in \mathbb{C}$, $K' \in PV$. The argument just completed shows that $K' \in L^q(S)$ with bound $O(|S|_q|T|_q)$, so Proposition 2 is also proved.

To prove Proposition 1 suppose that $K \in PV$ restricts to an L^q function on S. Introduce an auxiliary function $\eta \in C_0^\infty(G)$, depending only on $|x|$ and vanishing identically in some neighborhood of 0. Thus ηK annihilates constants. Choose it so that

$$\int_0^\infty \eta(\delta_s(x))\, ds/s \equiv 1$$

on $G\setminus\{0\}$. Set $K_t(x) = \eta(\delta_t^{-1}x)K(x)$. Then $\int_0^\infty K_t\, dt/t$ converges in $\mathcal{S}'$ and equals K. Set $T_t f = f*K_t$ for all $f \in \mathcal{S}$.

Lemma 7. *There exist $C < \infty$ and $\epsilon > 0$ such that for all $s, t > 0$,*

$$\|T_t T_s^*\|_{2,2} + \|T_t^* T_s\|_{2,2} \leqslant C \min(s/t, t/s)^{\epsilon}.$$

L^2 boundedness of $f \mapsto f*K$ then follows from the Cotlar-Stein almost-orthogonality lemma.

SKETCH OF PROOF. Consider $T_s T_t^*$ in the case when $t \leqslant s$; the case $s < t$ as well as $T_s^* T_t$ may be handled by the same method. Rescaling via the dilation δ_s then reduces matters to the case $s = 1$. Since $K_t \in L^1$ uniformly in t, the issue

is the factor of t^ϵ as $t \to 0$. Let $\tilde{K}_s(x)$ be the complex conjugate of $K_s(-x)$. As in [C1] it suffices to prove that

$$\|\tilde{K}_t * (K_1 * \tilde{K}_1 * K_1 * \tilde{K}_1 * \cdots)\|_{L^q} \leqslant Ct^\epsilon \tag{2}$$

where the convolution product has a total of N factors of K_1, $\tilde{K}_1$. This holds with $\epsilon = 0$ when $q = 1$, so by interpolation it will suffice to prove it for some $\epsilon > 0$ when $q = \infty$.

Consider N points $\theta_1, \ldots, \theta_N \in S$, and regard S as a subset of the Lie algebra $\boldsymbol{g}$. Write $s = (s_1, \ldots, s_N) \in \mathbb{R}^N$ and $\theta = (\theta_1, \ldots, \theta_N) \in S^N$. Consider the map from $\mathbb{R}^N \times S^N$ to G given by $F(s, \theta) = \exp(s_1\theta_1)\exp(s_2\theta_2)\cdots\exp(s_N\theta_N)$, where the product is taken with respect to the group structure on G. Denote by $J(s, \theta)$ the determinant of the $N \times N$ matrix of first partial derivatives of F with respect to s, a real analytic function of both variables. Whenever $\theta_1, \ldots, \theta_N$ are linearly independent elements of $\boldsymbol{g}$, $J(0, \theta) \neq 0$. Therefore J is real analytic and not identically zero. Now $K_1 * \tilde{K}_1 * \cdots * (x)\,dx$ is the push-forward by F of the measure $\prod_{j=1}^{N} (\Omega_j(\theta_j))\phi(s_j)\,d\theta_j\,ds_j$, for a certain function $\phi \in C_0^\infty(\mathbb{R}^+)$, where Ω_j denotes the restriction of either K_1 or $\tilde{K}_1$ to S, with the choice alternating in j. Since $\tilde{K}_t$ annihilates constants and is supported on $\{x \in G: |x| \leqslant Ct\}$, (2) now follows directly from a combination of the proofs of [C1, Lemma 5.4] and [C3, Lemma 2.2]. See also [RS] and [C5] for arguments of this type.

Inequality (2) remains valid, with the same proof, if $\tilde{K}_t$ is replaced by any measure supported on $\{x \in G: |x| \leqslant Ct\}$ which annihilates constants; the constant on the right-hand side of (2) depends only on the total mass of the measure. Therefore the boundedness of any element of A_q on L^p, for any $p, q \in (1, \infty)$, follows from the methods of Duoandikoetxea and Rubio de Francia [DR] and the author [C1].

References

[B] Beurling, A. Sur les intégrales de Fourier absolument convergentes et leur applications à une transformation fonctionelle, in *Nionde skandinaviska Matematikerkongressen* (1938), Mercators Tryckeri, Helsinki, 1939.

[CZ1] Calderón, A. P. and Zygmund, A. On singular integral operators, *Amer. J. Math.* **78**(1956), 289-309.

[CZ2] —, Algebras of certain singular integral operators, *Amer. J. Math.* **78**(1956), 310-320.

[C1] Christ, M. Hilbert transforms along curves, I. Nilpotent groups. *Annals of Math.* **122**(1985), 576-596.

[C2] —, On the regularity of inverses of singular integral operators, *Duke Math. J.*, **57**(1988), 459-484.

[C3] —, Weak type (1,1) bounds for rough operators, *Annals of Math.*, **128**(1988), 19-42.

[C4] —, Regularity properties of the $\bar{\partial}_b$ equation on weakly pseudoconvex *CR* manifolds of dimension three, *Journal Amer. Math. Soc.*, **1**(1988), 587-646.

[C5] —, The strong maximal fucntion on a nilpotent group, *Trans. Amer. Math. Soc.*, to appear.

[DR] Duoandikoetxea, J. and Rubio de Francia, J. L. Maximal and singular integral operators via Fourier transform estimates, *Invent. Math.* **84**(1986), 541-561.

[K] Katznelson, Y. *An Introduction to Harmonic Analysis*, Dover Publications, New York, 1976.

[RS] Ricci, F. and Stein, E. M. Harmonic analysis on nilpotent groups and singular integrals. II: Singular kernels supported on submanifolds, preprint.

M. Christ*
University of California
Los Ángeles, Ca. 90024,
U.S.A.

* Mathematical Sciences Research Institute. Berkeley, Calfornia.

REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.° 2, 1988

Fatou Theorems for Some Nonlinear Elliptic Equations

E. Fabes, N. Garofalo, S. Marín-Malave and S. Salsa

Introduction

A classical Fatou theorem states:

For any given nonnegative harmonic function $u(x)$ defined in the unit ball B of $\mathbb{R}^n$ there exists a set of boundary points E_u, with surface measure equal to the surface measure of the entire boundary, such that for each point $P \in E_u$,

$$\lim_{\substack{x \to P \\ x \in B \cap \Gamma_P}} u(x) \text{ exists},$$

where Γ_P is any finite cone with vertex P and interior contained in B. This behavior of u is often described by saying «u has a nontangential limit at almost every (with respect to surface measure) boundary point».

The above Fatou theorem has been generalized in many directions. It has been extended to nonnegative harmonic functions defined in nonsmooth domains ([6], [7]) and to nonnegative solutions of general second order elliptic equations with nonsmooth coefficients ([1], [2]). The statement of the Fatou theorem in these situations remains the same except that surface measure must be replaced by the harmonic measure associated with the governing partial differential operator.

In this work we wish to extend the Fatou theory to several classes of second order nonlinear elliptic operators. The primary motivating equation for us is the p-Laplacian,

$$\operatorname{div}(|\nabla u|^{p-2}\nabla u) = 0 \qquad (1 < p < \infty).$$

Our Fatou theorem in this case takes the following form:

Let D be a bounded Lipschitz domain in $\mathbb{R}^n$ and suppose u is a nonnegative weak solution of $\operatorname{div}(|\nabla u|^{p-2}\nabla u) = 0$ in D. The set of boundary points E_u at which u has a nontangential limit has Hausdorf dimension $\geqslant \beta$, a positive number depending only on p, n, and the Lipschitz character of D. Recent examples of T. Wolff [15] and J. Lewis [11] show that even when the domain D is smooth the set E_u could have surface measure zero. The positive result at least guarantees that E_u is somewhat far from being empty.

The above Fatou theorem for the p-Laplacian in smooth domains of $\mathbb{R}^n$ and with $1 < p < 3 + (2/(n-2))$ was first proved by Manfredi and Weitsman in [12]. Several ideas in their paper were very useful and we would like to thank the authors for sharing them with us early on in their work. Also, while we were preparing the present paper Manfredi and Weitsman communicated to us a proof, different than the one presented here, of the above theorem in smooth domains of $\mathbb{R}^n$ and for any $1 < p < \infty$.

The same type of Fatou theorem remains valid for nonnegative solutions of other classes of nonlinear elliptic equations. We illustrate this in the case of a generalization of the p-Laplacian in the form

$$\operatorname{div}\left(\frac{f'(|\nabla u|)}{|\nabla u|}\nabla u\right) = 0$$

and in the case of solutions of a fully nonlinear equation of the type

$$F(D^2u, x) = 0.$$

In both situations the Hausdorf dimension of the set of nontangential limits is estimated from below by a positive number depending only on the structure constants (properties of f and F), dimension n, and the Lipschitz character of D. (See Sections II.2 and II.3.) Since the methods we employ for the general classes of equations are minor modifications of those used for the p-Laplacian, $\operatorname{div}(|\nabla u|^{p-2}\nabla u) = 0$, we have decided to give a detailed proof of the Fatou theorem only in this special case (Section II.1). The extension of the result to the other classes is discussed briefly in the final two sections of the paper.

Our technique for establishing a Fatou theory in the nonlinear setting is itself linear in nature. It frequently occurs that results for linear equations if achieved in sufficient generality can be successfully applied to nonlinear situations. The present work is one more example of this phenomenum. In order to find a Fatou theorem in the nonlinear setting we first develop a potential theory for linear uniformly elliptic equations in nondivergence form and for the corresponding adjoint equations. The coefficients of the linear operator are assumed to be smooth, but, more importantly for applications, all the

constants involved in the basic estimates do *not* depend in any quantitative manner on the smoothness of the coefficients. These constants depend only on dimension, ellipticity, L^∞-bounds of the coefficients, and the Lipschitz character of the domain. In this respect our work is a continuation and refinement of Bauman's paper [1].

The primary results in the linear theory are what we refer to as Comparison Theorems for nonnegative solutions of either a nondivergence form elliptic equation or of the associated adjoint equation. In brief terms these theorems state that two nonnegative solutions of the equation (or adjoint equation) which vanish on an open portion of the boundary must vanish there at the same rate. For a precise statement of these results see Theorems I.2.2 and I.3.7.

The paper is divided into two parts. Part I is devoted to the potential theory for second order linear nonvariational elliptic operators and their adjoints, always aiming towards establishing the above mentioned Comparison Theorems. Part II applies the linear theory to prove the previously mentioned Fatou theorems for the p-Laplacian, its generalization, and for some fully nonlinear elliptic equations. Part II may be read somewhat independently of Part I, at least from the point of view of understanding how the linear theory enters in establishing the positive Hausdorf dimension of the set of boundary points at which the solution of the nonlinear problem has a nontangential limit.

I. POTENTIAL THEORY FOR NONDIVERGENCE FORM OPERATORS AND THEIR ADJOINTS

Before we begin the main body of this paper we would like to recall the basic definitions and introduce the primary notation which will be extensively used throughout the work.

Definition A. *A bounded domain D of $\mathbb{R}^n$ is called a Lipschitz domain if*

(i) *for each $Q \in \partial D$ there exists a coordinate system $(x', x_n) \in \mathbb{R}^{n-1} \times \mathbb{R}$, a number r_0 and a function $\varphi\colon \mathbb{R}^{n-1} \to \mathbb{R}$ satisfying $\|\nabla\varphi\|_{L^\infty(\mathbb{R}^{n-1})} \leqslant m$ such that*

(ii) $$B_{r_0}(Q) \cap D = \{(x', x_n)\colon x_n > \varphi(x')\} \cap B_{r_0}(Q),$$

($B_{r_0}(Q)$ is the ball in $\mathbb{R}^n$ with center Q and radius r_0) and

$$B_{r_0}(Q) \cap \partial D = \{(x', \varphi(x')\} \cap B_{r_0}(Q).$$

We may assume the numbers m and r_0 are the same for each $Q \in \partial D$ and we will say that these numbers determine the Lipschitz character of D.

Given the coordinate system about $Q \equiv (Q', Q_n) \in \partial D$, *we define for* $r \leqslant r_0$,

$$T_r(Q) = \{(x', x_n): |x' - Q'| < r, |x_n - Q_n| < mr\}$$
$$\Delta_r(Q) = T_r(Q) \cap \partial D, \qquad \psi_r(Q) = T_r(Q) \cap D,$$

and $A_r(Q) = (Q', Q_n + mr)$. *When the point* Q *is understood and unimportant in the discussion we will use the notation* T_r, Δ_r, ψ_r *and* A_r.

In Section 1 we consider elliptic operators of the form

$$Lu(x) = \sum_{i,j=1}^{n} a_{ij}(x) D^2_{x_i x_j} u(x) \qquad (x \in \mathbb{R}^n).$$

We assume the coefficients are smooth and the matrix $a(x) = (a_{ij}(x))$ is bounded, symmetric, and positive definite, uniformly in x, *i.e.* there exist positive numbers λ and Λ such that

$$\lambda|\xi|^2 \leqslant \sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \leqslant \Lambda|\xi|^2$$

for all x and ξ in $\mathbb{R}^n$. We wish again to emphasize that the assumption of smoothness of $a_{ij}(x)$ is only a qualitative one. In our estimates the dependence of the constants on the coefficients will only be in terms of the ellipticity parameters λ and Λ and the dimension n.

Corresponding to the operator L we have the adjoint operator L^* defined by

$$L^*v(y) = \sum_{i,j=1}^{n} D^2_{y_i y_j}(a_{ij}(y)v(y)).$$

For a given Lipschitz domain D we let $g_D(x, y) = g_{D,L}(x, y)$ be the Green's function corresponding to the operator L and domain D. In particular

$$L(g_D(\cdot, y))(x) = 0 \quad \text{for} \quad x \in D \setminus \{y\},$$
$$L^*(g_D(x, \cdot))(y) = 0 \quad \text{for} \quad y \in D \setminus \{x\}, \quad \text{and}$$
$$g(Q, y) = 0 = g(x, Q) \quad \text{for} \quad Q \in \partial D, \quad x \in D \quad \text{and} \quad y \in D.$$

I.1. The Notion of a Normalized Adjoint Solution and the Boundary Harnack Principle

Definition B. *For a ball* $\mathcal{B}$ *let* $t\mathcal{B}$ *denote the ball concentric with* $\mathcal{B}$ *and radius equal to* t *times the radius of* $\mathcal{B}$. *Assume* $\frac{1}{4}\mathcal{B} \supset \bar{D}$ *and fix a point* $P \in \frac{3}{4}\mathcal{B} \setminus \frac{1}{2}\mathcal{B}$. *A normalized adjoint solution for* L^* *and* D (*briefly* n.a.s.) *is any function* $\tilde{v}$ *of the form*

$$\tilde{v}(y) = \frac{v(y)}{g_{\mathcal{B}}(P, y)}$$

where v is a solution of the adjoint equation $L^*v = 0$ in D. (We shall see that for our purposes there is no quantitative dependence of $\tilde{v}$ on $\mathcal{B}$ and $g_{\mathcal{B}}(P, y)$. Therefore we do not highlight them in the name.)

The notion of a normalized adjoint solution was used extensively in [1]. The main purpose of Part I is to refine the results there by dropping the dependence on the modulus of continuity of the coefficients and to extend the results to include a Comparison Theorem for two nonnegative solutions of the adjoint equation vanishing on a portion of the boundary. (See Theorem I.3.7.) We begin by recalling some basic facts from [1] concerning these functions.

Theorem I.1.1. (Interior Harnack Principle for n.a.s.) *Suppose $\tilde{v}$ is a nonnegative* n.a.s. *in a ball B_{2R} of radius $2R$. There exists a constant C depending only on n, λ and Λ such that for all $s \leqslant R$*

$$\sup\{\tilde{v}(y): y \in B_s\} \leqslant C \inf\{\tilde{v}(y): y \in B_s\}.$$

(B_s and B_{2R} are assumed concentric.)

Theorem I.1.2. (The Dirichlet Problem for n.a.s.) *Given $f \in C(\partial D)$ there exists a unique normalized adjoint solution, $\tilde{v}$, such that $\tilde{v} \in C(\bar{D})$ and $\tilde{v} = f$ on ∂D. Moreover the maximum and minimum values of $\tilde{v}$ occur on ∂D.*

Definition C. *Under the conditions of Theorem* I.1.2. *for fixed $y \in D$, the map $f \mapsto \tilde{v}(y)$ is a positive continuous linear functional on $C(\partial D)$. So by the Riesz Representation Theorem*

$$v(y) = \int_{\partial D} f(Q)\tilde{\omega}^y(dQ)$$

where $\tilde{\omega}^y$ is a regular Borel measure on ∂D. We will call $\tilde{\omega}^y$ the normalized adjoint measure at y (corresponding to L^* and D).

Notice that when D is a smooth domain

$$\tilde{\omega}^y(dQ) = \frac{\partial}{\partial v_Q} g(Q, y) g_{\mathcal{B}}(P, Q)\sigma(dQ)/g_{\mathcal{B}}(P, y)$$

where g denotes the Green's function for L and D, v_Q is the inward conormal to ∂D at Q. ($v_Q = a(Q)(N_Q)$ with $a = (a_{ij})$ and N_Q is the unit inward normal to ∂D at Q),

$$\frac{\partial}{\partial v_Q} g(Q,y) = \nabla_x g(x,y)|_{x=Q} \cdot v_Q,$$

and $\sigma(dQ)$ is surface measure. The normalized adjoint measure $\tilde{\omega}^y$ will play an important role in the study of normalized adjoint solutions.

Lemma I.1.3. *Let $\tilde{\omega}^y_{2r}$ denote the normalized adjoint measure corresponding to L^* and a ball B_{2r}. For $Q \in \partial B_{2r}$ and $0 < \delta < 1$ set $\Delta_{\delta r} = \partial B_{2r} \cap B_{\delta r}(Q)$. There exists a positive constant c depending only on the ellipticity parameters, n, and δ such that*

$$\inf_{y \in B_r} \tilde{\omega}^y_{2r}(\Delta_{\delta r}) \geqslant c$$

(B_r is assumed concentric with B_{2r}.)

PROOF. By a translation we may assume B_{2r} is centered at the origin and by a dilation we may assume $r = 1$,

$$\tilde{\omega}^y_2(\Delta_\delta) = \int_{\Delta_\delta} \frac{\partial}{\partial v_Q} g_2(Q,y) g_{\mathcal{B}}(P,Q)\sigma(dQ)/g_{\mathcal{B}}(P,y)$$

where $g_2(x,y)$ is the Green's function corresponding to B_2 and elliptic operator L. Fix a point $A \in \partial B_{3/2}$. All the succeeding constants in this lemma will depend at most on the dimension n, the parameters of ellipticity, λ and Λ, for L, and δ.

From Hopf's Lemma and Harnack's Inequality [4], [9] there exists $c_1 > 0$ such that for all $y \in B_1$, and $Q \in \partial B_2$,

$$\frac{\partial}{\partial v_Q} g_2(Q,y) \geqslant c_1 g_2(A,y).$$

(See [1, Lemma 4.3, p. 166].) Hence, for $y \in B_1$

$$\tilde{\omega}^y_2(\Delta_\delta) \geqslant c_1 g_2(A,y) \int_{\Delta_\delta} g_{\mathcal{B}}(P,Q)\sigma(dQ)/g_{\mathcal{B}}(P,y).$$

Let $g_3(x,y)$ denote the Green's function for B_3 and L. Then Harnack's property for normalized adjoint solutions (Theorem I.1.1.) implies

$$\frac{g_{\mathcal{B}}(P,Q)}{g_3(A,Q)} \geqslant c_2 \frac{g_{\mathcal{B}}(P,y)}{g_3(A,y)}$$

for all $Q \in \partial B_2$ and $y \in B_1$. Hence

$$\tilde{\omega}^y_2(\Delta_\delta) \geqslant c_1 c_2 \frac{g_2(A,y)}{g_3(A,y)} \int_{\Delta_\delta} g_3(A,Q)\sigma(dQ).$$

Once again using Harnack's inequality (for n.a.s.) we have for $y \in B_1$.

$$\frac{g_2(A,y)}{g_3(A,y)} \geqslant c_3 \frac{\int_{B_1} g_2(A,z)\,dz}{\int_{B_1} g_3(A,z)\,dz} \geqslant c_4 > 0.$$

We are now left to show

Lemma I.1.4. *Let $g_3(x,y)$ denote the Green's function for L and B_3. For $Q \in \partial B_2$ and $0 < \delta < 1$, set $\alpha_\delta = \partial B_2 \cap B_\delta(Q)$. There exists a positive constant c depending only on n, λ, Λ, and δ such that for all $A \in \partial B_{3/2}$*

$$\int_{\alpha_\delta} g_3(A,Q)\sigma(dQ) \geqslant c > 0.$$

PROOF. Let $D \subset B_2 \setminus B_{7/4}$ be a smooth domain such that $\alpha_\delta \subset \partial D$ and D contains a ball B' whose radius depends only on δ. Now let $D' \subset B_3$ be another smooth domain containing D such that $\alpha_\delta \subset D'$ and $\partial D' \cap \partial D = \partial D \setminus \alpha_\delta$. We pick a point $A' \in D' \setminus \bar{D}$ so that the positive constants in the following chain of inequalities depend only on the parameters of ellipticity, λ and Λ, the dimension n, and δ:

$$\int_{\alpha_\delta} g_3(A,Q)\sigma(dQ) \geqslant c_1 \int_{\alpha_\delta} g_3(A',Q)\sigma(dQ) \geqslant c_2 \int_{\alpha_\delta} g_{D'}(A',Q)\sigma(dQ).$$

In the final inequality above $g_{D'}$ denotes the Green's function for D' and L and the inequality results from the fact $g_{D'} \leqslant g_3$ since $D' \subset B_3$.

Since $\partial D \setminus \alpha_\delta = \partial D \cap \partial D'$, $g_{D'}(A',Q) = 0$ for $Q \in \partial D \setminus \alpha_\delta$. Therefore

$$\int_{\alpha_\delta} g_{D'}(A',Q)\sigma(dQ) = \int_{\partial D} g_{D'}(A',Q)\sigma(dQ).$$

The function $g_{D'}(A',y)$, as a function of y, satisfies

$$L^*(g_{D'}(A',\cdot))(y) = 0 \quad \text{for} \quad y \in D.$$

The function

$$W(x) = -\int_{B'} g_D(x,y)\,dy$$

satisfies $LW = \chi_{B'}$, the characteristic function of B', and $W|_{\partial D} = 0$. Hence, an integration by parts gives

$$\int_{B'} g_{D'}(A',y)\,dy = \int_D g_{D'}(A',y) LW(y)\,dy = \int_{\partial D} g_{D'}(A',Q) \frac{\partial}{\partial v_Q} W(Q)\sigma(dQ).$$

Again from [1, Lemma 4.3, p. 166] there exists $C > 0$ such that

$$\frac{\partial W}{\partial v_Q}(Q) \leqslant C \sup_D \int_{B'} g_D(x,y)\,dy \leqslant \tilde{C},$$

$\tilde{C}$ depending only on λ, Λ, n, and δ. We now have

$$\int_{\alpha_\delta} g_{D'}(A', Q)\sigma(dQ) \geqslant c \int_{B'} g_{D'}(A', y)\, dy,$$

and using a maximum principle argument the last integral is bounded below by a positive constant depending only λ, Λ, n, and δ. (See [4, Proof of Lemma 3.3].) This concludes the proof of Lemma I.1.4 and so also the proof of Lemma I.1.3.

As a consequence of Lemma I.1.3 we obtain the Hölder continuity of a nonnegative normalized adjoint solution at the open parts of the boundary where it vanishes.

Theorem I.1.5. *Let D be a Lipschitz domain in $\mathbb{R}^n$ with constants r_0 and m determining the Lipschitz character of D. Let*

$$L = \sum_{i,j=1}^{n} a_{ij}(x) D^2_{x_i x_j}$$

be a uniformly elliptic operator with parameters of ellipticity λ and Λ. Fix $Q \in \partial D$, $r \leqslant r_0/2$, and assume $\tilde{v}$ is a nonnegative normalized adjoint solution for L^* and $\psi_{2r} = \psi_{2r}(Q)$ which continuously vanishes on $\Delta_{2r} = \Delta_{2r}(Q)$. There exists a Θ, $0 < \Theta < 1$ depending only on λ, Λ, n, and m such that

$$\sup_{\psi_r} \tilde{v} \leqslant \Theta \sup_{\psi_{2r}} \tilde{v}.$$

As a consequence there exist positive constants C and α depending only on the above parameters such that for all $y \in \psi_{2r}$,

$$\tilde{v}(y) \leqslant c\left(\frac{|y - Q|}{r}\right)^{\alpha} \sup_{\psi_{2r}} \tilde{v}.$$

PROOF. We may assume $\sup_{\psi_{2r}} \tilde{v} = 1$. Let $\tilde{\omega}^y_{2r}$ denote the normalized adjoint measure for L^* and $B_{2r} = B_{2r}(Q)$. We can find a positive number δ, depending on m, and a point $\tilde{Q} \in \partial B_{2r}$ such that

$$\alpha_{\delta r} = B_{\delta r}(\tilde{Q}) \cap \partial B_{2r} \subset \partial B_{2r} \setminus \bar{D}.$$

From the maximum principle for n.a.s. (Theorem I.1.2)

$$1 - \tilde{v}(y) \geqslant \tilde{\omega}^y_{2r}(\alpha_{\delta r}).$$

Therefore, using Lemma I.1.4.

$$\inf_{\psi_r} (1 - \tilde{v}(y)) \geqslant c > 0$$

with c depending only on λ, Λ, n, and m. Then

$$\sup_{\psi_r} v \leqslant 1 - c = \Theta.$$

To conclude the Hölder continuity we observe that the last inequality above implies: for all $y \in \psi_2 - k_r(Q)$,

$$\tilde{v}(y) \leqslant \Theta^{k+1} \sup_{\psi_{2r}} \tilde{v}, \qquad k = 0, 1, 2, \ldots$$

This immediately gives the existence of positive numbers c and α depending only on λ, Λ, n, m, such that for $y \in \psi_{2r}(Q)$

$$\tilde{v}(y) \leqslant c\left(\frac{|y - Q|}{r}\right)^{\alpha} \sup_{\psi_{2r}} \tilde{v}.$$

Theorems I.1.1 and I.1.5 allow us to repeat verbatim the arguments given in [1, Lemma 2.4, p. 157], to prove a strengthened version of Theorem I.1.5.

Theorem I.1.6. (Boundary Harnack principle for n.a.s.) *Under the hypotheses of Theorem* I.1.5. *assume again $\tilde{v}$ is a nonnegative normalized adjoint solution for L^* and $\psi_{2r}(Q)$ which vanishes continuously on $\Delta_{2r}(Q)$, $Q \in \partial D$. There exists positive constants C and α depending only on λ, Λ, n, and m, such that for $y \in \psi_r(Q)$*

$$\tilde{v}(y) \leqslant C\left(\frac{|y - Q|}{r}\right)^{\alpha} \tilde{v}(A_r(Q))$$

where, recall for $Q = (Q', Q_n)$, $A_r(Q) = (Q', Q_n + mr)$.

A consequence of Theorem I.1.6 is the validity of all the results in [1] with constants depending only on the parameters of ellipticity of L, dimension, and the Lipschitz character of D. In the next section we collect some of these results.

I.2. Potential Theory for Solutions of $Lu = 0$; the Doubling Property for L-Harmonic Measure

We remind the reader that

$$L = \sum_{i,j=1}^{n} a_{ij}(x) D^2_{x_i x_j}$$

is an elliptic operator satisfying

$$\lambda|\xi|^2 \leqslant \sum_{i,j} a_{ij}(x)\xi_i\xi_j \leqslant \Lambda|\xi|^2$$

for all x and ξ in $\mathbb{R}^n$. (λ and Λ are positive constants.) Also D is a Lipschitz domain with m and r_0 describing the Lipschitz character. (See Definition A.)

Theorem I.2.1. *For $r \leqslant r_0/6$ let $g_r(x,y)$ denote the Green's function for L and $\psi_{4r}(Q)$, $Q \in \partial D$. Pick $\delta > 0$ such that $B_{2\delta r}(A_r(Q)) \subset D$. There exists a constant C depending on λ, Λ, n, and m such that for all $x \in \psi_{4r}(Q)\setminus\psi_{2r}(Q)$*

$$\int_{\psi_r(Q)} g_r(x,y)\,dy \leqslant C\int_{B_{\delta r}(A_r(Q))} g_r(x,y)\,dy.$$

PROOF. Let $\mathcal{B}$ be a ball with $\frac{1}{4}\,\mathcal{B} \supset \bar{D}$, P a point in $\frac{3}{4}\,\mathcal{B}\setminus\frac{1}{4}\,\mathcal{B}$, and $g_{\mathcal{B}}(x,y)$ the Green's function for L and $\mathcal{B}$. Also set

$$\tilde{v}(y) = \frac{g_r(x,y)}{g_{\mathcal{B}}(x,y)} \quad \text{for} \quad y \in \psi_{Ar}(Q), \qquad y \neq x.$$

Using Theorem I.1.6, for $x \in \psi_{4r}(Q)\setminus\psi_{2r}(Q)$,

$$\int_{\psi_r(Q)} g_r(x,y)\,dy \leqslant C\tilde{v}(A_r(Q))\int_{\psi_r(Q)} g_{\mathcal{B}}(P,y)\,dy.$$

The (interior) doubling property for nonnegative solutions of the adjoint equation [4] implies

$$\int_{\psi_r(Q)} g_{\mathcal{B}}(P,y)\,dy \leqslant C\int_{B_{\delta r}(A_r(Q))} g_{\mathcal{B}}(P,y)\,dy.$$

Harnack's inequality for $\tilde{v}$ (Theorem I.1.1) gives

$$\tilde{v}(A_r(Q))\int_{B_{\delta r}(A_r(Q))} g_{\mathcal{B}}(P,y)\,dy \leqslant C\int_{B_{\delta r}(A_r(Q))} g_r(x,y)\,dy$$

for all $x \in \psi_{4r}(Q)\setminus\psi_{2r}(Q)$.

Theorem I.2.1. is the main result used by Bauman in [1] to pove a Comparison Theorem for solutions of $Lu = 0$. Her argument [1, p. 160-161], can be repeated to obtain

Theorem I.2.2. (Comparison Theorem for solutions of $Lu = 0$.) *Let $Q \in \partial D$ and assume u_1 and u_2 are positive solutions of $Lu = 0$ in $\psi_{2r}(Q)$ which vanish on $\Delta_r(Q)$, $r \leqslant r_0$. Then*

$$\sup_{\psi_r} \frac{u_1}{u_2} \leqslant C\frac{u_1(A_r(Q))}{u_2(A_r(Q))}$$

where as usual C depends only on λ, Λ, n, *and the Lipschitz character of* D.

Definition D. *Given an elliptic operator*

$$L = \sum_{i,j} a_{ij}(x) D^2_{x_i x_j}$$

with smooth coefficients and a bounded Lipschitz domain D, *for any given* $g \in C(\partial D)$ *there exists a unique solution* u *to the Dirichlet problem*

$$Lu = 0 \quad in \quad D, \qquad u|_{\partial D} = g.$$

Fixing a point $x \in D$, *the maximum principle implies that the map* $g \mapsto u(x)$ *is a positive linear functional on* $C(\partial D)$. *Hence there exists a unique regular Borel measure* ω^x *on* ∂D *such that for every* $g \in C(\partial D)$

$$u(x) = \int_{\partial D} g(Q)\omega^x(dQ).$$

The measure $\omega^x(\partial D)$ *is called the L-harmonic measure for D evaluated at x. At times we will emphasize the dependence of* ω^x *on L and D by using the notation* $\omega^x_{L,D}$ *or* ω^x_D.

Let us say that two objects A and B (numbers or functions) are equivalent and write $A \sim B$ if there exists a positive constant C depending at most on ellipticity parameters, dimension, and the Lipschitz character of D such that

$$\frac{1}{C} A \leqslant B \leqslant CA.$$

A consequence of the Comparison Theorem for solutions I.2.2 is

Theorem I.2.3. (Doubling property for L-harmonic measure.) *Let* ω^x *denote the L-harmonic measure for D evaluted at x. Take* $r \leqslant r_0/4$ (r_0 *as in Definition A*). *Then for all* $Q \in \partial D$ *and* $x \in D \setminus \psi_{2r}(Q)$

$$\omega^x(\Delta_r(Q)) \sim \omega^x(\Delta_{2r}(Q)).$$

PROOF. By applying Theorem I.2.2 and the maximum principle to the functions $\omega^x(\Delta_{2r})$ and $\omega^x(\Delta_r)$ in the domain $D \setminus \psi_{4r}$, we obtain

$$\omega^x(\Delta_{2r}) \leqslant C \frac{\omega^{A_{4r}}(\Delta_{2r})}{\omega^{A_{4r}}(\Delta_r)} \omega^x(\Delta_r)$$

for all $x \in D \setminus \psi_{4r}$. There exists positive constants C and α, depending on λ, Λ, n, and m, such that for $x \in \psi_{r/2}$,

$$0 \leqslant 1 - \omega^x(\Delta_r) \leqslant C\left(\frac{|x-Q|}{r}\right)^{\alpha}$$

and

$$0 \leqslant 1 - \omega^x(\Delta_{2r}) \leqslant C\left(\frac{|x-Q|}{r}\right)^{\alpha}.$$

(See [1, Lemma 2.3, p. 157].) This observation and Harnack's inequality [4], [9] imply $\omega^{A_4 r}(\Delta_{2r})$ and $\omega^{A_4 r}(\Delta_r)$ are each equivalent to an absolute constant depending on λ, Λ, n, and m.

Other consequences of Theorem I.2.2 are concerned with the so-called kernel function. Proofs can be found in [2].

Definition E. *The kernel function $K(x, Q)$, normalized at $x_0 \in D$, is defined for $x \in D$ and $Q \in \partial D$ by*

$$K(x, Q) = \lim_{y \to Q} \frac{g_D(x, y)}{g_D(x_0, y)}$$

where g_D denotes the Green's function corresponding to L and D. This is the same as

$$K(x, Q) = \frac{d\omega^x}{d\omega^{x_0}}(Q).$$

In fact we have the following characterization of the kernel function.

Theorem I.2.4. *The kernel function is uniquely determined by the conditions*

(*a*) $L(K(\cdot, Q))(x) = 0$ *for* $x \in D$.
(*b*) $K(x_0, Q) = 1$ *for all* $Q \in \partial D$.
(*c*) *If* $Q' \in \partial D$ *and* $Q' \neq Q$, *then* $\lim_{x \to Q'} K(x, Q) = 0$.

The relationships between L-harmonic measure evaluated at x_0 and the kernel function are given in the next theorem.

Theorem I.2.5. *Let $K(x, Q)$ be the kernel function for L and D, normalized at x_0, and set $\omega^x = \omega^x_{L,D}$. Then there exists a sequence of constants $\{C_j\}$ such that C_j depends only on λ, Λ, n, m, r_0, and the* dist $(x_0, \partial D)$, $\sum C_j < \infty$, *and*

(i) $K(A_r(Q), Q) \sim \omega^{x_0}(\Delta_r(Q))^{-1} \left(Q \in \partial D, r \leqslant \frac{r_0}{2}\right)$

(ii) *for t sufficiently small and all $\tilde{Q} \in \partial D$*

$$\sup \{K(A_t(\tilde{Q}), Q): Q \in \Delta_{2j+1t}(\tilde{Q}) \setminus \Delta_{2jt}(\tilde{Q})\} \leqslant \frac{C_j}{\omega^{x_0}(\Delta_{2jt}(\tilde{Q}))}.$$

We conclude Section I.2 by remarking that via the techniques in [8], Theorems I.2.1.-I.2.3. imply the Hölder continuity of $K(x, \cdot)$ on ∂D with x fixed in D. The Hölder exponent can be taken depending only on λ, Λ, n, and m, and the Hölder norm can be bounded by a constant depending on these numbers, r_0, and the distances of x and x_0 to ∂D. This Hölder continuity and the properties described in Theorem I.2.5 of $K(x, Q)$ are main ingredients in the proofs of the Fatou theorems for nonlinear equations described in Part II. We approach the Hölder continuity of the kernel function from a different point of view than the one taken in [8]. We will prove that for fixed x and $x_0 \in D$ the function

$$g_D(x, y)/g_D(x_0, y)$$

is Hölder continuous in y near and up to the boundary of D. This is the subject of the next section.

I.3. A Comparison Theorem for Normalized Adjoint Solutions

We begin Section I.3 by establishing the doubling property for normalized adjoint measure.

Lemma I.3.1. *Let $Q \in \partial D$, $0 < r \leqslant r_0/2$, and denote by $\tilde{\omega}_r^y$ the normalized adjoint measure corresponding to L^* and $\psi_r(Q)$. Let*

$$\alpha_r = \partial\psi_r(Q) \cap D \quad \text{and} \quad \beta_r = B_{\delta_r/2}(A_r(Q)) \cap \partial\psi_r(Q),$$

where $\delta_r = \min\{\operatorname{dist}(A_r(Q), \partial D), r\}$. Then for $y \in \psi_{r/2}(Q)$,

$$\tilde{\omega}_r^y(\alpha_r) \sim \tilde{\omega}_r^y(\beta_r).$$

PROOF. Since β_r is a smooth portion of the boundary,

$$\tilde{\omega}_r^y(\beta_r) = \int_{\beta_r} \frac{\partial}{\partial v_Q} g_r(Q, y) g_{\mathcal{B}}(P, Q) \sigma(dQ)/g_{\mathcal{B}}(P, y)$$

where g_r is the Green's function corresponding to L and $\psi_r(Q)$. For $y \in \psi_{r/2}(Q)$, Hopf's Lemma [1, Lemma 4.3, p. 166], gives

$$\tilde{\omega}_r^y(\beta_r) \sim \frac{g_r(A_{r/2}(A), y)}{r g_{\mathcal{B}}(P, y)} \int_{\beta_r} g_{\mathcal{B}}(P, Q)\sigma(dQ). \tag{I.3.2}$$

The remaining part of the proof is dedicated to showing that for $y \in \psi_{r/2}(Q)$

$$\tilde{\omega}_r^y(\alpha_r) \leqslant C \frac{g_r(A_{r/2}(Q), y)}{r g_{\mathcal{B}}(P, y)} \int_{\beta_r} g_{\mathcal{B}}(P, Q)\sigma(dQ).$$

Choose h a smooth function in $\mathbb{R}^n$ such that $h = 1$ in $T_r(Q) \setminus T_{r/3}(Q)$, $h = 0$ in $T_{r/4}(Q)$, and $0 \leqslant h \leqslant 1$ everywhere. Then

$$\tilde{\omega}^y(\alpha_r) \leqslant \tilde{\omega}_r^y(h) = \tilde{\omega}_r^y(h) - h(y) \quad \text{for} \quad y \in \psi_{r/4}. \tag{I.3.3}$$

The function $g_{\mathcal{B}}(P, y)\tilde{\omega}_r^y(h)$ is a solution in ψ_r of the adjoint equation $L^*v = 0$ with boundary values $g_{\mathcal{B}}(P, Q)h(Q)$. Hence

$$g_{\mathcal{B}}(P, y)[\tilde{\omega}_r^y(h) - h(y)] = \int_{\psi_r} L^*[g_{\mathcal{B}}(P, z)h(z)]g_r(z, y)\, dz. \tag{I.3.4}$$

Now

$$L^*[g_{\mathcal{B}}(P, z)h(z)] = g_{\mathcal{B}}(P, z)Lh(z) + \sum_{i,j} D_{z_i}(a_{ij}(z)g_{\mathcal{B}}(P, z))D_{z_j}h(z).$$

We may choose h so that

$$|Lh| \leqslant \frac{C}{r^2}, \qquad |\nabla h| \leqslant \frac{C}{r},$$

and ∇h is supported in $T_{r/3}(Q) \setminus T_{r/4}(Q)$. Therefore, for $y \in \psi_{r/6}(Q)$ the integral in I.3.4 can be written, after an integration by parts, as

$$2 \int_{\psi_{r/3} \setminus \psi_{r/4}} g_{\mathcal{B}}(P, z)Lh(z)g_r(z, y)\, dz + 2 \int_{\psi_{r/3} \setminus \psi_{r/4}} g_{\mathcal{B}}(P, z)a(z)\nabla h(z) \cdot \nabla_z g_r(z, y)\, dz \equiv I + II.$$

Applying to $g_r(\cdot, y)$ the «boundary Harnack principle for positive solutions of $Lu = 0$", [1 Lemma 2.4, p. 157], $g_r(z, y) \leqslant C g_r(A_{r/2}(Q), y)$ for all $z \in \psi_{r/3} \setminus \psi_{r/4}$ and $y \in \psi_{r/6}$. Hence

$$|I| \leqslant \frac{C}{r^2} g_r(A_{r/2}(Q), y) \int_{\psi_{r/3} \setminus \psi_{r/2}} g_{\mathcal{B}}(P, z)\, dz, \qquad y \in \psi_{r/6}.$$

Again from the interior doubling property of the measure $g_{\mathcal{B}}(P, z)\, dz$, [4],

$$\int_{\psi_{r/3}} g_{\mathcal{B}}(P, z)\, dz \leqslant C \int_{B_{\delta_r/2}(A_r(Q))} g_{\mathcal{B}}(P, z)\, dz.$$

Let $g_{\delta_r}(x, y)$ denote the Green's function for L and $B_{\delta_r}(A_r(Q))$. Then for suitably chosen $\eta > 0$ and from Theorem I.1.1.,

$$\int_{B_{\delta_r/2}(A_r(Q))} g_{\mathcal{B}}(P, z)\, dz \leqslant C \frac{g_{\mathcal{B}}(P, A_r(Q))}{g_{\delta_r}(A_{\eta r}(Q), A_r(Q))} \int_{B_{\delta_r/2}(A_r(Q))} g_{\delta_r}(A_{\eta r}(Q), z)\, dz.$$

Using a dilation with center $A_r(Q)$ we have

$$\int_{B_{\delta_r/2}(A_r(Q))} g_{\delta_r}(A_\eta^r(Q), z)\, dz \sim r^2.$$

With this same dilation and Lemma I.1.4.

$$\int_{\beta_r} g_{\delta_r}(A_{\eta r}(Q), Q)\sigma(dQ) \sim r.$$

Hence

$$|I| \leqslant \frac{C}{r} g_r(A_{r/2}(Q), y) \frac{g_{\mathcal{B}}(P, A_r(Q))}{g_{\delta_r}(A_{\eta r}(Q), A_r(Q))} \int_{\beta_r} g_{\delta_r}(A_{\eta r}(Q), Q)\sigma(dQ),$$

and from Theorem I.1.1

$$|I| \leqslant \frac{C}{r} g_r(A_{r/2}(Q), y) \int_{\beta_r} g_{\mathcal{B}}(P, Q)\sigma(dQ) \qquad (y \in \psi_{r/6}). \tag{I.3.5}$$

We now handle

$$II \equiv 2 \int_{\psi_{r/3} \setminus \psi_{r/4}} g_{\mathcal{B}}(P, z) a(z) \nabla h(z) \cdot \nabla_z g_r(z, y)\, dz.$$

By Schwartz's inequality

$$|II| \leqslant \frac{C}{r} \left(\int_{\psi_r \setminus \psi_{r/4}} g_{\mathcal{B}}(P, z)\, dz \right)^{1/2} \left(\int_{\psi_r \setminus \psi_{r/4}} g_{\mathcal{B}}(P, z) |\nabla_z g_r(z, y)|^2\, dz \right)^{1/2}.$$

Choose $l(z) \in C_0^\infty(\mathbb{R}^n)$, $0 \leqslant l \leqslant 1$, satisfying $l = 1$ on $T_r \setminus T_{r/4}$, $l = 0$ on $T_{r/5}$, $|\nabla l| \leqslant cr^{-1}$ and $|D^2_{z_i z_j} l| \leqslant cr^{-2}$ for all i and j. For $z \in T_r \setminus T_{r/4}$ and $y \in \psi_{r/6}$,

$$\begin{aligned} |\nabla_z (L(z) g_r(z, y))|^2 &\leqslant c a(z) \nabla_z (l(z) g_r(z, y)) \cdot \nabla_z (l(z) g_r(z, y)) \\ &\leqslant c \{ L[(l(z) g_r(z, y))^2] - 2 l(z) L(l(z)) g_r^2(z, y) \\ &\quad + g_r^2((z, y) a(z) \nabla l(z)) \cdot \nabla l(z) \\ &\quad - 2 l(z) g_r(z, y) a(z) \nabla l(z) \cdot \nabla_z (l(z) g_r(z, y)) \}. \end{aligned}$$

Since $l(\cdot) g_r(\cdot, y) = 0$ on $\partial(\psi_r \setminus \psi_{r/5})$, for $y \in \psi_{r/6}$

$$\int_{\psi_r \setminus \psi_{r/5}} g_{\mathcal{B}}(P, z) L[(l(z) g_r(z, y))^2]\, dz = 0.$$

Therefore, applying again to $g_r(\cdot, y)$ the «boundary Harnack principle for positive solutions of $Lu = 0$» [1, Lemma 2.4, p. 157], it easily follows that for $y \in \psi_{r/6}$

$$\int_{\psi_r \setminus \psi_{r/5}} g_{\mathcal{B}}(P, z) |\nabla_z (l(z) g_r(z, y))|^2 \, dz \leqslant \frac{C}{r^2} g_r^2(A_{r/2}(Q), y) \int_{\psi_r \setminus \psi_{r/5}} g_{\mathcal{B}}(P, z) \, dz.$$

Finally from the arguments controlling I

$$|II| \leqslant \frac{C}{r^2} g_r(A_{r/2}(Q), y) \int_{\psi_r \setminus \psi_{r/5}} g_{\mathcal{B}}(P, z) \, dz \tag{I.3.6}$$

$$\leqslant \frac{C}{r} g_r(A_{r/2}(Q), y) \int_{\beta_r} g_{\mathcal{B}}(P, Q) \sigma(dQ).$$

The conclusion from I.3.2-I.3.6 is: for $y \in \psi_{r/6}(Q)$,

$$\tilde{\omega}_r^y(\alpha_r) \leqslant \frac{C}{r} \frac{g_r(A_{r/2}(Q), y)}{g_{\mathcal{B}}(P, y)} \int_{\beta_r} g_{\mathcal{B}}(P, Q) \sigma(dQ) \sim \tilde{\omega}_r^y(\beta_r).$$

We are now in position to prove

Theorem I.3.7. (Comparison Theorem for adjoint solutions.) *Let D be a Lipschitz domain whose Lipschitz character is determined by the numbers r_0 and m. Assume L is a uniformly elliptic operator with parameters of ellipticity λ and Λ (and having smooth coefficients). Let v and w be two nonnegative adjoint solutions, i.e. $L^*v = L^*w = 0$, in $\psi_{2r}(Q)$, $Q \in \partial D$, continuously vanishing on $\Delta_{2r}(Q)$ $(r \leqslant r_0/2)$.*

Then

$$\frac{v(y)}{w(y)} \sim \frac{v(A_r(Q))}{w(A_r(Q))} \quad \text{for all} \quad y \in \psi_r(Q).$$

PROOF. The functions

$$\tilde{v}(y) = \frac{v(y)}{g_{\mathcal{B}}(P, y)} \quad \text{and} \quad \tilde{w}(y) = \frac{w(y)}{g_{\mathcal{B}}(P, y)}$$

are normalized adjoint solutions satisfying the hypothesis of Theorem I.1.5. From this theorem, Lemma I.3.1, and the technique in [1, Proof of Theorem 2.1, p. 160], I.3.7 follows for v and w replaced by $\tilde{v}$ and $\tilde{w}$. Since $\tilde{v}/\tilde{w} = v/w$ the conclusion also holds for v and w.

Corollary I.3.8. *Let v and w satisfy the hypotheses of Theorem* I.3.7 *in $\psi_{r_0}(Q)$, $Q \in \partial D$. Then there exist positive constants C and α depending on λ,*

Λ, n, *and* m *such that*

$$\left|\frac{v(y)}{w(y)} - \frac{v(y')}{w(y')}\right| \leqslant C\frac{v(A_{r_0}(Q))}{w(A_{r_0}(Q))}\left(\frac{|y-y'|}{r_0}\right)^{\alpha}$$

for all y *and* y' *in* $\psi_{r_0/2}(Q)\cap D$.

PROOF. The proof follows the method of Moser who obtains Hölder continuity from the uniform Harnack principle [12]. In fact that argument and Theorem I.I.5 give the above result when $y \in D \cap \overline{\psi_{r_0/2}(Q)}$ and $y' \in B_{s/2}(y)$, with $s = \operatorname{dist}(y, \partial D)$. Hence it is enough to show the Hölder continuity at a boundary point $Q_0 \in \Delta_{r_0/2}(Q)$. We may also assume $v(A_{r_0}(Q)) = w(A_{r_0}(Q)) = 1$.

Set

$$M(s) = \sup\left\{\frac{v(y)}{w(y)} : y \in \psi_s(Q_0)\right\}$$

and

$$m(s) = \inf\left\{\frac{v(y)}{w(y)} : y \in \psi_s(Q_0)\right\}.$$

Then for $s \leqslant r_0/4$,

$$M(s) - \frac{v}{w} = \frac{M(s)w - v}{w}$$

and

$$\frac{v}{w} - m(s) = \frac{v - m(s)w}{w}$$

are quotients of positive adjoint solutions in $\psi_s(Q_0)$ which vanish on $\Delta_s(Q_0)$. By Theorem I.3.7,

$$\sup_{\psi_{s/2}(Q_0)}\left(M(s) - \frac{v}{w}\right) = C\inf_{\psi_{s/2}(Q_0)}\left(M(s) - \frac{v}{w}\right)$$

and

$$\sup_{\psi_{s/2}(Q_0)}\left(\frac{v}{w} - m(s)\right) \leqslant C\inf_{\psi_{s/2}(Q_0)}\left(\frac{v}{w} - m(s)\right)$$

i.e.

$$M(s) - m(s/2) \leqslant C(M(s) - M(s/2))$$

and

$$M(s/2) - m(s) \leqslant C(m(s/2) - m(s)).$$

Adding the final two inequalities we obtain

$$M(s/2) - m(s/2) \leqslant \Theta(M(s) - m(s))$$

where

$$0 < \Theta = \frac{C-1}{C+1} < 1.$$

Iterating gives

$$M(s) - m(s) \leqslant \Theta^{-1}\left(\frac{s}{r_0}\right)^{\alpha}(M(r_0) - m(r_0))$$

with

$\alpha = -\log_2 \Theta$.

This concludes the proof of the Corollary.

When v and w represent the Green's function for an elliptic operator at two fixed poles we obtain the following important

Theorem I.3.9. *Assume $L = \sum a_{ij}(x)D^2_{x_i x_j}$, an elliptic operator, and D, a Lipschitz domain, satisfy the hypotheses of Theorem* I.3.7. *Let $g(x, y)$ denote the Green's function corresponding to L and D. If x and x_0 are fixed points of D there exist positive constants C and α, α depending only on λ, Λ, n, m, r_0 but C depending in addition on $\delta = \min\{\operatorname{dist}(x, \partial D), \operatorname{dist}(x_0, \partial D)\}$, such that*

$$\left|\frac{g(x, y)}{g(x_0, y)} - \frac{g(x, y')}{g(x_0, y')}\right| \leqslant C|y' - y|^{\alpha}$$

for all y and y' belonging to $\{z \in D: \operatorname{dist}(z, \partial D) \leqslant \delta/2\}$. As a consequence, the kernel function $K(x, Q)$ corresponding to L and D and normalized at x_0 satisfies

$$|K(x, Q) - K(x, Q')| \leqslant C|Q - Q'|^{\alpha}$$

for all Q and Q' on ∂D; C and α as above.

II. FATOU THEOREMS FOR SOLUTIONS OF NONLINEAR EQUATIONS

In Part II, for the sake of simplicity but without any loss of generality, we consider Lipschitz domains D starshaped with respect to the origin.

II.1. The p-Laplacian

The first application we give of the potential theory developed in Part I is to the boundary behavior of nonnegative p-harmonic functions.

Definition F. *Fix $1 < p < \infty$. A function $u(x)$, $x \in D$, is a solution of the equation*

$$\Delta_p u = \operatorname{div}(|\nabla u|^{p-2}\nabla u) = 0 \text{ in } D \tag{II.1.1}$$

if $u \in W^{1,p}_{\text{loc}}(D) = \{u \in L^p_{\text{loc}}(D): \nabla u \in L^p_{\text{loc}}(D)\}$ and

$$\int_D |\nabla u|^{p-2}\nabla u \cdot \nabla \varphi \, dx = 0$$

for all $\varphi \in C_0^\infty(D)$. A function u which is a solution of II.1.1. *is called a p-harmonic function and Δ_p is called the p-Laplacian.*

Despite the degenerate character of II.1.1. solutions belong to $C^{1,\alpha}_{\text{loc}}(D)$ with $\alpha = \alpha(p, n) > 0$. (See [3] and [10].) Concerning boundary behavior we have

Theorem II.1.2. *Let u be a nonnegative p-harmonic function in D. Then $\{Q \in \partial D$: u has nontangential limit at $Q\}$ has Hausdorf dimension $\geqslant \beta > 0$ where β depends only on p, n, and the Lipschitz character of D.*

PROOF. Since D is starshaped with respect to the origin, for $0 < r < 1$, we set $u_r(x) = u(rx)$ and obtain a p-harmonic function u_r in D which is $C^{1,\alpha}$ in $\bar{D}$, the closure of D. Now consider the regularized p-harmonic operator

$$\Delta_p^\epsilon v = \operatorname{div}((|\nabla v|^2 + \epsilon)^{(p-2)/2}\nabla v), \qquad \epsilon > 0, \tag{II.1.3}$$

and denote by $v_{\epsilon,r}$ the unique solution of the Dirichlet problem

$$\Delta_p^\epsilon v = 0 \quad \text{in} \quad D, \qquad v|_{\partial D} = u_r|_{\partial D}.$$

Because of the nondegeneracy of Δ_p^ϵ, $v_{\epsilon,r}$ belongs to $C^\infty(D)$. Therefore we can perform the differentiation in II.1.3 and, after dividing by $(\epsilon + |\nabla v_{\epsilon,r}|^2)^{(p-2)/2}$, we see that $v_{\epsilon,r}$ is a solution of the linear equation

$$L^u_{\epsilon,r} w = \sum_{i,j=1}^n \left(\delta_{ij} + (p-2)\frac{D_{x_i} v_{\epsilon,r} D_{x_j} v_{\epsilon,r}}{\epsilon + |\nabla v_{\epsilon,r}|^2}\right) D^2_{x_i x_j} w = 0.$$

(We have added the superscript u to the operator to keep in mind the dependence of the coefficients on the p-harmonic function u.) The operator $L^u_{\epsilon,r}$ has parameters of ellipticity λ, Λ depending only on p; namely, $\lambda = \min\{1, p-1\}$ and $\Lambda = \min\{1, p-1\}$, and the coefficients belong to $C^\infty(D)$.

A key property of $v_{\epsilon,r}$ is that $v_{\epsilon,r} \to u_r$ in $W^{1,p}(D)$ as $\epsilon \to 0$, [10]. Also as $\epsilon \to 0$ $v_{\epsilon,r} \to u_r$ uniformly on compact subsets of D.

We make one more dilation: the function $v_{\epsilon,r}(rx)$ satisfies an elliptic equation $\tilde{L}^u_{\epsilon,r} w = 0$ of the same form as that defined by $L^u_{\epsilon,r}$. The parameters of ellipticity can be taken dependent only on p and the coefficients now belong to $C^\infty(\bar{D})$. For $v_{\epsilon,r}(rx)$ we have the following representation formula:

$$v_{\epsilon,r}(rx) = \int_{\partial D} v_{\epsilon,r}(rQ)\tilde{K}^u_{\epsilon,r}(x, Q)\tilde{\omega}^u_{\epsilon,r}(dQ) \tag{II.1.4}$$

where $\tilde{\omega}^u_{\epsilon,r}$ denotes the $\tilde{L}^u_{\epsilon,r}$-harmonic measure for D evaluated at the origin and $\tilde{K}^u_{\epsilon,r}(x, Q)$ denotes the kernel function associated with $\tilde{L}^u_{\epsilon,r}$ and D, normalized at the origin.

Since

$$\int_{\partial D} \tilde{\omega}^u_{\epsilon,r}(dQ) = 1,$$

we can select a sequence $\epsilon_j \to 0$ such that $\tilde{\omega}^u_{\epsilon_j,r}(dQ)$ converges weakly (as $j \to \infty$) to a regular Borel measure $\tilde{\omega}^u_r(dQ)$. Also, from the local uniform convergence of $v_{\epsilon,r}$ to u_r in D,

$$\lim_{j\to\infty} v_{\epsilon_j,r}(rQ) = u(r^2Q)$$

uniformly on ∂D. Finally recall that each kernel function $\tilde{K}^u_{\epsilon,r}(x,Q)$ is a solution in x of a second order elliptic equation with parameters of ellipticity depending only on p. Since $\tilde{K}^u_{\epsilon,r}(0, Q) = 1$. $\tilde{K}^u_{\epsilon,r}(x, Q)$ is locally Hölder continuous in X with local Hölder exponent and norm independent of Q, ϵ, r, and u. Also from Theorem I.3.9 when x varies over a fixed compact subset of D. $\tilde{K}^u_{\epsilon,r}(x, Q)$ is Hölder continuous on ∂D, with Hölder exponent depending only on p, n, and the Lipschitz character of D, and Hölder norm bounded independently of ϵ and r. Hence we may assume the sequence $\{K^u_{\epsilon_j,r}(x, Q)\}$ converges uniformly on ∂D for each fixed $x \in D$. The function

$$\tilde{K}^u_r(x, Q) = \lim_{j\to\infty} \tilde{K}^u_{\epsilon_j,r}(x, Q)$$

satisfies the same Hölder continuity just described for each kernel function of the sequence. We can now allow ϵ_j tend to zero in II.1.4 and obtain the representation

$$u(r^2x) = \int_{\partial D} u(r^2Q)\tilde{K}^u_r(x, Q)\tilde{\omega}^u_r(dQ). \tag{II.1.5}$$

We repeat once more the arguments of the previous paragraph:

$$u(0) = \int_{\partial D} u(r^2Q)\tilde{\omega}^u_r(dQ) \quad \text{and} \quad 1 = \int_{\partial D} \tilde{\omega}^u_r(dQ)$$

imply there exists a sequence $r_j \nearrow 1$ and regular Borel measures $\mu^u(dQ)$ and $\tilde{\omega}^u(dQ)$ such that as $r_j \nearrow 1$, $u(r_j^2 Q)\tilde{\omega}^u_{r_j}(dQ)$ converges weakly to $\mu^u(dQ)$ while $\tilde{\omega}^u_{r_j}(dQ)$ converges weakly to $\tilde{\omega}^u(dQ)$. As already noted in the previous paragraph $\tilde{K}^u_r(x, Q)$ is Hölder continuous in x and Q for $Q \in \partial D$ and x restricted to a compact subset of D. The Hölder norm can be bounded and the Hölder exponent can be written independently of r. Hence as $r_j \nearrow 1$ we may assume

$$\tilde{K}^u_{r_j}(x, Q) \to \tilde{K}^u(x, Q)$$

uniformly on ∂D for each $x \in D$. Letting $r_j \nearrow 1$ we obtain the final representation

$$u(x) = \int_{\partial D} \tilde{K}^u(x,Q)\mu^u(dQ). \tag{II.1.6}$$

The measures $\tilde{\omega}^u_{\epsilon,r}(dQ)$ enjoy the doubling condition of Theorem I.2.3 (for $x = 0$) with a doubling constant depending only on p, n, and the Lipschitz character of D. The same property holds for the weak limits $\tilde{\omega}^u_r$ and $\tilde{\omega}^u$. On the other hand, the relationships between $\tilde{K}^u_{\epsilon,r}$ and $\tilde{\omega}^u_{\epsilon,r}$ expressed in Theorem I.2.5 carry over to $\tilde{K}^u$ and $\tilde{\omega}^u$. These relationships and the doubling property of $\tilde{\omega}^u$ allow us to apply the procedure in [2] obtaining the existence of nontangential limits of u at all points $Q \in \partial D$ except for a set of $\tilde{\omega}^u$ measure zero.

In particular $\tilde{\omega}^u(\{Q \in \partial D: u \text{ has nontangential limit at } Q\}) = \tilde{\omega}^u(\partial D) = 1$. But the doubling property of $\tilde{\omega}^u$; namely.

$$\tilde{\omega}^u(\Delta_{2r}) \leqslant C\tilde{\omega}^u(\Delta_r) \qquad (r \leqslant r_0)$$

with C depending only on p, n, and the Lipschitz character of D implies there exist positive constants C and β also depending only on these parameters such that

$$\tilde{\omega}^u(\Delta_r) \leqslant Cr^\beta \qquad (r \leqslant r_0).$$

In particular the Hausdorf dimension of the $\{Q \in \partial D$: u has nontangential limit at $Q\}$ is $\geqslant \beta$.

In the next two sections we discuss two other classes of equations for which one can obtain a Fatou theorem of the type of Theorem II.1.2. Since the proofs follow so closely the one for the p-Laplacian, we will only briefly indicate the arguments.

II.2. The Equation $\operatorname{div}(f'(|\nabla u|)\nabla u/|\nabla u|) = 0$

We begin with a function $f(t)$ defined for $t > 0$, in $C^2(0, \infty)$, and which is positive, increasing and convex in $(0, \infty)$ see [5]. We also assume there exists

$C > 0$ such that (II.2.1)

(i) $\frac{1}{C}(t^p - 1) \leqslant tf'(t) \leqslant C(t^p + 1)$, $t > 0$, p fixed $1 < p < \infty$.

(ii) $\frac{1}{C} \leqslant t\frac{f''(t)}{f'(t)} \leqslant C, \qquad t > 0.$

We note that II.2.1 (ii) implies $\lim_{t \to 0^+} f'(t) = 0$.

Definition G. *We say that $u \in W^{1,p}(D)$ is a weak solution in D of*

$$\operatorname{div}\left(\frac{f'(|\nabla u|)}{|\nabla u|}\nabla u\right) = 0 \tag{II.2.2}$$

if

$$\int_D \frac{f'(|\nabla u(x)|)}{|\nabla u(x)|}\nabla u(x) \cdot \nabla \varphi(x) = 0$$

for all $\varphi \in C_0^\infty(D)$. (The integrand in the integral is intepreted to be 0 at each x where $\nabla u(x) = 0$.)

Theorem II.2.3. *Assume u is a nonnegative weak solution of* II.2.2 *in a (starshaped) Lipschitz domain D. There exists $\beta > 0$ depending only on the constant C in* (II.2.1) (i) *and* (ii), *dimension, and the Lipschitz character of D, such that the Hausdorf dimension of $\{Q \in \partial D$: u has nontangential limit at $Q\}$ is $\geqslant \beta$.*

PROOF. We proceed in the manner of Theorem II.1.2. For $0 < r < 1$ set $u_r(x) = u(rx)/r$. This function is a solution of (II.2.2) with data $u(rQ)/r$. For $\epsilon > 0$ let $u_{\epsilon,r}$ be the solution to the Dirichlet problem

$$\operatorname{div}\left(\frac{f'[(|\nabla v|^2 + \epsilon)^{1/2}]}{(|\nabla v|^2 + \epsilon)^{1/2}}\nabla v\right) = 0,$$

$$v|_{\partial D} = u_r|_{\partial D}.$$

Again the nondegenerate nature of this problem implies $u_{\epsilon,r}$ is smooth in D. ($u_{\epsilon,r}$ at least belongs to $W^{2,p}(D)$ for all finite p.) $u_{\epsilon,r}$ is itself a solution of the nonvariational equation

$$0 = L_{\epsilon,r}(w)(x) = \sum_{i,j=1}^{n} a^{\epsilon,r}(x) D^2_{x_i x_j} w(x),$$

where for $t = (|\nabla u_{\epsilon,r}(x)|^2 + \epsilon)^{1/2}$

$$a_{ij}^{\epsilon,r}(x) = \delta_{ij} + \frac{(f''(t)t - f'(t))}{f'(t)} \frac{D_{x_i} u_{\epsilon,r}(x) D_{x_j} u_{\epsilon,r}(x)}{(|\nabla u_{\epsilon,r}(x)|^2 + \epsilon)}.$$

Condition (II.2.1)-(ii) implies $L_{\epsilon,r}$ is uniformly elliptic with parameters of ellipticity depending on the constant C in (II.2.1)-(ii).

The function $u_{\epsilon,r} \to u_r$ as $\epsilon \to 0$ uniformly on compact subsets of D and the function $u_{\epsilon,r}(x) = u_{\epsilon,r}(rx)/r$ satisfies the same elliptic equation as $u_{\epsilon,r}$ with $\nabla u_{\epsilon,r}(x)$ in the coefficients replaced by $\nabla u_{\epsilon,r}(rx)$. We now repeat verbatim the remaining arguments in the case of the p-Laplacian to conclude the proof of Theorem II.2.3.

II.3. Fully Nonlinear Equations

As indicated in the introduction, a Fatou theorem holds for nonnegative solutions of fully nonlinear equations. Specifically we consider equations of the type

$$F(D^2u(x), x) = 0 \quad \text{in} \quad D, \tag{II.3.1}$$

where $F(M, x)$ is smooth with respect to the $n \times n$ matrix variables M and n-dimensional variable x in D. We further assume a uniformly ellipticity condition

$$\lambda|\xi|^2 \leqslant \sum_{i,j=1}^{n} \frac{\partial F}{\partial M_{ij}}(M, x)\xi_i\xi_j \leqslant \Lambda|\xi|^2$$

for all $M \in \mathbb{R}^{n^2}$, $x \in D$, $\xi \in \mathbb{R}^n$.

Theorem II.3.2. *Assume $u \in C^2(D)$ is a nonnegative solution of* II.3.1 *in D, a Lipschitz domain starshaped with respect to* 0. *There exists a positive number β depending only on λ, Λ, n, and the Lipschitz character of D such that if $F(0, x) \in L^n(D)$, then the Hausdorf dimension of $\{Q \in \partial D$: u has nontangential limit at $Q\}$ is $\geqslant \beta$.*

PROOF. Again the proof follows the techniques used for p-harmonic functions. In fact, u satisfies the linear equation

$$Lv = \sum_{i,j} a_{ij}(x) D^2_{x_i x_j} v(x) = f(x)$$

where

$$a_{ij}(x) = \int_0^1 \frac{\partial F}{\partial M_{ij}}(tD^2u(x), x)\, dt \quad \text{and} \quad f(x) = -F(0, x).$$

The function

$$u_r(x) = u(rx), \qquad 0 < r < 1,$$

satisfies

$$L_r u_r = \sum_{i,j} a_{ij}(rx) D^2_{x_i x_j} u_r(x) = f(rx) r^2$$

and has the representation

$$u(rx) = \int_{\partial D} u(rQ) K^u_r(x, Q) \omega^u_r(dQ) - \int_D r^2 g^u_r(x, y) f(ry)\, dy$$

where K^u_r, ω^u_r, and g^u_r denote respectively the kernel function (normalized at 0), the L_r-harmonic measure (evaluated at 0), and the Green's function corresponding to L_r and D. Therefore the conclusions of Theorem II.3.2 follow if for every $Q \in \partial D$

$$\lim_{x \to Q} \int_D g^u_r(x, y) f(ry)\, dy = 0 \tag{II.3.3}$$

uniformly for $Q \in \partial D$ and r near 1. This follows easily from the observations

$$\sup_x \int_D g^u_r(x, y)^{n/(n-1)}\, dy \leqslant C(\lambda, \Lambda, n, \operatorname{diam} D),$$

[14] and for $x \in \psi_{s/2}(Q)$,

$$\left| \int_{D \setminus \psi_s(Q)} g^u_r(x, y) f(ry)\, dy \right| \leqslant C \left(\frac{|x - Q|}{s} \right)^{\alpha} \int_{D \setminus \psi_s(Q)} g^u_r(A_{s/2}(Q), y) |f(ry)|\, dy,$$

with $0 < C$ and $0 < \alpha$ depending only on λ, Λ, n, and the Lipschitz character of D. The last observation results from an application to $g^u_r(\cdot, y)$ of the «boundary Harnack principle» and Hölder continuity for nonnegative solutions of $L_r w = 0$ ([1, Lemmas 2.3 and 2.4, p. 157]).

References

[1] Bauman, P. Positive solutions of elliptic equations in nondivergence form and their adjoints, *Arkiv fur Matematik,* **22**(1984), 153-173.

[2] Caffarelli, L., Fabes, F., Mortola, S., Salsa, S. Boundary behavior of nonnegative solutions of elliptic operators in divergence form, *Indiana J. Math.,* **30**(1981), 621-640.

[3] Evans, L. C. A new proof of local $C^{1,\alpha}$ regularity for solutions of certain degenerate elliptic P.D.E., *J. Differential Equations,* **45**(1982), 356-373.

[4] Fabes, E. B., Stroock, D. W. The L^p-integrability of Green's functions and fundamental solutions for elliptic and parabolic equations, *Duke Math. J.*, **51**(1984), 997-1016.
[5] Garofalo, N., Lewis, J. L. A symmetry result related to some overdetermined boundary value problems, *Amer. J. of Math.* **111**(1989), 9-33.
[6] Hunt, R. R., Wheeden, R. L. On the boundary values of harmonic functions, *Trans. Amer. Math. Soc.* **132**(1968), 307-322.
[7] Jerison, D. S., Kenig, C. E. Boundary behavior of harmonic functions in nontangentially accessible domains, *Advances in Math.* **46**(1982), 80-147.
[8] —, The logarithm of the Poisson kernel for a C^1 domain has vanishing mean oscillation, *Trans. Amer. Math. Soc.* **273**(1982), 781-794.
[9] Krylov, N. V., Safanov, M. V. A certain property of solutions of parabolic equations with measurable coefficients, *Math. USSR Izvestija,* **16**(1981), 151-164; English translation in *Izv. Akad. Nauk SSSR.* **44**(1980), 81-98.
[10] Lewis, J. L. Regularity of the derivatives of solutions to certain degenerate elliptic equations, *Indiana J. Math.,* **32**(1983), 849-858.
[11] —, Note on a theorem of Wolff, in *Holomorphic function theory workshop.* Ed. by F. Gehring, MSRI.
[12] Manfredi, J., Weitsman A. On the Fatou theorem for p-harmonic functions, *Comm. P.D.E.* **13**(1988), 651-668.
[13] Moser, J. On Harnack's inequality for elliptic differential equations, *Comm. Pure Appl. Math.,* **14**(1961), 577-591.
[14] Pucci, C. Limitazioni per soluzioni di equazioni ellitiche, *Annali di Matematica Pura ed Applicata,* **74**(1966), 15-30.
[15] Wolff, T. Gap series construction for the p-Laplacian, *Annali Scuola Norm. Sup. Pisa*, to appear.

E. Fabes[1]
U. of Minnesota
Minneapolis, Minn 55455
U.S.A.

N. Garofalo
Northwestern U.
Evanston, Ill 60201,
U.S.A.

S. Marin-Malave[2]
U. of Minnesota
Minneapolis, Minn 55455
U.S.A.

S. Salsa
Politecnico di Milano
20133 Milano
ITALY

[1,2] Partially supported by NSF/DMS-8421377-03.

REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.º 2, 1988

H^p-Theory on Euclidean Space and the Dirac Operator

John E. Gilbert and Margaret A. M. Murray

To the memory of José Luis Rubio de Francia

Introduction

Classical H^p-theory is the study of the boundary regularity of one particular linear first-order differential operator: the Cauchy-Riemann $\bar{\partial}$-operator. This is an elliptic operator, but in applications to planar potential theory great use is made of its *analytic over-determinedness*, meaning that the imaginary part V of any solution $F = U + iV$ of $\bar{\partial}F = 0$ is essentially determined by its real part U (and vice-versa). Put another way, every analytic function F can be written locally as $\partial\Phi$ with Φ real-valued harmonic function. Such analytic over-determinedness is the basis for *real* H^p-theory. The close relationship to the Dirac operator that $\bar{\partial}$ and ∂ have been known to have for many years (cf. [20]) has been used in more recent times to develop a theory of analytic functions on Euclidean space of any dimension, complex-valued functions being replaced by Clifford algebra-valued functions ([6]). A routine application of Stokes theorem allows one to prove a Cauchy Integral Theorem, for instance, in this more general context. In [15], we then used the L^p-boundedness of the principal-value Cauchy integral operator associated with Lipschitz domains in $\mathbb{R}^{n+1}$ to establish a boundary value theory for H^p-spaces defined on such domains by a family of first-order differential operators said to be of *Dirac type*. These include not only the classical geometric differential operators, but also the rotation-invariant systems ∂_τ introduced in [9]. Such Dirac type operators are virtually the only ones for which a Cauchy integral theorem can exist ([16]). One purpose of the present paper is to outline these ideas in their «ultimate form» by exploiting the theory of Clifford algebras and Clifford

modules more fully than in the earlier paper. But more importantly we shall formulate a class of boundary value problems for these Hardy spaces modelled on the usual Riemann-Hilbert problem and its generalizations for $\bar{\partial}$. Just as the latter contains the prototypical Dirichlet, Neumann and Oblique-derivative problems, so this general class contains the well-known higher-dimensional versions of these problems as special cases (cf. [22]). The point to be emphasized is that the splitting $\mathbb{C} = \mathbb{R} \oplus i\mathbb{R}$ used previously to write a complex-valued function as $F = U + iV$ has a natural interpretation in the context of Clifford theory. Now classical H^p-theory, representation theory for classical Lie groups, and algebraic geometry questions concerning, say, compact Riemann surfaces all fall within complex analytic function theory. Hence if such operators of Dirac type as ∂_τ are to have a deeper meaning than being merely *analogues of* $(\bar{\partial}, \partial)$ *for higher-dimensional Euclidean space*, they should be defined on more general Riemannian manifolds, with elliptic boundary value problems and H^p-theory being just part of the analysis associated with them in the case of Euclidean space. Indeed, as Seeley has so clearly pointed out ([29]), the Riemann-Hilbert problem and its variants for $\bar{\partial}$ lead very naturally to elliptic singular integral equations whose solution is the forerunner of the Atiyah-Singer Index theorem for manifolds with or without boundary. These are the ideas basic to [14] and to the series of papers beginning with [9], [10] and [13]; but here we shall mention them only briefly. They have been formulated over a long period of time during which many conversations with Kathy Davis and Ray Kunze were immensely useful in helping fix the fundamental concepts. More recent conversations with John Ryan were very helpful too.

1. Classical Case

Let Ω be the interior of a Jordan curve in the complex plane or the region $\{z \in \mathbb{C}: \operatorname{Im}(z) > \phi(\operatorname{Re} z)\}$ above the graph of a function $\phi: \mathbb{R} \to \mathbb{R}$; the unit disk or upper half-plane are natural examples. We shall think of Ω as being «flat» when it is equipped with the Euclidean metric and «curved» when a Riemannian metric on it gives it non-zero curvature as in the case of the hyperbolic metric. More generally still, Ω could be any Riemann surface. Geometrically, the flat case arises from identifying $\mathbb{C}$ with the tangent space to a curved Ω at some base point $\omega_0 \in \Omega$. In the flat case the Cauchy-Riemann $\bar{\partial}$-operator has constant coefficients, whereas it will have non-constant coefficients when Ω has curvature. Our principal focus, of course, is on the flat case where the Hardy spaces form one natural family of spaces of analytic functions.

First let Ω be the upper half-plane in $\mathbb{C}$. The Hardy $H^p(\Omega)$-spaces here consist of all analytic functions on Ω such that

$$\|F\|_{H^p(\Omega)} = \sup_{y>0}\left(\int_{-\infty}^{\infty} |F(x+iy)|^p\,dx\right)^{1/p} \qquad (0<p<\infty) \tag{1.1}$$

is finite. If $\mathfrak{H}$ is any finite-dimensional complex Hilbert space, then corresponding $H^p(\Omega)$-spaces of $\mathfrak{H}$-valued analytic functions can be defined in exactly the same way. All the H^p-theory for scalar-valued functions carries over component by component to the vector-valued case.

For $p>1$ straightforward use of analyticity and weak*-compactness gives (1.2)

(i) (*characterization of boundary values*) each F in $H^p(\Omega)$ has boundary values $F^+(x)$ a.e. on $\partial\Omega \sim \mathbb{R}$ and the space

$$H^p(\partial\Omega) = \{F^+ : F \in H^p(\Omega)\}$$

coincides with the subspace of all complex-valued functions f in $L^p(\partial\Omega)$ such that $(I - i\mathcal{H})f = 0$ where

$$\mathcal{H}f(x) = P.V.\frac{1}{\pi}\int_{\partial\Omega}\frac{f(y)}{x-y}\,dy \qquad (x\in\partial\Omega)$$

is the Hilbert transform of f;

(ii) (*analytic overdeterminedness*) the Cauchy integral

$$f \to \mathcal{C}f(z) = \frac{1}{2\pi i}\int_{\partial\Omega}\frac{f(x)}{x-z}\,dx \qquad (z\in\Omega)$$

is an isomorphism from the space of all real-valued functions f in $L^p(\partial\Omega)$ onto $H^p(\Omega)$.

In proving the existence of boundary values in $L^p(\partial\Omega)$, $p>1$, harmonicity of F rather than analyticity would have been sufficient. But when $p=1$, the requirement of analyticity is essential. Thus one *role of $\bar\partial$ is to improve on the boundary regularity of harmonic functions*; in fact, boundary values F^+ exist for every F in $H^p(\partial\Omega)$, $p>0$. Stein-Weiss ([31]) showed that this fundamental boundary regularity property could be derived from the subharmonicity of $z \to |F(z)|^p$; indeed,

$$\Delta|F(z)|^p \geqslant \frac{1}{4}p^2|F(z)|^{p-2}|\partial F(z)|^2 \geqslant 0 \qquad (0<p<2) \tag{1.3}$$

whenever F is analytic and non-zero, as one easily calculates using the *factorization* $\Delta = \partial\bar\partial = \bar\partial\partial$.

For an arbitrary flat Ω having Lipschitz (or smoother) boundary $\partial\Omega$, there is an analogous definition of $H^p(\partial\Omega)$, replacing (1.1) with the usual non-tangential maximal function norm

$$\|F\|_{H^p(\Omega)} = \left(\int_{\partial\Omega} N(F)(s)^p \, ds\right)^{1/p} \tag{1.4}$$

obtained by integrating along $\partial\Omega$ with respect to arc-length measure. But now the boundary regularity theory becomes decidedly more difficult as the smoothness of $\partial\Omega$ decreases. In his thesis Kenig used conformal mapping techniques to reduce the study of a general $H^p(\Omega)$-space to that of a weighted H^p-space on the upper half-plane ([21]), but the L^p-boundedness of the principal-value Cauchy integral operator on $\partial\Omega$ can also be used (cf. Section 3). Using such techniques it can be shown for each Ω there is a constant $p(\Omega)$, $1 \leqslant p(\Omega) < 2$, such that properties (i) and (ii) in (1.2) still hold provided $p > p(\Omega)$. If $\partial\Omega$ is C^1-smooth, then $p(\Omega) = 1$ as in the case of the upper half-plane, but $p(\Omega) > 1$ if $\partial\Omega$ has only Lipschitz smoothness. Results of Kenig and Pipher suggest that there should be an atomic type decomposition of $H^p(\partial\Omega)$ for $p \leqslant p(\Omega)$ and Lipschitz $\partial\Omega$ ([23]).

One way of discovering the relationship between $(\bar\partial, \partial)$ and the Dirac operator for $\mathbb{R}^2$ is to identify the Clifford algebra for $\mathbb{R}^2$ with the quaternions

$$\mathbb{H} = \{x_0 + x_1\boldsymbol{i} + x_2\boldsymbol{j} + x_3\boldsymbol{k} : x_j \in \mathbb{R}\} \tag{1.5}$$

where

$$\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = -1, \qquad \boldsymbol{ij} = \boldsymbol{k} = -\boldsymbol{ji} \tag{1.6}$$

Thus $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ are imaginary units; in particular,

$$\mathbb{C} = \{x_0 + \boldsymbol{j}x_2 : x_j \in \mathbb{R}\}, \qquad \mathbb{H} = \{z + \boldsymbol{i}w : z, w \in \mathbb{C}\}. \tag{1.7}$$

Under the embedding of $\mathbb{R}^2$ into $\mathbb{H}$ given by

$$(x, y) \to x\boldsymbol{i} + y\boldsymbol{k} = \boldsymbol{i}(x + \boldsymbol{j}y),$$

the standard Dirac operator takes the form

$$D = \boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{k}\frac{\partial}{\partial y} = \boldsymbol{i}\bar\partial = \partial\boldsymbol{i}, \tag{1.8}$$

setting

$$\bar\partial = \frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y}, \qquad \partial = \frac{\partial}{\partial x} - \boldsymbol{j}\frac{\partial}{\partial y}. \tag{1.8$'$}$$

Now let $F = F(x, y) = \Phi + \boldsymbol{i}\Psi$ be an $\mathbb{H}$-valued function on Ω, where Φ, Ψ are complex-valued functions determined by the decomposition $\mathbb{H} = \mathbb{C} \oplus \boldsymbol{i}\mathbb{C}$ of $\mathbb{H}$. Then

$$DF = \boldsymbol{i}\frac{\partial F}{\partial x} + \boldsymbol{k}\frac{\partial F}{\partial y} = \boldsymbol{i}\bar\partial\Phi - \partial\Psi. \tag{1.9}$$

Hence $DF = 0$ on Ω if and only if $\bar{\partial}\Phi = 0$, $\partial\Psi = 0$; and so $\bar{\partial}$ (resp. ∂) arises from restricting the Dirac operator to $\mathbb{H}$-valued functions having range in $\mathbb{C}$ (resp. $i\mathbb{C}$). Notice that these range spaces $\mathbb{C}$ and $i\mathbb{C}$ are just the respective $+1$ and -1 eigenspaces of the involution $z + iw \to z - iw$ on $\mathbb{H}$; notice also that if Φ is $\mathbb{C}$-valued, then $D\Phi$ is $i\mathbb{C}$-valued, whereas $D\Phi$ is $\mathbb{C}$-valued when Φ is $i\mathbb{C}$-valued, so that the effect of D is to interchange the eigenspaces of an involution. This is exactly how the important geometric differential operators associated with the Dirac operator on more general manifolds are defined (cf. [14, Chap. V]; [19]).

There is another interesting way of obtaining $(\bar{\partial}, \partial)$ from the Dirac operator by realizing $\mathbb{H}$ as the real sub-algebra

$$\mathbb{H} = \left\{ \begin{bmatrix} z_1 & -\bar{z}_2 \\ z_2 & \bar{z}_1 \end{bmatrix} : z_1, z_2 \in \mathbb{C} \right\} \tag{1.10}$$

of $\mathbb{C}^{2\times 2}$ where

$$i = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \qquad j = \begin{bmatrix} \sqrt{-1} & 0 \\ 0 & -\sqrt{-1} \end{bmatrix}, \qquad k = \begin{bmatrix} 0 & \sqrt{-1} \\ \sqrt{-1} & 0 \end{bmatrix}.$$

For then

$$D = i\frac{\partial}{\partial x} + k\frac{\partial}{\partial y} = \begin{bmatrix} 0 & -\partial \\ \bar{\partial} & 0 \end{bmatrix}, \tag{1.11}$$

and if $F = \left[\begin{smallmatrix} \Phi \\ \Psi \end{smallmatrix}\right]$ is a column vector of complex-valued functions, then

$$DF = \begin{bmatrix} 0 & -\partial \\ \bar{\partial} & 0 \end{bmatrix} \begin{bmatrix} \Phi \\ \Psi \end{bmatrix} = \begin{bmatrix} -\partial\Psi \\ \bar{\partial}\Phi \end{bmatrix}.$$

Hence $DF = 0$ if and only if $\bar{\partial}\Phi = 0$, $\partial\Psi = 0$; once again the effect of D is to interchange the eigenspaces of an involution, this time the involution being defined by

$$\begin{bmatrix} a \\ b \end{bmatrix} \to \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \\ -b \end{bmatrix}. \tag{1.12}$$

Generalizations of this derivation of $(\bar{\partial}, \partial)$ from D use the so-called Spin representations of Clifford algebras and are valid on arbitrary spin manifolds ([14, chap. V]; [5, Section 5]).

As a constant coefficient operator in the flat case, $(\bar{\partial}, \partial)$ commute with translations; but, more importantly, they also commute with the rotation operators $F \mapsto F_\theta$, $F_\theta(z) = F(e^{i\theta}z)$, in the sense that

$$\bar{\partial}(F_\theta) = e^{i\theta}(\bar{\partial}F)_\theta, \qquad \partial(F_\theta) = e^{-i\theta}(\partial F)_\theta. \tag{1.13}$$

These properties of invariance under the rigid motions of the complex plane are the reason why analytic function theory can be incorporated so successfully into Fourier series and Fourier Transform theory, and ultimately therefore for the importance of H^p-theory. On the other hand, when Ω has curvature, the Cauchy-Riemann operator is invariant with respect to the semi-simple automorphism group G of Ω, and hence is the reason for the extensive use of analytic function theory in realizing representations of such G. Our emphasis on the invariance properties of operators of Dirac type is thus well-founded.

Let us now study in greater detail the elliptic boundary value problems associated with $H^p(\Omega)$ for a domain in $\mathbb{C}$. Qualitatively, the classical Riemann-Hilbert problem can be formulated as

Riemann-Hilbert Problem 1.14. *Given real-valued functions a, b, and f on $\partial\Omega$, find a function F in $\Omega \cup \partial\Omega$ such that*

(i) $\bar{\partial}F = 0$ *on* Ω
(ii) $L_0(F) = \mathrm{Re}\,((a - ib)F) = f$ *on* $\partial\Omega$.

The Dirichlet problem, of course, is the special case $a = 1$, $b = 0$. In real terms the general boundary condition $L_0(F) = f$ is

$$a(z)U(z) + b(z)V(z) = f(z) \qquad (z \in \partial\Omega)$$

when $F = U + iV$, and so (1.14)(ii) imposes a linear relation on the real and imaginary parts of F at each point of $\partial\Omega$. This was how Riemann first formulated the problem. Thus from a conceptual point of view, the problem must exploit the analytic over-determinedness of $\bar{\partial}$ if there is to be any hope of a solution existing and having some degree of uniqueness. By allowing more general boundary conditions in (1.14)(ii), we get other classical problems as special cases. For instance, given complex-valued functions A_0, A_1 on $\partial\Omega$, F can be required to satisfy the boundary condition

$$\text{(1.14)(ii)}' \qquad L_1(F) = \mathrm{Re}\,(\overline{A_0(z)}F(z) + A_1(z)\,\partial F(z)) = f(z) \qquad (z \in \partial\Omega)$$

on $\partial\Omega$; here ∂F denotes the boundary values of the derivative ∂F of F in Ω. The so-called Generalized Riemann-Hilbert problem imposes conditions such as (1.14)(ii)′ on F and ∂F, as well as possibly on higher order derivatives $\partial^2 F, \ldots, \partial^m F$ (and integral terms too) (cf. [12]; [26]).

To study the boundary condition $L_1(F) = f$ in detail, let $t = t(s)$ be the parameterization of $\partial\Omega$ by arc length, and let $\theta = \theta(s)$ be the angle between the positive tangent to $\partial\Omega$ at t and the x-axis. Thus $e^{i\theta(s)} = t'(s)$. In a famous study of the theory of tides, Poincaré considered the following problem.

Poincaré Problem 1.15. *Given real-valued functions ϕ, ψ, c and f on $\partial\Omega$, find a real-valued function U on $\Omega \cup \partial\Omega$ satisfying*

(i) $\Delta U = 0$ *on* Ω

(ii) $\phi(s)\dfrac{\partial U}{\partial s} + \psi(s)\dfrac{\partial U}{\partial n} + c(s)U = f$ *on* $\partial\Omega$

where $\partial U/\partial s$ is the tangential derivative and $\partial U/\partial n$ is the normal derivative of U at $t = t(s)$.

This contains both the Dirichlet and Neumann problems, as well as the Oblique-derivative problem, as special cases. But on setting

$$A_0(s) = c(s), \qquad A_1(s) = \frac{1}{2}e^{i\theta(s)}(\phi + i\psi)(s) \qquad (t(s) \in \partial\Omega),$$

it is easily checked that the Poincaré problem can be solved by finding an analytic function $F = U + iV$ satisfying $L_1(F) = f$ on $\partial\Omega$.

In the Russian literature (cf. [12]; [26]) the Riemann-Hilbert problem and its variants were most often studied quantitatively in the context of open sets Ω whose boundary is (possibly piecewise-) $C^{1+\epsilon}$, with the boundary functions being in Lip (α)-spaces. The extra smoothness on $\partial\Omega$ ensured that Fredholm theory could be applied and that the principal value Cauchy integral operator was bounded on Lip (α) (Privalov's theorem). As we now know, $C^{1+\epsilon}$-smoothness is not essential for this last property, but the Riemann-Hilbert problem does not appear to have been studied for general Lipschitz domains. For concreteness therefore, we shall assume that $\partial\Omega$ is $C^{1+\epsilon}$-smooth for some $\epsilon > 0$ and replace Lip (α)-spaces by boundary L^p-spaces, $1 < p < \infty$. We will however assume that $A = A(z)$ is a complex-valued function on $\partial\Omega$ which is in Lip (α) for some $\alpha > 0$. In this setting (1.14) becomes what we shall call

Hardy-Riemann-Hilbert Problem 1.16. *Given a real-valued function f in $L^p(\partial\Omega)$, find F in $H^p(\Omega)$ such that* $\mathrm{Re}\,(\overline{A(z)}F(z)) = f(z)$ *on* $\partial\Omega$.

In view of the boundary regularity theory for H^p-spaces, this problem amounts to solving the equation

$$\mathfrak{I}_A(F) = f$$

where

$$\mathfrak{I}_A : H^p(\partial\Omega) \to L^p(\partial\Omega, \mathbb{R}), \qquad (\mathfrak{I}_A F)(z) = \mathrm{Re}\,(\overline{A(z)}F(z)), \tag{1.17}$$

and $L^p(\partial\Omega, \mathbb{R})$ is the Lebesgue space of real-valued functions on $\partial\Omega$. From the well-known solution of (1.14) for Lip (α)-spaces (cf. [12]; [26]) we obtain

Theorem 1.18. *If $A = A(z)$ is non-vanishing on ∂Q, then $\mathfrak{I}_A : H^p(\partial\Omega) \to L^p(\partial\Omega, \mathbb{R})$ is a Fredholm operator with index*

$$\text{Index}\,(\mathfrak{I}_A) = 2\omega_A + 1$$

where ω_A is the winding number of A.

Since both the kernel and co-kernel of $\mathfrak{I}_A$ are known explicitly, the Hardy-Riemann-Hilbert problem is thus completely solved. Extensions to less smooth domains and other function classes for A ($A \in BMO(\partial\Omega)$?) would clearly be of interest. The Riemann-Hilbert problem has been formulated and extensively studied for $\mathfrak{H}$-valued functions also (cf., for instance, [35]). On the other hand, the Riemann-Roch for compact Riemann surfaces relating the index of $\bar{\partial}$ to the genus of the surface is another of the classical index problems associated with analytic functions and was a particularly influential example in the development of the Atiyah-Singer Index Theorem.

The study of Hardy spaces and Bergman spaces of analytic functions is also the starting point for the realization of unitary representations of semi-simple Lie groups. Indeed the group

$$SL(2, \mathbb{R}) = \left\{ g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} : \det(g) = 1 \right\}$$

of 2×2 real matrices having determinant 1 acts on the compactified plane $\mathbb{C} \cup \{\infty\}$ by fractional linear transformations

$$g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} : z \mapsto z \cdot g = \frac{b + zd}{a + cz} \qquad (z \in \mathbb{C}). \tag{1.19}$$

Its orbits consist of the upper and lower half-planes $\Omega_\pm$ in $\mathbb{C}$ together with their compactified mutual boundary $\mathbb{R} \cup \{\infty\}$. There is also an induced action of $SL(2, \mathbb{R})$ on the Hardy $H^2(\Omega_+)$-space and on the Lebesgue space $L^2(\mathbb{R})$ of complex-valued functions on $\mathbb{R}$:

$$(\pi(g)F)(z) = \frac{1}{a + zc} F(z \cdot g) \qquad (z \in \Omega_+) \tag{1.20 (i)}$$

defines a unitary representation of $SL(2, \mathbb{R})$ on $H^2(\Omega_+)$, while

$$(\sigma(g)f)(v) = \frac{1}{a + vc} f(v \cdot g) \qquad (v \in \mathbb{R}) \tag{1.20 (ii)}$$

defines a unitary representation on $L^2(\mathbb{R})$. There are several implications of these results for Euclidean harmonic analysis:

(i) the familiar transformations of translation, «rotation» and dilation on $\mathbb{R}$ all arise as special cases of (1.19), as does inversion $v \to -1/v$;

(ii) the Cauchy integral

$$f \to F(z) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{f(v)}{v - z}\, dv = \mathcal{C}f(z)$$

intertwines the representations in (1.20) in the sense that $\pi(g)F = \mathcal{C}(\sigma(g)f)$;

(iii) the Hilbert transform

$$\mathcal{H}f(x) = P.V. \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f(v)}{x - v}\, dv$$

intertwines the representation in (1.20)(ii) in the sense that $\sigma(g)(\mathcal{H}f) = \mathcal{H}(\sigma(g)f)$.

Hence all the familiar and important results in H^2-theory have a group-theoretic interpretation in terms of representations of $SL(2, \mathbb{R})$. But $SL(2, \mathbb{R})$ is just the two-fold covering of the conformal group of $\mathbb{R}$; in addition, if Ω_+ is given the Poincaré metric, then $SL(2, \mathbb{R})$ is just the two-fold covering of the (identity component of the) isometry group of Ω_+. A discussion of H^2-theory in higher dimensions, therefore, should naturally include corresponding representations of the two-fold covering of the conformal group of $\mathbb{R}^n$ and of the isometry group of $(n + 1)$-dimensional hyperbolic space. This will be done in Section 6 as part of a general development of Hardy H^p-theory on Euclidean space. But first a detailed analytic and algebraic study of the higher-dimensional replacements for the Cauchy-Riemann $(\bar{\partial}, \partial)$-operators has to have been made. These are the Dirac $\mathfrak{D}$ and $\bar{\mathfrak{D}}$ operators, and the dicussion in Sections 2, 3 and 4 show that they are intrinsically associated with Euclidean space. Then in Section 5 we show how the operators $\mathfrak{D}$ and $\bar{\mathfrak{D}}$ lead to operator ∂_τ associated with any irreducible representation of the Euclidean rotation group, before finally turning to aspects of the general theory in Section 6.

2. The Dirac $\mathfrak{D}$-operator

Let $\mathfrak{H}$ be a finite-dimensional real or complex Hilbert space and let $\mathfrak{L}(\mathfrak{H})$ be the usual C^*-algebra of bounded linear operators on $\mathfrak{H}$, the adjoint of an operator a in $\mathfrak{L}(\mathfrak{H})$ being denoted by $\bar{a}$ (cf. [18]). The norm on $\mathfrak{H}$ will be denoted by $|\cdot|$, the inner product by $(\cdot, \cdot)$, and the operator norm by $\|\cdot\|$; thus

$$(au, v) = (u, \bar{a}u), \qquad \|\bar{a}a\| = \|a\|^2 \tag{2.1}$$

hold for all a in $\mathfrak{L}(\mathfrak{H})$ and u, v in $\mathfrak{H}$. Now fix skew-adjoint elements $e_1, \ldots, e_n$ in $\mathfrak{L}(\mathfrak{H})$ such that

$$e_j e_k + e_k e_j = -2\delta_{jk} e_0 \qquad (1 \leqslant j, k \leqslant n) \tag{2.2}$$

where e_0 $(= I)$ is the identity operator on $\mathfrak{H}$. For instance, $\mathfrak{H}$ could be the complex numbers $\mathbb{C}$ or quaternions $\mathbb{H}$; then multiplication on $\mathbb{C}$ by the imaginary unit i $(= \sqrt{-1})$ or on $\mathbb{H}$ by its imaginary units $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ defines skew-adjoint operators satisfying (2.2). More generally still, scalar multiplication

$$i\colon \xi \mapsto i\xi \qquad (\xi \in \mathfrak{H}) \tag{2.3}$$

by i $(= \sqrt{-1})$ on any finite-dimensional complex Hilbert space $\mathfrak{H}$ is skew-adjoint.

Definition 2.4. *The Dirac $\mathfrak{D}$-operator and its adjoint $\bar{\mathfrak{D}}$ are the first-order systems of differential operators on $C^\infty(\Omega, \mathfrak{H})$ defined by*

$$\mathfrak{D} = \sum_{j=0}^{n} e_j \frac{\partial}{\partial x_j}, \qquad \bar{\mathfrak{D}} = \sum_{j=0}^{n} \bar{e}_j \frac{\partial}{\partial x_j}$$

for any open set Ω in $\mathbb{R}^{n+1}$.

When $\mathfrak{H}$ is any complex Hilbert space, for example, $\mathfrak{D}$ and $\bar{\mathfrak{D}}$ are just the classical Cauchy-Riemann operators $\bar{\partial}$ and ∂ respectively, while $\mathfrak{D}$ is the Fueter operator when $\mathfrak{H} = \mathbb{H}$ ([6, Chap. 2]). As the notation and terminology suggest, the skew-adjointness of the e_j ensure that $\mathfrak{D}$ and $\bar{\mathfrak{D}}$ are adjoint in the sense that

$$\int_\Omega (\mathfrak{D}F(x), \Phi(x))\, dx = \int_\Omega (F(x), \bar{\mathfrak{D}}\Phi(x))\, dx \tag{2.5}$$

when F, Φ are compactly supported functions in $C^\infty(\Omega, \mathfrak{H})$. On the other hand, since

$$\mathfrak{D}\bar{\mathfrak{D}} = \sum_{j=0}^{n} \left(\frac{\partial}{\partial x_j}\right)^2 = \bar{\mathfrak{D}}\mathfrak{D}, \tag{2.6}$$

every solution of $\mathfrak{D}F = 0$ or $\bar{\mathfrak{D}}F = 0$ has harmonic components. Consequently, $\mathfrak{D}$ and $\bar{\mathfrak{D}}$ are Generalized Cauchy-Riemann (GCR-) systems as introduced by Stein-Weiss ([33, p. 231]); in particular, they are *injectively elliptic* meaning that the symbol mappings

$$\xi \to \left(\sum_{j=0}^{n} \lambda_j e_j\right)\xi, \qquad \xi \to \left(\sum_{j=0}^{n} \lambda_j e_j\right)^{-} \xi \qquad (\xi \in \mathfrak{H}) \tag{2.7}$$

are injective on $\mathfrak{H}$ for every non-zero $\lambda = (\lambda_0, \ldots, \lambda_n)$ in $\mathbb{R}^{n+1}$ (cf. [10, p. 79]; [33, p. 231]). In fact, as we shall see, the algebraic structure in (2.2) forces the mappings in (2.7) to be surjective as well as injective, and so $\mathfrak{D}$, $\bar{\mathfrak{D}}$ are *elliptic first-order systems* in the usual sense of the term elliptic. But the consequences of this algebraic structure go much deeper.

Denote by $\nu: \mathbb{R}^{n+1} \to \mathfrak{L}(\mathfrak{H})$ the linear embedding

$$\nu: x = (x_0, \ldots, x_n) \to \sum_{j=0}^{n} x_j e_j \qquad (x \in \mathbb{R}^{n+1}) \tag{2.8}$$

of $\mathbb{R}^{n+1}$ into $\mathfrak{L}(\mathfrak{H})$. Then, because of (2.2),

$$\nu(x)^- \nu(x) = |x|^2 e_0 = \nu(x)\nu(x)^- \qquad (x \in \mathbb{R}^{n+1})$$

where $|x| = (\Sigma_j |x_j|^2)^{1/2}$ is the Euclidean length of x; on the other hand,

$$\|\nu(x)\|^2 = \|\nu(x)^- \nu(x)\| = |x|^2 \|e_0\| = |x|^2 \tag{2.9}$$

because of the C^*-algebra norm property. Hence $\nu: \mathbb{R}^{n+1} \to \mathfrak{L}(\mathfrak{H})$ is an *isometric embedding of* $\mathbb{R}^{n+1}$ *into* $\mathfrak{L}(\mathfrak{H})$; we shall therefore identify $\mathbb{R}^{n+1}$ with a subspace of $\mathfrak{L}(\mathfrak{H})$, regarding $e_0, \ldots, e_n$ simply as an orthonormal basis for $\mathbb{R}^{n+1}$ and omitting any mention of the mapping ν in (2.8). With this convention, let $\mathfrak{A}(\mathfrak{H})$ be the (real) $\mathcal{C}^*$-subalgebra of $\mathfrak{L}(\mathfrak{H})$ generated by $\mathbb{R}^{n+1}$ ([18]). Now, in the special cases of $\mathfrak{H} = \mathbb{C}$ or $\mathbb{H}$,

$$\mathfrak{A}(\mathfrak{H}) = \mathbb{C} \sim \mathbb{R}^2, \qquad \mathfrak{A}(\mathfrak{H}) = \mathbb{H} \sim \mathbb{R}^4 \tag{2.10}$$

and

$$\|wz\| = \|w\| \, \|z\| = |w| \, |z| \qquad (w, z \in \mathfrak{A}(\mathfrak{H})); \tag{2.11}$$

consequently, in addition to the linear structure, there is a multiplicative structure on $\mathbb{R}^2$ and $\mathbb{R}^4$ with respect to which the metric is multiplicative, *i.e.*, satisfies (2.11). Although this property characterizes $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{H}$ as C^*-algebras ([18, p. 78]), it will be exceedingly important and highly instructive to think of $\mathfrak{A}(\mathfrak{H})$ as being a sufficiently strong substitute for any $\mathfrak{H}$. Indeed, by (2.2),

$$\|wz\| = \|(wz)^-(wz)\| = |w| \, |z| \qquad (w, z \in \mathbb{R}^{n+1}), \tag{2.12 (i)}$$

and more generally,

$$\|a \cdots wz\| = |a| \cdots |w| \, |z| \tag{2.12 (ii)}$$

for any finite product $a \cdots wz$ in $\mathfrak{A}(\mathfrak{H})$ of elements $a, \ldots, w, z$ in $\mathbb{R}^{n+1}$. What fails in general is the multiplicativity on linear combinations of such finite

products. Nonetheless, all of this and Adams' solution of the vector fields on spheres problem are much more than is needed to establish the following results.

Theorem 2.13. *The Dirac $\mathfrak{D}$- and $\bar{\mathfrak{D}}$-operators are first-order elliptic systems of differential operators on $C^\infty(\Omega, \mathfrak{H})$ for any open set Ω in $\mathbb{R}^{n+1}$; in particular, if $n = 8d + r$, $0 \leqslant r \leqslant 7$, then*

(i) $\dim_{\mathbb{R}}(\mathfrak{H}) \geqslant (r+1)16^d$
(ii) $\dim_{\mathbb{C}}(\mathfrak{H}) \geqslant 2^{[r/2]}16^d$.

PROOF. Since

$$\left(\sum_{j=0}^{n} \lambda_j e_j\right)\left(\sum_k \lambda_k e_k\right)^- = |\lambda|^2,$$

it is clear that the symbol mappings

$$\xi \to \left(\sum_{j=0}^{n} \lambda_j e_j\right)\xi, \qquad \xi \to \left(\sum_{j=0}^{n} \lambda_j e_j\right)^- \xi$$

are invertible on $\mathfrak{H}$ for each $\lambda \neq 0$. Hence $\mathfrak{D}$ and $\bar{\mathfrak{D}}$ are elliptic. As a consequence of this ellipticity, the dimension of $\mathfrak{H}$ cannot be too small in relation to n (cf. [1], [30]). More precisely, if

$$\dim_{\mathbb{F}}(\mathfrak{H}) = (2k+1)2^c 16^\delta \qquad (0 \leqslant c \leqslant 3, k \geqslant 0),$$

then

(i) $n + 1 \leqslant 8\delta + 2^c \quad (\mathbb{F} = \mathbb{R})$,
(ii) $n + 1 \leqslant 8\delta + 2c + 2 \quad (\mathbb{F} = \mathbb{C})$.

Now suppose $n = 8d + r$, $0 \leqslant r \leqslant 7$. Then

$$\dim_{\mathbb{R}}(\mathfrak{H}) \geqslant (r+1)16^d, \qquad \dim_{\mathbb{C}}(\mathfrak{H}) \geqslant 2^{[r/2]}16^d,$$

completing the proof. □

The significance of these estimates for $\dim_{\mathbb{R}}(\mathfrak{H})$ will become clear later. Let

$$C(x) = \frac{1}{\omega_{n+1}}\left(\frac{\bar{x}}{|x|^{n+1}}\right) \qquad (x \in \mathbb{R}^{n+1}, x \neq 0) \tag{2.14}$$

be the Cauchy kernel on $\mathbb{R}^{n+1}$. It is well-defined as an $\mathfrak{A}(\mathfrak{H})$-valued solution of $\mathfrak{D}\Phi = 0$ on $\mathbb{R}^{n+1}\setminus\{0\}$ since

$$C(x) = \frac{1}{(1-n)\omega_{n+1}}\bar{\mathfrak{D}}\left(\frac{1}{|x|^{n-1}}\right)$$

expresses the Cauchy kernel as a constant multiple of $\bar{\mathfrak{D}}(\Gamma_{n+1})$, where Γ_{n+1} is the fundamental solution of the Laplacian on $\mathbb{R}^{n+1}$. Similarly,

$$d\sigma(x) = \sum_{j=0}^{n} (-1)^j e_j \, dx_0 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_n \tag{2.15}$$

is an $\mathfrak{A}(\mathfrak{H})$-valued n-form on $\mathbb{R}^{n+1}$. The usual argument using Stokes theorem (cf., for intance, [6, p. 51]; [14, Chap. II, Section 3]) then gives

Theorem 2.16. *Let M be a compact set in Ω having suitably smooth boundary ∂M. Then the Cauchy integral*

$$f \to \mathcal{C}f(x) = \frac{1}{\omega_{n+1}} \int_{\partial M} \frac{\bar{y} - \bar{x}}{|y - x|^{n+1}} \, d\sigma(y) f(y)$$

of any f in $C(\partial M, \mathfrak{H})$ satisfies $\mathfrak{D}(\mathcal{C}f) = 0$ on $\mathbb{R}^{n+1} \backslash \partial M$, while

$$\frac{1}{\omega_{n+1}} \int_{\partial M} \frac{\bar{y} - \bar{x}}{|y - x|^{n+1}} \, d\sigma(y) F(y) = \begin{cases} F(x) & (x \in M \backslash \partial M) \\ 0 & (x \in \Omega \backslash M) \end{cases}$$

whenever $\mathfrak{D}F = 0$ on Ω.

3. Hardy Spaces

In complete analogy with the classical case, the Cauchy integral results lead naturally to Hardy spaces on domains in $\mathbb{R}^{n+1}$.

Let Ω be a domain in $\mathbb{R}^{n+1}$ which is either a bounded Lipschitz domain or a special Lipschitz domain in the unbounded case. Thus for each x in $\partial\Omega$ having inward unit normal $\eta(x)$, there is a cone

$$\Gamma_\alpha(x) = \{ z \in \mathbb{R}^{n+1} : |z - x| < \delta, (z - x, \eta(x)) > \alpha |z - x| \} \tag{3.1}$$

having vertex at x and lying wholly inside Ω for some choice of α, δ independent of x. Now denote by $\mathcal{HC}^p(\Omega, \mathfrak{H})$ the *harmonic Hardy space* of all harmonic functions in $C^\infty(\Omega, \mathfrak{H})$ whose non-tangential maximal function

$$N(F)(x) = \sup_{z \in \Gamma_\alpha(x)} |F(z)| \qquad (x \in \partial\Omega) \tag{3.2}$$

is L^p-integrable on $\partial\Omega$. Since the Laplacian is elliptic, $\mathcal{HC}^p(\Omega, \mathfrak{H})$ is a Banach space under the norm

$$\|F\|_p = \left(\int_{\partial\Omega} N(F)(x)^p \, dS \right)^{1/p} \qquad (1 \leqslant p < \infty) \tag{3.3}$$

defined with respect to scalar surface measure on $\partial\Omega$.

Definition 3.4. *The Hardy* $H^p(\Omega, \mathfrak{H})$ *consists of all solutions of* $\mathfrak{D}F = 0$ *in* $C^\infty(\Omega, \mathfrak{H})$ *for which* (3.3) *is finite.*

There are corresponding spaces $\mathcal{HC}^p(\Omega, \mathfrak{H})$ and $H^p(\Omega, \mathfrak{H})$ when $0 < p < 1$, though they are not Banach spaces of course. Ellipticity and GCR-properties of $\mathfrak{D}$ ensure that $H^p(\Omega, \mathfrak{H})$ is a closed subspace of $\mathcal{HC}^p(\Omega, \mathfrak{H})$.

The key to studying Hardy spaces is the L^p-boundedness of the principal-value Cauchy integral

$$\mathcal{HC}f(x) = P.V.\left(\frac{2}{\omega_{n+1}} \int_{\partial\Omega} \frac{\bar{y} - \bar{x}}{|y - x|^{n+1}} d\sigma(y) f(y)\right) \qquad (x \in \partial\Omega) \tag{3.5}$$

for $\mathfrak{H}$-valued functions on $\partial\Omega$. In the seminal work of Coifman, McIntosh and Meyer ([7]) establishing this boundedness for *complex-valued* functions through use of P_t, Q_t-operator techniques, essential though unconscious use is made of the multiplicativity property $|wz| = |w|\,|z|$ for all w, z in $\mathbb{C}$ in controlling the operator norm of iterated powers of P_t and Q_t. The fundamental multiplicative norm property (2.12) for arbitrary $\mathfrak{H}$ maintains the same control on the corresponding operators in the general case; one then proves (cf. [14])

Theorem 3.6. *Let* Ω *be a special Lipschitz domain in* $\mathbb{R}^{n+1}$. *Then for each* p, $1 < p < \infty$,

(i) *the principal-value Cauchy integral operator* $f \to \mathcal{HC}f$ *is bounded on* $L^p(\partial\Omega, \mathfrak{H})$;

(ii) *the Cauchy integral* $f \to \mathcal{C}f$ *maps* $L^p(\partial\Omega, \mathfrak{H})$ *boundedly into* $H^p(\Omega, \mathfrak{H})$;

(iii) *the Cauchy integral* $\mathcal{C}f$ *of any* f *in* $L^p(\partial\Omega, \mathfrak{H})$ *has non-tangential boundary values on* $\partial\Omega$ *such that*

$$\lim_{z \to x} \mathcal{C}f(z) = \frac{1}{2}(f(x) + \mathcal{HC}f(x))$$

almost everywhere on $\partial\Omega$.

Because of the GCR-property enjoyed by $\mathfrak{D}$, the solution of the Dirichlet problem for Ω with scalar-valued L^p-boundary data can be applied componentwise to the Hardy spaces in studying the boundary value behaviour of functions in $H^p(\Omega, \mathfrak{H})$ (cf. [22]). Let $p(\Omega)$ be the critical index for the scalar Dirichlet problem with L^p-data, *i.e.*, for each real-valued function f in $L^p(\partial\Omega)$, $p > p(\Omega)$, there is a function U so that

$$\text{(3.7) (i)} \qquad \Delta U = 0 \quad \text{in} \quad \Omega$$

$$\text{(ii)} \qquad U|_{\partial\Omega} = f.$$

In the case of any C^1-domain, $p(\Omega) = 1$; but in general, $p(\Omega) > 1$. By a simple adaptation of the proof of Theorem 3.19 in Chapter 2 in [33], one readily sees that each F in $H^p(\Omega, \mathfrak{H})$, $p > p(\Omega)$, has non-tangential boundary values F^+ in $L^p(\partial\Omega, \mathfrak{H})$.

Definition 3.8. *Denote by $H(\partial\Omega, \mathfrak{H})$ the subspace in $L^p(\partial\Omega, \mathfrak{H})$ of the non-tangential boundary values F^+ of all functions F in $H^p(\Omega, \mathfrak{H})$, $p > p(\Omega)$.*

In [15] we proved the following result characterizing this space $H^p(\partial\Omega, \mathfrak{H})$ of boundary values.

Theorem 3.9. *Let Ω be a special Lipschitz domain in $\mathbb{R}^{n+1}$. Then, for each p, $p(\Omega) < p < \infty$,*

(i) *the Cauchy integral $f \to \mathcal{C}f$ is a bounded mapping from $L^p(\partial\Omega, \mathfrak{H})$ onto $H^p(\Omega, \mathfrak{H})$;*
(ii) *every F in $H^p(\Omega, \mathfrak{H})$ is the Cauchy integral $\mathcal{C}F^+$ of its non-tangential boundary values F^+ in $H^p(\partial\Omega, \mathfrak{H})$;*
(iii) *a function f in $L^p(\partial\Omega, \mathfrak{H})$ is in $H^p(\partial\Omega, \mathfrak{H})$ if and only if $f = \mathcal{K}f$.*

Two results are conspicuously absent from (3.9): the first dealing with characterizations for $p \leqslant p(\Omega)$, and the second with an analytic over-determinedness characterization such as (1.2)(ii). Carlos Kenig has remarked to us that the ideas in his joint work with Jill Pipher ([23]) may well yield characterizations of $H^p(\partial\Omega, \mathfrak{H})$ below the critical index $p(\Omega)$. As regards the second omission, let us assume that for each imaginary unit $\boldsymbol{j}$ in $\mathbb{R}^{n+1}$ ($\subseteq \mathfrak{A}(\mathfrak{H})$) there is a splitting $\mathfrak{K} = \mathfrak{H}_0 \oplus \boldsymbol{j}\mathfrak{H}_0$ of $\mathfrak{H}$ for some subspace $\mathfrak{H}_0$ of $\mathfrak{H}$. We shall then say that $\mathfrak{H}$ has the *splitting property*. From the structural characterizations of $\mathfrak{H}$ described in the next section, it likely follows that every $\mathfrak{H}$ has this property. Notice that each such splitting determines an involution $x + \boldsymbol{j}y \to x - \boldsymbol{j}y$ on $\mathfrak{H}$. The prototypical example of course is the splitting $\mathbb{C} = \mathbb{R} \oplus \sqrt{-1}\,\mathbb{R}$ where the involution is complex conjugation. For this reason we shall denote by Re, Im: $\mathfrak{H} \to \mathfrak{H}_0$ the linear mappings

$$z \to x = \mathrm{Re}\,(z), \qquad z \to y = \mathrm{Im}\,(z), \qquad z = x + \boldsymbol{j}y \qquad (x, y \in \mathfrak{H}_0) \tag{3.10}$$

that a splitting $\mathfrak{H} = \mathfrak{H}_0 \oplus \boldsymbol{j}\mathfrak{H}_0$ determines.

Conjecture 3.11. (Real H^p-theory.) *Suppose $\mathfrak{H} = \mathfrak{H}_0 \oplus \boldsymbol{j}\mathfrak{H}_0$ is a splitting of $\mathfrak{H}$. Then, for each p, $p(\Omega) < p < \infty$,*

(i) *the Cauchy integral $f \to \mathcal{C}f$ is an isomorphism from $L^p(\partial\Omega, \mathfrak{H}_0)$ onto $H^p(\Omega, \mathfrak{H})$, and conversely,*
(ii) *the boundary operator $F \to \mathrm{Re}\,(F|_{\partial\Omega})$ is continuous from $H^p(\Omega, \mathfrak{H})$ onto $L^p(\partial\Omega, \mathfrak{H}_0)$.*

A Fourier Transform argument confirms this conjecture for special choices of $\mathfrak{H}$ when Ω is the upper half-space

$$\Omega_+ = \{x + jy : x \in \mathbb{R}^n, y > 0\} \tag{3.12}$$

in $\mathbb{R}^{n+1}$ (cf. [14, Chap. II]).

4. Clifford Algebras and $\mathfrak{A}(\mathfrak{H})$

The definition of $\mathfrak{A}(\mathfrak{H})$ ensures that it is always a C^*-algebra representation of the Universal Clifford algebra $\mathfrak{A}_n$ for $\mathbb{R}^n$; algebraically, therefore, $\mathfrak{H}$ is simply an ungraded Clifford module ([4]). This makes available all the algebraic constructions associated with both of these classes of algebras, and we shall use freely results for these algebras established in [14] and [18] to which the reader is referred for all unsubstantiated assertions (and omitted references). The culmination of this section will be characterizations of $\mathfrak{A}(\mathfrak{H})$ and $\mathfrak{H}$, but the algebraic details should not be allowed to obscure the fact that these characterizations show that $\mathfrak{A}(\mathfrak{H})$ and $\mathfrak{H}$ are intrinsically associated with $\mathbb{R}^n$.

Recall that a Clifford algebra for $\mathbb{R}^n$ consists of a pair $(\mathbb{A}_n, \beta)$ where $\mathbb{A}_n$ is an associative algebra over $\mathbb{R}$ with identity 1 and $\beta: \mathbb{R}^n \to \mathbb{A}_n$ is a linear embedding of $\mathbb{R}^n$ into $\mathbb{A}_n$ such that

$$\begin{aligned} &\text{(i)} && \mathbb{A}_n \textit{ is generated by } \{\beta(x) : x \in \mathbb{R}^n\} \textit{ and } \{\lambda 1 : \lambda \in \mathbb{R}\}, \\ &\text{(ii)} && \beta(x)^2 = -|x|^2 1 \quad (x \in \mathbb{R}^n). \end{aligned} \tag{4.1}$$

The algebra $\mathfrak{A}(\mathfrak{H})$, for instance, always has these properties because of hypothesis (2.2). Clearly,

$$\dim_{\mathbb{R}}(\mathbb{A}_n) \leqslant 2^n. \tag{4.2}$$

Let $(\mathfrak{A}_n, \alpha)$ be a Clifford algebra for $\mathbb{R}^n$ having maximal dimension 2^n. It is universal in the sense that for every $(\mathbb{A}_n, \beta)$ there is a unique algebra homomorphism $\pi: \mathfrak{A}_n \to \mathbb{A}_n$ such that $(\pi \circ \alpha)(x) = \beta(x)$, $x \in \mathbb{R}^n$. Hence $\mathfrak{A}_n$, the so-called Universal Clifford algebra for $\mathbb{R}^n$, is unique up to isomorphism; we shall regard $\mathbb{R}$ and $\mathbb{R}^n$ as subspaces of $\mathfrak{A}_n$, identifying $\mathbb{R}^{n+1}$ with $\mathbb{R} \oplus \mathbb{R}^n$. Universality also ensures that the mapping $x \to -x$ on $\mathbb{R}^n$ extends to a unique automorphism $a \to a'$ of $\mathfrak{A}_n$, *the principal automorphism* of $\mathfrak{A}_n$, and a unique anti-automorphism $a \to \bar{a}$ of $\mathfrak{A}_n$, *conjugation* on $\mathfrak{A}_n$, such that

$$x' = -x = \bar{x} \qquad (x \in \mathbb{R}^n) \tag{4.3}$$

(cf. [14, Chap. I]; [27, p. 245]). The composition $a \to a^* = (a')^- = (\bar{a})'$ of conjugation and the principal automorphism is called the *principal anti-*

automorphism of $\mathfrak{A}_n$. Known basic structural properties of $\mathfrak{A}(\mathfrak{H})$ are collected together in the next theorem. We include its proof for convenience and clarity of exposition.

Theorem 4.4. *The algebra $\mathfrak{A}(\mathfrak{H})$ is a Clifford algebra for $\mathbb{R}^n$ such that the homomorphism $\pi: \mathfrak{A}_n \to \mathfrak{A}(\mathfrak{H})$ satisfies*

(i) *π is always an isomorphism when $n \neq 4l + 3$, $l \geqslant 0$;*
(ii) *π fails to be an isomorphism if and only if $e_1 \cdots e_n = \pm I$;*
(iii) *if π is not an isomorphism, then there is a central idempotent E in $\mathfrak{A}_n$ such that π is an isomorphism from the ideal $\{aE: a \in \mathfrak{A}_n\}$ in $\mathfrak{A}_n$ onto $\mathfrak{A}(\mathfrak{H})$, and* $\ker \pi = \{a(I - E): a \in \mathfrak{A}_n\}$.

Furthermore, whether π is an isomorphism or not,

(iv) $\pi(a)^- = \pi(\bar{a}) \quad (a \in \mathfrak{A}_n)$.

Since algebraically-*isomorphic C^*-algebras are always isometrically-*isomorphic ([18, p. 87]), the C^*-algebras $\mathfrak{A}(\mathfrak{H})$ are all isometrically-*isomorphic when $n \neq 4l + 3$, $l \geqslant 0$, or when $n = 4l + 3$ and the generators $e_1, \ldots, e_n$ satisfy $e_1 \cdots e_n \neq \pm I$. *In fact, they are all C^*-algebra realizations of $\mathfrak{A}_n$ in which the adjoint operation on $\mathfrak{A}(\mathfrak{H})$ coincides with conjugation on $\mathfrak{A}_n$.* The example of $\mathfrak{A}(\mathfrak{H}) = \mathfrak{H} = \mathbb{H}$, with e_1, e_2, e_3 multiplication by the imaginary units $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$, is a case where $e_1 e_2 e_3 = -I$.

PROOF OF THEOREM 4.4. Let $\epsilon_1, \ldots, \epsilon_n$ be generators for $\mathfrak{A}_n$ such that

$$\epsilon_j \epsilon_k + \epsilon_k \epsilon_j = -2\delta_{jk}, \qquad \pi(\epsilon_j) = e_j \qquad (1 \leqslant j \leqslant n)$$

with respect to the homomorphism $\pi: \mathfrak{A}_n \to \mathfrak{A}(\mathfrak{H})$. Now the kernel of π is a two-sided ideal in $\mathfrak{A}_n$, so $\pi: \mathfrak{A}_n \to \mathfrak{A}(\mathfrak{H})$ will be an isomorphism if and only if $\mathfrak{A}_n$ is simple. But $\mathfrak{A}_n$ is always simple when $n \neq 4l + 3$; on the other hand, when $n = 4l + 3$, it is simple if and only if $\epsilon_1 \cdots \epsilon_n \notin \mathbb{R}$ ([14, Chap. I, Corollary 3.6, Theorem 3.19]). This proves (i). It also proves the necessity half of (ii) because

$$e_1 \cdots e_n = \pi(\epsilon_1 \cdots \epsilon_n) = \lambda I \qquad (\lambda \in \mathbb{R})$$

if $\epsilon_1 \cdots \epsilon_n \in \mathbb{R}$; note that $\lambda = \pm 1$ since

$$(e_1 \cdots e_n)^2 = (e_1 \cdots e_n)(e_1 \cdots e_n) = (-1)^{n(n+1)/2} = 1$$

when $n = 4l + 3$. On the other hand,

$$\dim_{\mathbb{R}}(\mathfrak{A}(\mathfrak{H})) < 2^n = \dim_{\mathbb{R}}(\mathfrak{A}_n)$$

when $e_1 \cdots e_n = \pm I$, for then $e_n = \mp e_1 \cdots e_{n-1}$; this ensures that it is not an isomorphism, completing the proof of (ii).

To establish (iii) set

$$E = \tfrac{1}{2}(I + \epsilon_1 \cdots \epsilon_n) \quad \text{or} \quad E = \tfrac{1}{2}(I - \epsilon_1 \cdots \epsilon_n) \tag{4.5}$$

according as $e_1 \cdots e_n = I$ or $-I$. In either case, E is a central idempotent in $\mathfrak{A}_n$ and

$$\mathfrak{A}_n = \{aE + a(I - E) : a \in \mathfrak{A}_n\}$$

where $I_+ = \{aE : a \in \mathfrak{A}_n\}$ and $I_- = \{a(I - E) : a \in \mathfrak{A}_n\}$ are two-sided ideals in $\mathfrak{A}_n$ such that $\mathfrak{A}_n = I_+ \oplus I_-$ (cf. [14, Chap. I]). Now by the definition of E,

$$\pi(aE) = \pi(a), \qquad \pi(a(I - E)) = 0.$$

Consequently, π maps I_+ onto $\mathfrak{A}(\mathfrak{H})$, while $\ker(\pi) \supseteq I_-$. But I_+ and I_- are the only non-trivial, proper, two-sided ideals in $\mathfrak{A}_n$ ([14, Chap. I, Theorem 3.19]). Hence $\pi : I_+ \to \mathfrak{A}(\mathfrak{H})$ is an isomorphism and $\ker(\pi) = I_-$. This proves (iii).

To prove (iv) observe first that

$$(\epsilon_{j_1} \cdots \epsilon_{j_k})^- = (-1)^{k(k+1)/2} \epsilon_{j_1} \cdots \epsilon_{j_k}$$

(cf. [14, Chap. I (2.25)]); a similar result

$$(e_{j_1} \cdots e_{j_k})^- = (-1)^{k(k+1)/2} e_{j_1} \cdots e_{j_k}$$

holds in $\mathfrak{A}(\mathfrak{H})$ since the adjoint operation is an anti-automorphism such that $^-: e_j \to -e_j$. Thus

$$\pi((\epsilon_{j_1} \cdots \epsilon_{j_k})^-) = (\pi(\epsilon_{j_1} \cdots \epsilon_{j_k}))^-$$

holds whether or not π is an isomorphism. Now the products $\epsilon_{j_1} \cdots \epsilon_{j_k}$, $1 \leqslant j_1 < \cdots < j_k \leqslant n$ together with the identity are a basis for $\mathfrak{A}_n$. Hence this last equality extends linearly to all of $\mathfrak{A}_n$, *i.e.*, $\pi(\bar{a}) = \pi(a)^-$, when π is an isomorphism. If, however, π is not an isomorphism, then $n = 4l + 3$ and

$$(\epsilon_1 \cdots \epsilon_n)^- = (-1)^{n(n+1)/2} \epsilon_1 \cdots \epsilon_n = \epsilon_1 \cdots \epsilon_n,$$

ensuring that $\bar{E} = E$. Consequently,

$$\pi((a - aE)^-) = \pi(\bar{a} - \bar{a}E) = 0 \qquad (a \in \mathfrak{A}_n).$$

Hence the equality $\pi(\bar{a}) = \pi(a)^-$ continues to hold on all of $\mathfrak{A}_n$ even if π is not an isomorphism. This completes the proof of (iv). □

A failure of $\mathfrak{A}(\mathfrak{H})$ to be isomorphic to $\mathfrak{A}_n$ when $n = 4l + 3$ is not serious because $\mathfrak{A}(\mathfrak{H})$ can then be embedded in a larger C^*-algebra which is isomorphic to $\mathfrak{A}_n$. For suppose $e_1 \cdots e_n = -I$, say, and regard $\mathfrak{H} \oplus \mathfrak{H}$ as the Hilbert space

$$\text{(4.6) (i)} \qquad \mathfrak{H} \oplus \mathfrak{H} = \left\{ \begin{bmatrix} u \\ v \end{bmatrix} : u, v \in \mathfrak{A} \right\}$$

of column vectors under the inner product

$$\text{(4.6) (ii)} \qquad \left(\begin{bmatrix} u_1 \\ v_1 \end{bmatrix}, \begin{bmatrix} u_2 \\ v_2 \end{bmatrix} \right) = (u_1, u_2) + (u_1, v_2)$$

derived from that on $\mathfrak{H}$. Then

$$\text{(4.7)} \qquad \gamma_1 = \begin{bmatrix} e_1 & 0 \\ 0 & -e_1 \end{bmatrix}, \ldots, \gamma_n = \begin{bmatrix} e_n & 0 \\ 0 & -e_n \end{bmatrix}$$

are skew-adjoint operators on $\mathfrak{H} \oplus \mathfrak{H}$ such that

$$\text{(i)} \ \gamma_j \gamma_k + \gamma_k \gamma_j = -2\delta_{jk} I, \qquad \text{(ii)} \ \gamma_1 \cdots \gamma_n = \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix},$$

since $n = 4l + 3$. Criterion (4.4) (ii) thus ensures that $\mathfrak{A}_n$ is isomorphic to the C^*-subalgebra generated in $\mathfrak{L}(\mathfrak{H} \oplus \mathfrak{H})$ by $\gamma_1, \ldots, \gamma_n$. This clearly is

$$\text{(4.8)} \qquad \left\{ \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} : a, b \in \mathfrak{A}(\mathfrak{H}) \right\} \cong \mathfrak{A}(\mathfrak{H}) \oplus \mathfrak{A}(\mathfrak{H}),$$

acting on the Hilbert space (4.6).

Now every finite-dimensional real C^*-algebra is the finite direct sum of full matrix rings $M(k, \mathbb{K})$ where $\mathbb{K} = \mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$. Since each is of these is simple, it follows that

$$\text{(4.9) (i)} \qquad \mathfrak{A}_n \cong M(k, \mathbb{K}) \qquad (n \neq 4l + 3)$$

for some choice of $(k, \mathbb{K})$, while

$$\text{(4.9) (ii)} \qquad \mathfrak{A}_n \cong M(k, \mathbb{K}) \oplus M(k, \mathbb{K}) \qquad (n = 4l + 3).$$

This is consistent with the well-known realizations

$$\text{(4.10)} \qquad \mathfrak{A}_0 \cong \mathbb{R}, \qquad \mathfrak{A}_1 \cong \mathbb{C}, \qquad \mathfrak{A}_2 \cong \mathbb{H}, \qquad \mathfrak{A}_3 \cong \mathbb{H} \oplus \mathbb{H}$$

and the less well-known

$$\text{(4.10) (i)} \quad \mathfrak{A}_4 \cong \mathbb{H}^{2\times 2}, \quad \mathfrak{A}_5 \cong \mathbb{C}^{4\times 4}, \quad \mathfrak{A}_6 \cong \mathbb{R}^{8\times 8}, \quad \mathfrak{A}_7 \cong \mathbb{R}^{8\times 8} \oplus \mathbb{R}^{8\times 8}.$$

The following fundamental Periodicity Theorem completes the characterization of $\mathfrak{A}_n$ for every n.

Theorem 4.10. (ii) *For each $n \geqslant 0$ there is a realization of $\mathfrak{A}_{n+8}$ as the algebra*

$$\mathfrak{A}_{n+8} = M(16, \mathfrak{A}_n)$$

of all 16×16 matrices having entries from $\mathfrak{A}_n$.

Consequently, when $n = 8d + r$, $0 \leqslant r < 8$, the choice of $\mathbb{K}$ in (4.9) is given by

(4.11) (i) $\mathbb{K} = \mathbb{R}$ $(r = 0, 6, 7)$, $\mathbb{K} = \mathbb{C}$ $(r = 1, 5)$, $\mathbb{K} = \mathbb{H}$ $(r = 2, 3, 4)$,

while k is determined by

(4.11) (ii) $k^2 \dim_{\mathbb{R}}(\mathbb{K}) = 2^{8d+r}$ $(r \neq 3, 7)$, $k^2 \dim_{\mathbb{R}}(\mathbb{K}) = 2^{8d+r-1}$ $(r = 3, 7)$

since $\dim_{\mathbb{R}}(\mathfrak{A}_n) = 2^n$.

In summarizing these results it will be convenient to introduce the following definition.

Definition 4.12. *A finite-dimensional Hilbert space $\mathfrak{H}$ is said to be an $\mathfrak{A}_n$-module when there exist skew-adjoint operators $e_1, \ldots, e_n$ in $\mathfrak{L}(\mathfrak{H})$ such that*

$$e_j e_k + e_k e_j = -2\delta_{jk} e_0 \qquad (1 \leqslant j, k \leqslant n),$$

where e_0 $(= I)$ is the identity operator.

As before, the real C^*-algebra of $\mathfrak{L}(\mathfrak{H})$ generated by $e_1, \ldots, e_n$ will be denoted by $\mathfrak{A}(\mathfrak{H})$. Then for each n there is a real Hilbert space $\mathfrak{R}_n \cong \mathbb{K}^k$, where $(k, \mathbb{K})$ are determined by (4.11), such that

(4.13) (i) $\mathfrak{A}(\mathfrak{H})$ *is isometrically-*isomorphic to the real C^*-algebra $\mathfrak{L}(\mathfrak{R}_n)$ when $n \neq 4l + 3$,*

(4.13) (ii) $\mathfrak{A}(\mathfrak{H})$ *is isometrically-*isomorphic to $\mathfrak{L}(\mathfrak{R}_n)$ or to $\mathfrak{L}(\mathfrak{R}_n) \oplus \mathfrak{L}(\mathfrak{R}_n)$ when $n = 4l + 3$.*

This space $\mathfrak{R}_n$ is called the space of *real spinors*. By identifying $\mathfrak{A}_n$ with $\mathfrak{R}_n \otimes \mathfrak{R}_n$ when $n \neq 4l + 3$, or with $\mathfrak{R}_n \otimes \mathfrak{R}_n \oplus \mathfrak{R}_n \otimes \mathfrak{R}_n$ when $n = 4l + 3$, the Universal Clifford algebra itself is an $\mathfrak{A}_n$-module under left multiplication. Thus Clifford analyticity, the study of $\mathfrak{A}_n$-valued solutions of $\mathfrak{D}F = 0$, is one special case of the theory of $\mathfrak{H}$-valued solutions. On the other hand, every real $M(k, \mathbb{K})$-module is isomorphic to the real tensor product $\mathbb{K}^k \otimes \mathbb{R}^m = \mathbb{K}^{k \times m}$ for some $m \geqslant 1$. Since the latter becomes a real Hilbert space under the inner product

$$(w, z) = \mathrm{tr}\,(\bar{z}w) \qquad (w, z \in \mathbb{K}^{k\times m}) \tag{4.14}$$

where $\bar{z} = [\bar{z}_{sr}] \in \mathbb{K}^{m\times k}$ is the conjugate of $z = [z_{rs}] \in \mathbb{K}^{k\times m}$, we obtain the following fundamental characterization familiar from real K-theory.

Theorem 4.15. *When $n = 8d + r$, $0 \leqslant r < 8$, and $(k, \mathbb{K})$ are specified by* (4.11), *then every $\mathbb{K}^{k\times m}$, $m \geqslant 1$, is an $\mathfrak{A}_n$-module under matrix multiplication of the left by $M(k, \mathbb{K})$. Conversely, every real Hilbert space that is also an $\mathfrak{A}_n$-module is isomorphic to some $M(k, \mathbb{K})$-module $\mathbb{K}^{k\times m}$ or to the direct sum of two such modules; this last case can occur only if $r = 3$ or 7.*

Thus real $\mathfrak{A}_n$-modules have a rigid structure with a very remarkable «period 8» property. The complex $\mathfrak{A}_n$-modules have a «period 2» property, and so have an even more rigid structure. To be more precise, let $\mathfrak{C}(\mathfrak{H})$ be the complex C^*-algebra generated in $\mathfrak{L}(\mathfrak{H})$ by $e_1, \ldots, e_n$ when $\mathfrak{H}$ is a complex $\mathfrak{A}_n$-module. Then $\mathfrak{C}(\mathfrak{H})$ is isometrically-*isomorphic to $M(2^m, \mathbb{C})$ when $n = 2m$ or to one of $M(2^m, \mathbb{C})$, $M(2^m, \mathbb{C}) \oplus M(2^m, \mathbb{C})$ when $n = 2m + 1$. Thus for each n there is a complex Hilbert space $\mathfrak{S}_n \cong \mathbb{C}^k$, $k = 2^{[n/2]}$, such that $\mathfrak{C}(\mathfrak{H})$ is isometrically-*isomorphic to the complex C^*-algebra $\mathfrak{L}(\mathfrak{S}_n)$ when $n = 2l$ or to one of $\mathfrak{L}(\mathfrak{S}_n)$, $\mathfrak{L}(\mathfrak{S}_n) \oplus \mathfrak{L}(\mathfrak{S}_n)$ when $n = 2l + 1$. This space $\mathfrak{S}_n$ is called the space of *complex spinors*. Every complex $\mathfrak{A}_n$-module, therefore, is isomorphic to an $M(k, \mathbb{C})$-module $\mathbb{C}^{k\times m}$ or to the direct sum of two such modules; this last case can occur only if n is odd.

To complete this section we return to the C^*-algebra representations of $\mathfrak{A}_n$ on an $\mathfrak{A}_n$-module $\mathfrak{H}$. There are always two such representations:

(4.16) (i) *the representation $\pi: \mathfrak{A}_n \to \mathfrak{A}(\mathfrak{H})$ derived from the universal property of $\mathfrak{A}_n$, and*

(4.16) (ii) *the representation $\sigma: \mathfrak{A}_n \to \mathfrak{A}(\mathfrak{H})$ defined by $\sigma(a) = \pi(a')$ where $a \to a'$ is the principal automorphism on $\mathfrak{A}_n$.*

Notice that

$$\pi(\bar{a}) = \pi(a)^-, \qquad \sigma(a^*) = \sigma(a)^-.$$

These will become crucial in the next section in realizing all the irreducible representations of the Euclidean rotation group.

5. Hardy Spaces $H^p(\Omega, \mathfrak{V})$

A large and very important class of Hardy spaces arise as subspaces of $H^p(\Omega, \mathfrak{H})$ by considering only those functions having range in a fixed subspace

of an $\mathfrak{A}_n$-module $\mathfrak{H}$. Let $\mathfrak{V}$ be such a subspace and regard $\mathfrak{D}$ as a first-order system of differential operators

$$\mathfrak{D}: C^\infty(\Omega, \mathfrak{V}) \to C^\infty(\Omega, \mathfrak{H}). \tag{5.1}$$

Such an operator is said to be of *Dirac type*. It inherits many of the properties that $\mathfrak{D}$ has as an operator on $C^\infty(\Omega, \mathfrak{H})$. It obviously has the GCR-property, for instance; and the Cauchy integral representation (2.16) is still valid for solutions in $C^\infty(\Omega, \mathfrak{V})$ of $\mathfrak{D}F = 0$. On the other hand, the Cauchy integral of a function f in $C(\partial\Omega, \mathfrak{V})$ need not be $\mathfrak{V}$-valued unless each e_j maps $\mathfrak{V}$ into $\mathfrak{V}$, *i.e.*, $\mathfrak{V}$ is an $\mathfrak{A}(\mathfrak{H})$-submodule of $\mathfrak{H}$; so in general the Cauchy integral cannot be used directly to construct $\mathfrak{V}$-valued solutions of $\mathfrak{D}F = 0$. But most importantly, the first-order system (5.1) is usually an *over-determined* elliptic system in the sense that its symbol mapping is injective without being surjective; the term *injectively elliptic* used in Section 2 would seem to be an appropriate one to emphasize the distinction with standard usage of the term elliptic. The operator $\mathfrak{D}$ will still be elliptic on $C^\infty(\Omega, \mathfrak{V})$ in this standard sense, however, when $\mathfrak{V}$ is an $\mathfrak{A}(\mathfrak{H})$-submodule of $\mathfrak{H}$.

Definition 5.2. *If $\mathfrak{V}$ is a subspace of $\mathfrak{H}$, the Hardy space $H^p(\Omega, \mathfrak{V})$, $0 < p < \infty$, is the set of all solutions in $C^\infty(\Omega, \mathfrak{V})$ of $\mathfrak{D}F = 0$ for which*

$$\|F\|_p = \left(\int_{\partial\Omega} N(F)(x)^p \, dS\right)^{1/p}$$

is finite.

The most important subspaces $\mathfrak{V}$ arise from representations of the Euclidean rotation group. Let

$$\Lambda_n = \{a \cdots wz : a, \ldots, w, z \in \mathbb{R}^{n+1}\} \tag{5.3}$$

be the multiplicative semi-group in $\mathfrak{A}_n$ of all finite products of elements $a, \ldots, w, z$ in $\mathbb{R}^{n+1}$; this is the subset of $\mathfrak{A}_n$ on which the C^*-algebra norm $\|\cdot\|$ is multiplicative. Thus the unit sphere

$$\{k \in \Lambda_n : \|k\| = 1\} \tag{5.4}$$

in Λ_n is a multiplicative group, the so-called Spin group Spin $(n+1)$. On the other hand, to each a in $\mathbb{R}^{n+1}$, $a \neq 0$, corresponds a «twisted» similarity transformation

$$\sigma(a): v \to av(a')^{-1} \qquad (v \in \mathbb{R}^{n+1}) \tag{5.5}$$

of $\mathbb{R}^{n+1}$ which can be shown to be the product of two reflections. Consequently, $\sigma(a)$ is a proper rotation of $\mathbb{R}^{n+1}$, *i.e.*, $\sigma(a) \in SO(n+1)$ for each non-

zero a in $\mathbb{R}^{n+1}$, and σ extends to a covering homomorphism

$$\sigma\colon \mathrm{Spin}(n+1) \to SO(n+1), \qquad \sigma(k)\colon v \to kv(k')^{-1} \tag{5.6}$$

from $\mathrm{Spin}(n+1)$ onto $SO(n+1)$. The restriction to k in the unit sphere of Λ_n ensures that σ is a two-fold covering. In the special cases of $\mathfrak{A}_1 = \mathbb{C}$ and $\mathfrak{A}_2 = \mathbb{H}$, $\mathrm{Spin}(n+1)$ reduces respectively to the unit circle $\{e^{i\theta/2}\colon 0 \leqslant \theta < 4\pi\}$ in $\mathbb{C}$ and the group $SU(2)$ regarded as the unit sphere in $\mathbb{H}$. It is well-known that these groups are two-fold coverings of $SO(2)$ and $SO(3)$. Hence, once more, we see how the Clifford theory contains the higher-dimensional analogues of classical low-dimensional results; this time it is the algebraic and geometric structure of the Euclidean rotation group, as expressed through (5.3), ..., (5.6), that is contained in Clifford theory. Even more basic to representation theory is the fact that every finite-dimensional $\mathrm{Spin}(n+1)$-module, and hence also any $SO(n+1)$-module, can be realized canonically as a subspace of an appropriate $\mathfrak{A}_n$-module $\mathfrak{H}$. These modules are then natural candidates for the $\mathfrak{V}$ in (5.2).

To be more precise, let $\mathfrak{R}_n$ be the Spinor module described in the previous section. By restricting to $\mathrm{Spin}(n+1)$ the two spin representations of $\mathfrak{A}_n$ on $\mathfrak{R}_n$ we obtain representations S_+ and S_-,

$$S_+(k)\colon x \to kx, \qquad S_-(k)\colon x \to k'x \tag{5.7}$$

of $\mathrm{Spin}(n+1)$ on $\mathfrak{R}_n$ (cf. (4.16)). Furthermore, these are irreducible representations of $\mathrm{Spin}(n+1)$ because the spin representation of $\mathfrak{A}_n$ on $\mathfrak{R}_n$ is irreducible and the linear span of $\mathrm{Spin}(n+1)$ is all of $\mathfrak{A}_n$. It will be convenient to denote these representations of $\mathrm{Spin}(n+1)$ by $(\mathfrak{R}_n^+, S_+)$ and $(\mathfrak{R}_n^-, S_-)$ respectively where, of course, $\mathfrak{R}_n^\pm = \mathfrak{R}_n$ and S_+, S_- are defined by (5.7). When $(n+1)$ is even, S_+ and S_- are inequivalent representations, but they are equivalent when $(n+1)$ is odd. Taking any finite tensor product (over $\mathbb{R}$)

$$\mathfrak{H} = \mathfrak{R}_n^{\epsilon_1} \otimes \cdots \otimes \mathfrak{R}_n^{\epsilon_r} \qquad (\epsilon_j = \pm 1) \tag{5.8}$$

we thus obtain an $\mathfrak{A}_n$-module on which $S_{\epsilon_1} \otimes \cdots \otimes S_{\epsilon_r}$ defines a reducible representation of $\mathrm{Spin}(n+1)$. From (5.6) and the Cartan-Weyl theory now follow the following fundamental result (cf. [14, Chap. III]).

Theorem 5.9. *Each irreducible representation of* $\mathrm{Spin}(n+1)$ *can be realized by restricting* $S_{\epsilon_1} \otimes \cdots \otimes S_{\epsilon_r}$ *to a canonically specified subspace of* $\mathfrak{R}_n^{\epsilon_1} \otimes \cdots \otimes \mathfrak{R}_n^{\epsilon_r}$ *for an appropriate choice of* $(\epsilon_1, \ldots, \epsilon_r)$.

If $\tau = (m_1, m_2, \ldots)$ is the signature of an irreducible representation of $\mathrm{Spin}(n+1)$, we shall denote by $\mathfrak{H}_r$ the $\mathfrak{A}_n$-module $\mathfrak{R}_n^{\epsilon_1} \otimes \cdots \otimes \mathfrak{R}_n^{\epsilon_r}$ and by $\mathfrak{V}_\tau$ the subspace of $\mathfrak{H}_\tau$ on which $S_{\epsilon_1} \otimes \cdots \otimes S_{\epsilon_r}$ realizes this representation. The

detailed prescription of $(\epsilon_1, \ldots, \epsilon_r)$ and the subspace $\mathfrak{V}_\tau$ of $\mathfrak{H}_\tau$ are given in [14, Chap. III]. To realize any finite-dimensional Spin $(n+1)$-module $\mathfrak{X}$ in an $\mathfrak{A}_n$-module, let $\mathfrak{X} = \mathfrak{X}_{\tau_1} \oplus \mathfrak{X}_{\tau_2} \oplus \cdots$ be the decomposition of $\mathfrak{X}$ into its irreducible Spin $(n+1)$-components labelled by their signature. Then $\mathfrak{X}$ is equivalent as a Spin $(n+1)$-module to the subspace $\mathfrak{V}_{\tau_1} \oplus \mathfrak{V}_{\tau_2} \oplus \cdots$ of $\mathfrak{H}_{\tau_1} \oplus \mathfrak{H}_{\tau_2} \oplus \cdots$. The fundamental representations of Spin $(n+1)$, for example, all arise from the representations $(S_+, \mathfrak{K}_+)$, $(S_-, \mathfrak{K}_-)$ and from restrictions of $S_{\epsilon_1} \otimes S_{\epsilon_2}$ to subspaces of $\mathfrak{K}_{\epsilon_1} \otimes \mathfrak{K}_{\epsilon_2}$, $\epsilon_j = \pm 1$. Again details are given more fully in [14, Chap. III].

6. $H^p(\Omega, \mathfrak{V})$-theory

In this section we shall describe how the fundamental properties of the classical Hardy spaces carry over to the Hardy spaces $H^p(\Omega, \mathfrak{V})$, $\Omega \subseteq \mathbb{R}^{n+1}$, associated with a Spin $(n+1)$-submodule $\mathfrak{V}$ of $\mathfrak{H}$. Some aspects of the theory have been examined in detail already, others less so, and some not at all. The complete picture has yet to be painted, though the broad outlines can now be sketched in. Much as in Section 1, the fundamental properties are divided into four broad categories: boundary values, first-order differential operators, boundary characterizations, and representation theory.

(*a*) Boundary theory. The L^p-boundary regularity was established for $p > p(\Omega)$ by using the solution of the scalar-valued Dirichlet problem for Ω. Since this corresponds in the classical case to treating an analytic function simply as a harmonic function, it means that the condition $\mathfrak{D}F = 0$ did not improve upon properties of harmonic functions; instead, it imposed algebraic restrictions on the boundary functions through the characterization of $H^p(\partial\Omega, \mathfrak{H})$ as the subspace of functions in $L^p(\partial\Omega, \mathfrak{H})$ left fixed by the principal-value Cauchy integral operator $\mathfrak{K}$. Clearly this is a major omission and would be a major weakness of the theory if extra regularity did not occur. The stumbling block to progress is knowing how to relate existence of boundary values on $\partial\Omega$ with the algebraic structure of $\mathfrak{D}$ and the geometry of $\partial\Omega$. The only technique currently known for overcoming this problem establishes first the subharmonicity properties that solutions of $\mathfrak{D}F = 0$ have. For instance, a general result of Calderón ensures that to each $\mathfrak{A}_n$-module $\mathfrak{H}$ there corresponds a value of p_0, $0 < p_0 < 1$, so that $\Delta|F|^p \geqslant 0$ holds on Ω for every $p > p_0$ and every solution of $\mathfrak{D}F = 0$ in $C^\infty(\Omega, \mathfrak{H})$ ([33, p. 233]). In fact one can be more precise (cf. [14, Chap. IV]).

Theorem 6.1. *The subharmonicity condition $\Delta|F|^p \geqslant 0$ holds on Ω for every $p > (n-1)/n$ and every solution of $\mathfrak{D}F = 0$ in $C^\infty(\Omega, \mathfrak{H})$ whatever the choice of $\mathfrak{A}_n$-module and open set Ω.*

In the case $n = 1$ this result reduces to the fact that $\Delta|F|^p \geqslant 0$ holds for every $p > 0$ when F is analytic. Now, if F is merely a harmonic function in $C^\infty(\Omega, \mathfrak{H})$, then the best one could say is that $\Delta|F|^p \geqslant 0$ holds with $p \geqslant 1$. Thus, just as in the classical case, the extra condition $\mathfrak{D}F = 0$ on a harmonic function decreases the lower bound on the allowed range of p. But it is natural to ask if this lower bound can be decreased still further by considering only those solutions of $\mathfrak{D}F = 0$ taking values in some fixed subspace $\mathfrak{V}$ of $\mathfrak{H}$. This idea goes back to an important paper of Stein-Weiss ([32]), and is one of the principal reason for introducing the Hardy spaces $H^p(\Omega, \mathfrak{V})$.

Definition 6.2. *Let*

$$\mathfrak{D}: C^\infty(\Omega, \mathfrak{V}_\tau) \to C^\infty(\Omega, \mathfrak{H}_\tau)$$

be the operator of Dirac type associated with the irreducible representation of Spin $(n + 1)$ *having signature* τ. *Then its critical index* p_τ *is the smallest value of* p_0 *so that* $\Delta|F|^p \geqslant 0$ *holds in* Ω *for every* $p > p_0$ *and every solution of* $\mathfrak{D}F = 0$ *in* $C^\infty(\Omega, \mathfrak{V}_\tau)$.

Obviously $0 < p_\tau \leqslant (n - 1)/n$, but the value of p_τ depends only on the way $\mathfrak{V}_\tau$ «lies» in $\mathfrak{H}_\tau$; in particular, it is independent of Ω ([14, Chap. IV]). By identifying particular examples of these operators of Dirac type with the operators ∂_τ in [9] (cf. also part (*b*) of this section), and hence with the operators introduced by Stein-Weiss, estimates for some p_τ can be deduced from the calculations in [32]. It would be very useful to have estimates for other p_τ (cf. [14, Chap. IV]).

The point of the condition $\Delta|F|^p \geqslant 0$ is that it can be used to establish L^p-boundary regularity, at least when Ω is an upper half-space, in complete analogy to the first half of (1.2) (i) (cf. [33, Chaps. II, IV]).

Theorem 6.3. *Let*

$$\Omega_+ = \{x + jy: x \in \mathbb{R}^n, y > 0\}$$

be an upper half-space in $\mathbb{R}^{n+1}_+$. *Then each* F *in* $H^p(\Omega_+, \mathfrak{V}_\tau)$, $p > p_\tau$, *has non-tangential boundary values* $F^+(x)$ *on* $\partial\Omega$ *for almost all* x; *furthermore*

$$\lim_{y \to 0^+} \int_{\mathbb{R}^n} |F(x + jy) - F^+(x)|^p \, dx.$$

With simple adaptations, the Stein-Weiss proof can be carried over to $H^p(\Omega, \mathfrak{V}_\tau)$ when Ω is the unit ball in $\mathbb{R}^{n+1}$, and with greater subtlety it may be possible to carry it over to C^1-domains. The possibility of extending it to any Lipschitz domain is a particularly interesting one because the proof depends

on being able to solve the Dirichlet problem in order to construct least harmonic majorants of subharmonic functions. Formally, we shall denote by

$$H^p(\partial\Omega, \mathfrak{B}_\tau) = \{F|_{\partial\Omega} : F \in H^p(\Omega, \mathfrak{B}_\tau)\} \qquad (p > p_\tau)$$

the space of boundary values, restricting Ω if necessary to ensure that they exist. For particular Ω and $\mathfrak{B}_\tau$, known atomic and molecular characterizations from real H^p-theory should provide corresponding characterizations of $H^p(\partial\Omega, \mathfrak{B}_\tau)$ (cf. [8], [11]) but little is known at present.

***(b)* First-order differential operators.** Much is known about the algebraic structure of $(\mathfrak{D}, \bar{\mathfrak{D}})$-operators and Dirac type operators, as well as about their position within the class of all first-order differential operators. For instance, the $(\mathfrak{D}, \bar{\mathfrak{D}})$-operators are as easily derived from a standard Dirac operator as the Cauchy-Riemann operators were. In the formalism of Section 2, let $\mathfrak{H}$ be a finite-dimensional Hilbert space on which $e_1, \ldots, e_n$ are skew-adjoint operators satisfying (2.2). Then

$$\gamma_0 = \begin{bmatrix} 0 & -e_0 \\ e_0 & 0 \end{bmatrix}, \qquad \gamma_j = \begin{bmatrix} 0 & e_j \\ e_j & 0 \end{bmatrix} \quad (1 \leqslant j \leqslant n) \tag{6.4}$$

are skew-adjoint operators on the Hilbert space $\mathfrak{H} \oplus \mathfrak{H}$ of (4.6), and

$$\gamma_j\gamma_k + \gamma_k\gamma_j = -2\delta_{jk}I \qquad (0 \leqslant j, k \leqslant n). \tag{6.4$'$}$$

Consequently, $\mathfrak{H} \oplus \mathfrak{H}$ is an $\mathfrak{A}_{n+1}$-module. The first-order system defined by

$$DF = \sum_{j=0}^{n} \gamma_j \frac{\partial F}{\partial x_j} \qquad (F \in C^\infty(\Omega, \mathfrak{H} \oplus \mathfrak{H})) \tag{6.5}$$

for any open set Ω in $\mathbb{R}^{n+1}$ is then a standard Dirac operator as the term was used in [15]: because of (6.4), D is an elliptic self-adjoint differential operator such that

$$D = \begin{bmatrix} 0 & -\bar{\mathfrak{D}} \\ \mathfrak{D} & 0 \end{bmatrix}, \qquad D^2 = -\Delta. \tag{6.6}$$

Thus

$$DF = \begin{bmatrix} 0 & -\bar{\mathfrak{D}} \\ \mathfrak{D} & 0 \end{bmatrix} \begin{bmatrix} \Phi \\ \Psi \end{bmatrix} = \begin{bmatrix} -\bar{\mathfrak{D}}\Psi \\ \mathfrak{D}\Phi \end{bmatrix}$$

on any $\mathfrak{H} \oplus \mathfrak{H}$-valued function

$$F = \begin{bmatrix} \Phi \\ \Psi \end{bmatrix},$$

so $(\mathfrak{D}, \bar{\mathfrak{D}})$ arise from restricting D to $\mathfrak{H}$-valued functions.

An alternative way of describing the Dirac type operators

$$\mathfrak{D}: C^\infty(\Omega, \mathfrak{V}) \to C^\infty(\Omega, \mathfrak{H}), \qquad \mathfrak{D}F = \sum_{j=0}^{n} e_j \frac{\partial F}{\partial x_j}, \tag{6.7}$$

helps bring out their algebraic structure more clearly. Since this second approach essentially defines a differential operator through its symbol, it is also a convenient one to use on more general manifolds. Let $\mathfrak{A}$, $\mathfrak{B}$ be finite dimensional real Hilbert spaces, and let $\mathfrak{A} \otimes \mathbb{R}^{n+1}$ be the usual Hilbert space tensor product. If $\{e_j\}_{j=0}^n$ is the standard basis for $\mathbb{R}^{n+1}$ this tensor product consists of the elements

$$\xi = \sum_{j=0}^{n} \lambda_j (a_j \otimes e_j) \qquad (\lambda_j \in \mathbb{R}, a_j \in \mathfrak{A}) \tag{6.8}$$

having Hilbert-Schmidt norm

$$\|\xi\| = \left(\sum_{j=0}^{n} \lambda_j^2 \|a_j\|_{\mathfrak{A}}^2 \right)^{1/2}. \tag{6.8$'$}$$

In particular, to each linear operator $A: \mathfrak{A} \otimes \mathbb{R}^{n+1} \to \mathfrak{B}$ there corresponds a family $A_0, \ldots, A_n$ of linear operators

$$A_j: \mathfrak{A} \to \mathfrak{B}, \qquad A_j(a) = A(a \otimes e_j) \qquad (0 \leqslant j \leqslant n).$$

Definition 6.9. *Let*

$$\partial_A F = (A \circ \nabla)F = \sum_{j=0}^{n} A_j \frac{\partial F}{\partial x_j}$$

be the differential operator obtained by composing the usual derivative

$$\nabla: C^\infty(\Omega, \mathfrak{A}) \to C^\infty(\Omega, \mathfrak{A} \otimes \mathbb{R}^{n+1}), \qquad \nabla F = \sum_{j=0}^{n} \frac{\partial F}{\partial x_j} \otimes e_j$$

of $\mathfrak{A}$-valued functions with a linear operator $A: \mathfrak{A} \otimes \mathbb{R}^{n+1} \to \mathfrak{B}$.

This defines for each A a linear first-order homogeneous constant-coefficient differential operator such that $\partial_A: C^\infty(\Omega, \mathfrak{A}) \to C^\infty(\Omega, \mathfrak{B})$; and every such operator can be put in this form for some choice of A. In (6.7), for instance,

$$A: \mathfrak{V} \otimes \mathbb{R}^{n+1} \to \mathfrak{H}, \qquad A: \sum_{j=0}^{n} v_j \otimes e_j \to \sum_{j=0}^{n} e_j v_j,$$

where $A_j: \mathfrak{V} \to \mathfrak{H}$, $1 \leqslant j \leqslant n$, is multiplication by the imaginary unit e_j, while $A_0: \mathfrak{V} \to \mathfrak{V}$ is the identity. In general, of course, the mapping

$$a \to A(a \otimes \lambda) = \sum_{j=0}^{n} \lambda_j (A_j a) \qquad (a \in \mathfrak{A}) \tag{6.10}$$

defined from $\mathfrak{A}$ to $\mathfrak{B}$ for each non-zero $\lambda = (\lambda_0, \ldots, \lambda_n)$ is just the symbol mapping of ∂_A. To understand the algebraic structure of ∂_A, therefore, we need to know how the properties of $A: \mathfrak{A} \otimes \mathbb{R}^{n+1} \to \mathfrak{B}$ determine those of $\partial_A: C^\infty(\Omega, \mathfrak{A}) \to C^\infty(\Omega, \mathfrak{B})$.

Let $\partial_B: C^\infty(\Omega, \mathfrak{A}) \to C^\infty(\Omega, \mathfrak{B}')$ be the differential operator determined by a linear operator $B: \mathfrak{A} \otimes \mathbb{R}^{n+1} \to \mathfrak{B}'$ having a possibly different range space $\mathfrak{B}'$ from that of $A: \mathfrak{A} \otimes \mathbb{R}^{n+1} \to \mathfrak{B}$. We shall say that ∂_A is equivalent to ∂_B if they have the same kernel, *i.e.*, $\partial_A F = 0$ if and only if $\partial_B F = 0$. Since the vanishing of $\partial_A F(x_0)$, $x_0 \in \Omega$, is a linear condition on the homogeneous first-order Taylor coefficients of F at x_0, it is easy to prove

Theorem 6.11. *The first-order differential operators ∂_A and ∂_B on $C^\infty(\Omega, \mathfrak{A})$ are equivalent for every Ω in $\mathbb{R}^{n+1}$ if and only if A, B have the same kernel in $\mathfrak{A} \otimes \mathbb{R}^{n+1}$, i.e., $A\xi = 0$ if and only if $B\xi = 0$, $\xi \in \mathfrak{A} \otimes \mathbb{R}^{n+1}$.*

In other words, there is a $1-1$ correspondence between subspaces of $\mathfrak{A} \otimes \mathbb{R}^{n+1}$ and the classes of first-order differential operators determined by this notion of equivalence. In particular, an operator ∂_A will be equivalent to the Dirac type operator (6.7) precisely when the kernel of A in $\mathfrak{V} \otimes \mathbb{R}^{n+1}$ satisfies

$$\ker A = \left\{ v \in \mathfrak{V} \otimes \mathbb{R}^{n+1} : v = \sum_{j=0}^{n} v_j \otimes e_j, \sum_{j=0}^{n} e_j v_j = 0 \right\}. \tag{6.12}$$

For the Spin $(n+1)$-modules $\mathfrak{V}_\tau$, group-theoretic methods were used in [10] to construct rotation-invariant operator ∂_τ whose symbol has property (6.12). Here again the idea goes back to the fundamental paper of Stein-Weiss ([32]). The crucial idea in the rotation-invariant case is to decompose the Spin $(n+1)$-module $\mathfrak{V}_\tau \otimes \mathbb{R}^{n+1}$ into its irreducible Spin $(n+1)$-submodules. The classical geometric differential operators are then the operators of Dirac type arising from the fundamental representations of Spin $(n+1)$, and the higher gradient operators introduced by Stein-Weiss are the Dirac type operators corresponding to representations on spherical harmonics. More details can be found in [10], [14].

A careful examination of the precise values of $(\mathbb{K}, k)$ in (4.11) explains the significance of the estimates of $\dim_{\mathbb{F}}(\mathfrak{H})$ in Theorem 2.13. For the results of Adams *et al.* ([1], [30]) show that if $\partial_A: C^\infty(\Omega, \mathfrak{A}) \to C^\infty(\Omega, \mathfrak{B})$ is any linear, elliptic, first-order system on an open set Ω in $\mathbb{R}^{n+1}$, $n = 8d + r$, $0 \leqslant r < 7$, then $\dim_{\mathbb{F}}(\mathfrak{A})$ satisfies the same estimates

$$\dim_{\mathbb{R}}(\mathfrak{A}) \geqslant (r+1)16^d, \qquad \dim_{\mathbb{C}}(\mathfrak{A}) \geqslant 2^{[r/2]}16^d,$$

as $\dim_{\mathbb{F}}(\mathfrak{H})$ did. Now the lower bounds on the dimension are attained for $\mathfrak{D}, \bar{\mathfrak{D}}$ on $\mathfrak{R}_n$- or S_n-valued functions. Hence in a very precise sense the Dirac $\mathfrak{D}$- and $\bar{\mathfrak{D}}$-operators are the «smallest» linear elliptic first-order systems of differential operators.

(c) Boundary value problems. Using the known solutions of the various Riemann-Hilbert problems for analytic functions and the Dirichlet-Neumann problems in higher dimensions as a guide, we can formulate a series of boundary value problems for the Hardy spaces $H^p(\Omega, \mathfrak{H})$ and $H^p(\Omega, \mathfrak{B}_\tau)$. But apart from the special case when Ω is a half-space (or ball) in $\mathbb{R}^{n+1}$ where Fourier Transform or group-theoretic techniques can be exploited, few results have been established as yet, so the conjectural nature of what follows must be borne in mind. There are really two distinct difficulties:

(i) setting up the *algebraic* formalism of the boundary value problems for an arbitrary set Ω, especially *group-theoretic* aspects, and
(ii) handling the extra *analytic* difficulties created by the geometry of $\partial\Omega$, more particularly by its lack of smoothness.

To avoid problems of the second kind at this stage, we shall tacitly assume that $\partial\Omega$ is C^∞ when formulating a boundary value problem; for instance, the example of $\partial\Omega = \Sigma_n$, the unit sphere in $\mathbb{R}^{n+1}$, is a natural test case. Very intuitively, our basic philosophy will be that the boundary theory should reflect the *over-determinedness* of $(\mathfrak{D}, \bar{\mathfrak{D}})$. Given, say a Hardy space $H^p(\Omega, \mathfrak{H})$ or $H^p(\Omega, \mathfrak{B}_\tau)$, the problem is to exhibit a «smaller» Banach space $\mathfrak{X}$ of functions or distributions on $\partial\Omega$ isomorphic to this Hardy space, isomorphic meaning that there is a potential type operator $\mathfrak{K}$ mapping the boundary space onto the Hardy space or a boundary type operator mapping the Hardy space onto the boundary space. A «smaller» Banach space would mean one where the range space of the functions or distributions in $\mathfrak{X}$ is a proper subspace of $\mathfrak{H}$ or $\mathfrak{B}_\tau$, or where they have to satisfy a boundary condition such as being fixed by the principal-value Cauchy integral operator or some such singular integral operator on $\partial\Omega$. These restrictions were present in the classical Riemann-Hilbert problem for analytic functions and in Theorem 3.9. To proceed in detail, however, a precise understanding of the meaning of over-determinedness is needed. Now $(\mathfrak{D}, \bar{\mathfrak{D}})$ and $(\bar{\partial}, \partial)$ are elliptic operators, and so are *determined* in the sense that the associated symbols are both surjective and injective. Nonetheless, the solutions of $\bar{\partial}F = 0$ are *analytically over-determined.* The splitting $\mathfrak{H} = \mathfrak{H}_0 \oplus j\mathfrak{H}_0$ was used in (3.11) to express the same analytic over-determinedness that $\mathfrak{D}$ is presumed to have as a differential operator on $C^\infty(\Omega, \mathfrak{H})$. Recall Re, Im: $\mathfrak{H} \to \mathfrak{H}_0$ were defined by

$$z = \mathrm{Re}(z) + j\,\mathrm{Im}(z) = x + jy \qquad (x, y \in \mathfrak{H}_0) \tag{6.13}$$

In an attempt to understand thoroughly the analytic over-determinedness of $\mathfrak{D}$, therefore, we introduce the following analogue of the Hardy-Riemann Hilbert problem for analytic functions (cf. (1.16)). Let

$$\Lambda(\mathfrak{H}) = \{a \cdots wz : a, \ldots, w, z \in \mathbb{R}^{n+1} \subseteq \mathfrak{H}\} \tag{6.14}$$

be the multiplicative semi-group in $\mathfrak{A}(\mathfrak{H})$ of all finite products of elements in $\mathbb{R}^{n+1}$ ($\subseteq \mathfrak{H}$). On this semi-group the C^*-algebra norm is multiplicative; in particular, if $A: \partial\Omega \to \Lambda(\mathfrak{H})$ is non-vanishing on $\partial\Omega$, then it is invertible.

(6.15) Hardy-Riemann-Hilbert Problem. *Let $A: \partial\Omega \to \Lambda(\mathfrak{H})$ be a sufficiently smooth function which is non-vanishing on $\partial\Omega$. Then, given f in $L^p(\partial\Omega, \mathfrak{H}_0)$, find F in $H^p(\partial\Omega, \mathfrak{H})$ such that* $\operatorname{Re}(\overline{A(z)}F(z)) = f(z)$ *on* $\partial\Omega$.

In view of the boundary regularity theory for H^p-spaces, this problem reduces to solving the equation $\mathfrak{I}_A(F) = f$ where

$$\mathfrak{I}_A : H^p(\partial\Omega, \mathfrak{H}) \to L^p(\partial\Omega, \mathfrak{H}_0), \qquad (\mathfrak{I}_A F)(z) = \operatorname{Re}(\overline{A(z)}F(z)), \tag{6.16}$$

just as in the classical case. The Dirichlet type problem in conjecture (3.11) is the special case $A = 1$. In the classical case one uses the Cauchy integral in an attempt to invert $\mathfrak{I}_A$, and presumably a similar attempt might solve (6.15). The availability of the Cauchy integral and all the related theory to solve boundary value problems such as (6.15) indicate the usefulness of a development of «analytic function theory» for higher dimensional Euclidean space. A topological characterization of the index of $\mathfrak{I}_A$, special case of the Index theorem for manifolds with boundary, would be particularly interesting also.

There are some natural geometric choices of A. Let $\eta(X)$ be the unit outward normal to $\partial\Omega$ at X. Then $X \to \eta(X)$ defines a function $\eta: \partial\Omega \to \Sigma_n$ ($\subseteq \Lambda(\mathfrak{H})$), the Gauss map, which is a natural candidate for A. For any such $A: \partial\Omega \to \Sigma_n$ the Hardy-Riemann-Hilbert is related to the classical Neumann or Oblique-derivative problem.

By restricting to the Dirac type operator

$$\mathfrak{D}: C^\infty(\Omega, \mathfrak{V}_\tau) \to C^\infty(\Omega, \mathfrak{H}_\tau), \tag{6.17}$$

the differential operator becomes algebraically over-determined, *i.e.*, injectively elliptic, as well as analytically over-determined, and so algebraic restrictions now enter. Special cases for the unit ball were studied in [24], [25], and [28], but the only known results valid for general domains rise in studies of Dirichlet-Neumann Problem (cf. [22]), particularly in the work of Fabes-Kenig and Verchota ([11], [36]). This last case corresponds to the standard representation of $SO(n+1)$ on $\mathbb{R}^{n+1}$, so that $\mathfrak{V}_\tau = \mathbb{R}^{n+1}$ and the Dirac type operator (6.17) is equivalent to the div-curl system

$$\text{(6.18)} \qquad \operatorname{div} F = 0 \qquad \operatorname{curl} F = 0, \qquad (F \in C^\infty(\Omega, \mathbb{R}^{n+1}))$$

on Ω ([33, p. 234]). Now at each boundary point, the unit outward normal $\eta(X)$ determines a geometric splitting

$$\text{(6.19)} \qquad \mathbb{R}^{n+1} = T_X \oplus N_X, \qquad v = x + \eta(X)y \qquad (x \perp \eta(X), y \in \mathbb{R})$$

of each $v \in \mathfrak{V}_\tau$ into tangential and normal components. Hence each $\mathfrak{V}_\tau$-valued function $F = F(X)$ on $\partial\Omega$ admits a decomposition $F(X) = \Phi(X) + \eta(X)\Psi(X)$ with $\Phi(X) \perp \eta(X)$ and $\Psi = \Psi(X)$ a scalar-valued function on $\partial\Omega$; technically, $F = \Phi + \eta\Psi$ with Φ, $\eta\Psi$ sections of the tangent and normal bundles on $\partial\Omega$ respectively. Solving the Neumann problem for Ω with L^p-data is equivalent to finding F in $H^p(\Omega, \mathfrak{V}_\tau)$ so that $F(X) = \Phi(X) + \eta(X)\Psi(X)$ where $\Psi = \Psi(X)$ is a specified real-valued function in $L^p(\partial\Omega)$. Save possibly for a vanishing moment condition, the choice of Ψ is arbitrary. But any Φ arising as the tangential component of $F|_{\partial\Omega}$ is far from arbitrary (cf. [22], [36]).

To proceed in general, let $\mathfrak{H}_\tau = \mathfrak{H}_0 \oplus j\mathfrak{H}_0$ be a *fixed* splitting of $\mathfrak{H}_\tau$ and set

$$\text{(6.20)} \qquad \mathfrak{T}_\tau = \{\operatorname{Re}(v): v \in \mathfrak{V}_\tau\}, \qquad \mathfrak{N}_\tau = \{\operatorname{Im}(v): v \in \mathfrak{V}_\tau\};$$

we shall think of these as the respective *tangential* and *normal* components of $\mathfrak{V}_\tau$ analogous to the splitting of $\mathbb{R}^{n+1}$ in (6.19). But instead of splitting at every boundary point, we can compensate by introducing a function $A: \partial\Omega \to \Lambda(\mathfrak{H})$; for as we saw earlier, the function

$$f(z) = \operatorname{Re}(\overline{A(z)}F(z)) \qquad (z \in \partial\Omega)$$

would give the «normal component» of F at $z \in \partial\Omega$ for any appropriate choice of A, and a different choice of A would also give the «tangential component». Thus, let $A: \partial\Omega \to \Lambda(\mathfrak{H})$ be a sufficiently smooth non-vanishing function on $\partial\Omega$.

(6.21) Problem. *Determine the pseudo-differential operator $\mathfrak{R}_\tau$ on $L^p(\partial\Omega, \mathfrak{N}_\tau)$ so that the equations*

$$\mathfrak{R}_\tau f = 0, \qquad \operatorname{Re}(\overline{A(z)}F(z)) = f(z) \qquad (z \in \partial\Omega)$$

can be solved for F in $H^p(\Omega, \mathfrak{V}_\tau)$ and f in $L^p(\partial\Omega, \mathfrak{N}_\tau)$.

(*d*) Representation theory. Just as the hyperbolic plane H_2 can be realized as the upper half-plane or unit disk in $\mathbb{C}$, hence in the Clifford algebra $\mathfrak{A}_1$, so $(n+1)$-dimensional real hyperbolic space H_{n+1} can be realized as an upper half-space or unit ball in $\mathbb{R}^{n+1}$, hence in the Clifford algebra $\mathfrak{A}_n$. What is more, the algebra structure that becomes available because of this embedding makes it possible to carry over virtually intact from $SL(2, \mathbb{R})$ to $\operatorname{Spin}_0(n+1, 1)$ the algebraic, geometric, and analytic results associated with the group of isometries

of H_{n+1}. The idea can be traced back to Vahlen and Maas, but it has been brought to the fore more recently by Ahlfors (cf. [3] for instance) and, independently, by Takahashi ([34]). Its significance as far as Hardy theory is concerned, arises for two reasons: when $\mathrm{Spin}_0(n+1,1)$ is the identity component of the two-fold covering of $SO_0(n+1,1)$, then

(i) $\mathrm{Spin}_0(n+1,1)$ *acts on* $\mathbb{R}^{n+1}$ *by fractional linear transformations each of which is a composition of translation, rotation, dilation and inversion*;

(ii) *there is an induced action of* $\mathrm{Spin}_0(n+1,1)$ *by unitary operators both on* $H^2(\Omega, \mathfrak{H})$, $\Omega = \mathbb{R}^{n+1}_+$, *and on the boundary space* $L^2(\mathbb{R}^n, \mathfrak{H})$, *that commutes with the Cauchy integral and the principal-value Cauchy integral operators.*

The representation-theoretic implications of these properties will be discussed elsewhere. Here we shall do little more than indicate how closely the results for $n > 1$ parallel those in the classical case $n = 1$ discussed in Section 1.

Fix an imaginary unit $\boldsymbol{j}$ in $\mathbb{R}^n$ ($\subseteq \mathfrak{A}_n$) and let $\mathfrak{A}_{n-1}$ be the Clifford subalgebra of $\mathfrak{A}_n$ generated by $e_1, \ldots, e_{n-1}$ where $\{e_1, \ldots, e_{n-1}, \boldsymbol{j}\}$ is an orthonormal basis for $\mathbb{R}^n$. Then

$$\mathfrak{A}_n = \mathfrak{A}_{n-1} \oplus \boldsymbol{j}\mathfrak{A}_{n-1} \tag{6.22}$$

and

$$\Omega_+ = \left\{ z : z = x + y\boldsymbol{j}, x = \sum_{k=0}^{n-1} x_k e_k \in \mathbb{R}^n, y > 0 \right\} \tag{6.23 (i)}$$

is an upper half-space in $\mathbb{R}^{n+1}$ having the same boundary

$$\partial\Omega = \left\{ x : x = \sum_{k=0}^{n-1} x_k e_k \in \mathbb{R}^n \right\} \tag{6.23 (ii)}$$

as the lower half-space

$$\Omega_- = \left\{ z : z = x + y\boldsymbol{j}, x = \sum_{k=0}^{n-1} x_k e_k \in \mathbb{R}^n, y < 0 \right\}. \tag{6.23 (iii)}$$

This imaginary unit $\boldsymbol{j}$ is the unit inward normal to Ω_+ and unit outward normal to Ω_-. Thus $\boldsymbol{j}$ plays exactly the same role for $\mathbb{R}^{n+1}$ as $i = \sqrt{-1}$ does classically for $\mathbb{C} = \mathbb{R} \oplus i\mathbb{R}$ when $n = 1$. In this last case, $\mathrm{Spin}_0(n+1,1)$ can be identified with $SL(2, \mathbb{R})$ acting on the compactified plane $\mathbb{C} \cup \{\infty\}$ by fractional linear transformations

$$g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} : z \to z \cdot g = \frac{b + zd}{a + cz} \qquad (z \in \mathbb{C})$$

as we saw in Section 1. In the general case, $\mathrm{Spin}_0(n+1,1)$ is realized as a multiplicative subgroup G of

$$M(2, \mathfrak{A}_{n-1}) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d \in \mathfrak{A}_{n-1} \right\}$$

acting on $\mathbb{R}^{n+1} \cup \{\infty\}$ by fractional linear transformations

$$g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} : z \to (a + zc)^{-1}(b + zd) \qquad (z \in \mathbb{R}^{n+1}), \tag{6.24}$$

so that the orbits of G in $\mathbb{R}^{n+1} \cup \{\infty\}$ consist of Ω_+, Ω_-, and their compactified mutual boundary $\partial\Omega \cup \{\infty\}$. Although there is an intrinsic definition of G (cf. [14, Chap. V]), it is more illuminating to proceed indirectly, constructing subgroups M, A and V of G with respect to which (6.24) reduces to rotations, dilations and translations respectively. Let Λ_{n-1} be the multiplicative semi-group generated in $\mathfrak{A}_{n-1}$ by $\partial\Omega$; when $n = 1$, Λ_{n-1} is just $\mathbb{R}$, of course. Set

$$M = \left\{ \begin{bmatrix} \alpha & 0 \\ 0 & \alpha' \end{bmatrix} : \alpha \in \Lambda_{n-1}, \|\alpha\| = 1 \right\}, \qquad A = \left\{ \begin{bmatrix} \lambda^{-1/2} & 0 \\ 0 & \lambda^{1/2} \end{bmatrix} : \lambda > 0 \right\},$$

and

$$V = \left\{ \begin{bmatrix} 1 & v \\ 0 & 1 \end{bmatrix} : v \in \partial\Omega_+ \right\}.$$

For these groups, (6.24) becomes

$$\begin{bmatrix} \alpha & 0 \\ 0 & \alpha' \end{bmatrix} : z \to \alpha^{-1} z \alpha' = \sigma(a^{-1})v \qquad (z \in \mathbb{R}^{n+1}),$$

which is known to be a *rotation* of $\mathbb{R}^{n+1}$ fixing the $\boldsymbol{j}$-direction (cf. (5.5)), while

$$\begin{bmatrix} \lambda^{-1/2} & 0 \\ 0 & \lambda^{1/2} \end{bmatrix} : z \to \lambda z \qquad (z \in \mathbb{R}^{n+1}),$$

defines *dilation* by λ, and

$$\begin{bmatrix} 1 & v \\ 0 & 1 \end{bmatrix} : z \to z + v \qquad (z \in \mathbb{R}^{n+1}),$$

defines *translation* by $v \in \partial\Omega$ in any hyperplane parallel to $\partial\Omega$. On the other hand, the fractional linear transformation determinaded by the Weyl group element

$$\omega = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

is the inversion

$$\omega = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} : z \to -1/z$$

which interchanges 0 and ∞. Finally, set

$$N = \omega V \omega^{-1} = \left\{ \begin{bmatrix} 1 & 0 \\ u & 1 \end{bmatrix} : u \in \partial\Omega \right\}.$$

It is easy to check that MAN is a multiplicative group in $M(2, \mathfrak{A}_{n-1})$. Assuming that $\mathrm{Spin}_0(n+1, 1)$ has been realized as an intrisically defined group G in $M(2, \mathfrak{A}_{n-1})$, one shows

Theorem 6.25. (Bruhat decomposition.) *There is a uniquely determined coset-space decomposition*

$$G = MANV \cup MAN\omega$$

of G with respect to its subgroup MAN such that the left coset-space $MAN\backslash G$ is diffeomorphic to $\partial\Omega \cup \{\infty\}$, identifying V with $\partial\Omega$ and ω with $\{\infty\}$; furthermore, when G acts on $\mathbb{R}^{n+1} \cup \{\infty\}$ by fractional linear transformations

$$g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} : z \to z \cdot g = (a + zc)^{-1}(b + zd),$$

the orbits of G are Ω_+, Ω_-, and their compactified mutual boundary $\partial\Omega \cup \{\infty\}$.

With this coset-space decomposition for G it is not hard to show that $z \to z \cdot g$ preserves the usual hyperbolic metric on Ω_+ as a realization of H_{n+1}; and, conversely, every isometry in the connected group of isometries of H_{n+1} arises as such a fractional linear transformation. On the other hand, on the compactified boundary $\partial\Omega \cup \{\infty\}$, the fractional linear transformations $x \to x \cdot g$ are precisely the (sense-preserving) conformal transformations of $\mathbb{R}^n \cup \{\infty\}$ (cf. [2]). The geometric and algebraic properties of G are thus completely encoded in the transformations $z \to z \cdot g$ of $\mathbb{R}^{n+1} \cup \{\infty\}$. Analytic properties show up in Hardy H^2-spaces for Ω_+.

Let $\mathfrak{H}$ be an $\mathfrak{A}_n$-module, and let $\mathfrak{A}(\mathfrak{H})$ be the associated C^*-algebra realization of $\mathfrak{A}_n$. Then

$$(6.26) \qquad g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} : z \to z \cdot g = (a + zc)^{-1}(b + zd) \qquad (z \in \mathbb{R}^{n+1})$$

is well-defined solely within $\mathfrak{A}(\mathfrak{H})$, regarding z and the matriz entries of g as

elements of $\mathfrak{A}(\mathfrak{H})$; in particular, any term such as $(a + zc)$ is an operator on $\mathfrak{H}$. Now let $H^2(\Omega_+, \mathfrak{H})$ be the Hardy space on Ω_+ whose norm is defined geometrically by the natural analogue

$$\|F\|_{H^2} = \sup_{y>0} \left(\int_{\partial\Omega} |F(x + jy)|^2 \, dx \right)^{1/2}$$

of (1.1).

Theorem 6.27. *For each g in G,*

$$(\pi_s(g)F)(z) = \frac{(a + zc)'}{|a + zc|^{n+1}} F(z \cdot g) \qquad (z \in \Omega_+)$$

is a unitary operator on $H^2(\Omega_+, \mathfrak{H})$, and $g \to \pi_s(g)$ is a unitary representation of G on $H^2(\Omega_+, \mathfrak{H})$.

As stated, Theorem 6.27 really hides the fundamental property of the Dirac $\mathfrak{D}$-operator that is needed. For $\pi(g)F$ must be a solution of $\mathfrak{D}(\pi(g)F) = 0$ if $\pi(g)F$ is to be in $H^2(\Omega_+, \mathfrak{H})$. This, however, follows immediately from the fact that

$$\mathfrak{D}(\pi_s(g)F(z \cdot g)) = \frac{1}{|a + zc|^2} \pi_s(g')(\mathfrak{D}F)(z \cdot g)$$

holds for any F in $C^\infty(\Omega_+, \mathfrak{H})$, an equivariance property which ensures that $\mathfrak{D}$ is invariant not only under translation and rotations, the transformations preserving the Euclidean distance, but also under transformations preserving hyperbolic distance on Ω_+, just as the Cauchy-Riemann operators were in the classical case.

To relate this action of G on $H^2(\Omega_+, \mathfrak{H})$ to one on spaces of boundary values, we prove

Theorem 6.28. *For each g in G,*

$$(\sigma_s(g)f)(x) = \frac{(a + xc)'}{|a + xc|^{n+1}} f(x \cdot g) \qquad (x \in \partial\Omega)$$

is a unitary operator on $L^2(\partial\Omega, \mathfrak{H})$ and $g \to \sigma_s(g)$ is a unitary representation of G such that

$$\text{(i)}\ \mathcal{K}(\sigma_s(g)f) = \sigma_s(g)(\mathcal{K}f) \qquad \text{(ii)}\ \mathcal{C}(\sigma_s(g)f) = \pi_s(g)(\mathcal{C}f)$$

where $\mathcal{K}$ is the principal-value Cauchy integral operator on $\partial\Omega$ and $\mathcal{C}$ is the Cauchy integral operator.

Details of these and other results will appear in [17].

References

[1] Adams, J. F., Lax, P. D. and Philips, R. S. On matrices whose real linear combinations are non-singular, *Proc. Amer. Math. Soc.* **16**(1965), 318-322.

[2] Ahlfors, L. V. Möbius transforms in several dimensions, University of Minnesota Lecture Notes, Minnesota, 1981.

[3] Ahlfors, L. V. Möbius transformations in $\mathbb{R}^n$ expressed through 2×2 matrices of Clifford Numbers, *Complex Variables* **5**(1986), 215-224.

[4] Atiyah, M. H., Bott, R. and Shapiro, A. Clifford Modules, *Topology* **3**(1964), 3-38.

[5] Atiyah, M. and Singer, I. M. The index of elliptic operators: III. *Ann. of Math.* **87**(1968), 546-604.

[6] Brackx, F., Delanghe, R. and Sommen, F. *Clifford Analysis*, Pitman, 1983.

[7] Coifman, R. R., McIntosh, A. and Meyer, Y. L'integrale de Cauchy définit un opérateur borné sur L^2 pour les courbes Lipschitziennes, *Annals of Math.* **116**(1982), 361-387.

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

[9] Davis, K. M., Gilbert, J. E. and Kunze, R. A. Elliptic differential operators in harmonic analysis, I. Generalized Cauchy-Riemann systems. Submitted to *Amer. J. Math.*

[10] Davis, K. M., Gilbert, J. E. and Kunze, R. A. Invariant differential operators in harmonic analysis on real hyperbolic space, in *Proc. Int. Conf. Harmonic Analysis and Operator Algebras*, Canberra, 1987. To appear.

[11] Fabes, E. and Kenig, C. E. On the Hardy H^1-space of a C^1-domain, *Ark. Mat.* **19**(1981), 1-22.

[12] Gakhov, F. D. *Boundary Value Problems*, Pergamon Press, 1966.

[13] Gilbert, J. E., Kunze, R. A., Stanton, R. J. and Tomas, P. A. Higher Gradients and Representations of Lie Groups, in *Conference on Harmonic Analysis in Honor of Antoni Zygmund*, Vol. II (W. Beckner, A. P. Calderón, R. Fefferman, P. Jones, eds.) Wadsworth International (1981), 416-426.

[14] Gilbert, J. E. and Murray, M. A. M. *Clifford Algebras and Dirac operators in harmonic analysis*, in Cambridge Studies in Advanced Mathematics. To appear.

[15] Gilbert, J. E. and Murray, M. A. M. The Dirac Operator and Hardy spaces of Clifford analytic functions on Lipschitz domains, submitted for publication.

[16] Gilbert, J. E., Murray, M. A. M. and Ryan, J. Generalized Cauchy-Riemann systems, the Cauchy integral formula, and operators of Dirac type, submitted for publication.

[17] Gilbert, J. E., Murray, M. A. M. and Ryan, J. Invariance properties of the Dirac operator and representation theory, in preparation.

[18] Goodearl, K. R. *Real and Complex C*-algebras*, Shiva Math. Series, No. 5, Birhauser, 1982.

[19] Gromov, M. and Lawson, B. Positive scalar curvature and the Dirac operator on complete Riemannian manifolds, *Publ. Math. IHES.* **58**(1983), 295-408.

[20] Iwanenko and Nikolski. Über den Zusammenhang zwischen den Cauchy-Riemann'schen und Dirac'schen Differentialgleichungen, *Zeit. Für Phys.*, **63**(1930), 129-137.

[21] Kenig, C. E. Weighted H^p spaces on Lipschitz domains, *Amer. J. Math.* **102**(1980), 129-163.

[22] Kenig, C. E. Elliptic boundary value problems on Lipschitz domains, in *Beijing Lectures in Harmonic Analysis*, Princeton Univ. Press, (1986), 131-183.
[23] Kenig, C. E. and Pipher, J. Hardy spaces and the Dirichlet problem on Lipschitz domains, *Revista Mat. Iberoamericana* **3**(1987), 191-247.
[24] Korányi, A. and Vági, S. Cauchy-Szegö integrals for systems of harmonic functions, *Ann. Scuola Norm. Sup. Pisa,* **26**(1972), 181-196.
[25] Korányi A. and Vági, S. Group-theoretic remarks on Riesz systems on balls, *Proc. Amer. Math. Soc.* **85**(1982), 200-205.
[26] Muskhelishvili, N. I. *Singular Integral Equations*, Wolters-Noordhoff Publishing, Groningen, 1972.
[27] Porteous, I. R. *Topological Geometry*, second edition, Cambridge University Press, 1981.
[28] Ricci, F. and Weiss, G. A characterization of $H^1(\Sigma_{n-1})$, *Proc. Symp. Pure Math.*, XXXV, (1979), 289-294.
[29] Seeley, R. T. Elliptic singular integral equations, *Proc. Symp. Pure Math.*, X (1967), 308-315.
[30] Solomjak, M. Z. Linear elliptic systems of the first order, *Soviet Doklady*, **4**(1963), 604-607.
[31] Stein, E. M. and Weiss, G. On the theory of harmonic functions of several variables, I. The theory of H^p-spaces, *Acta Math.* **103**(1960), 25-62.
[32] Stein, E. M. and Weiss, G. Generalization of the Cauchy-Riemann equations and representations of the rotation group, *Amer. J. Math.* **90**(1968), 163-196.
[33] Stein, E. M. and Weiss, G. *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, 1971.
[34] Takahashi, R. Série discrète pour les groupes de Lorentz $SO_0(2n, 1)$, in *Colloque sur les Functions Sphériques et la Théorie des Groupes*, Nancy, 1971.
[35] Vekua, N. P. *Systems of Singular Integral Equations*, P. Noordhoff, Groningen, 1967.
[36] Verchota, G. C. Layer potentials and boundary value problems for Laplace's equation in Lipschitz domains, *J. Funct. Anal.* **59**(1984), 572-611.

John E. Gilbert*
Department of Mathematics
University of Texas at Austin
Austin, TX 78712.
U.S.A.

Margaret A. M. Murray†
Department of Mathematics
Virginia Polytechnic Institute
and State University
Blacksburg, VA 24061.
U.S.A.

* Research partially supported by NSF grant DMS-8502352.
† Research partially supported by NSF grant DMS-8603234.

REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.º 2, 1988

BMO Harmonic Approximation in the Plane and Spectral Synthesis for Hardy-Sobolev Spaces

Joan Mateu and Joan Verdera

En record de José Luis Rubio de Francia

Introduction

The spectral synthesis theorem for Sobolev spaces of Hedberg and Wolff [7] has been applied in combination with duality, to problems of L^q approximation by analytic and harmonic functions. In fact, such applications were one of the main motivations to consider spectral synthesis problems in the Sobolev space setting. In this paper we go the opposite way in the context of the *BMO*-H^1 duality: we prove a *BMO* approximation theorem by harmonic functions and then we apply the ideas in its proof to produce a spectral synthesis result for variants of Sobolev spaces involving the Fefferman-Stein Hardy space H^1. It should be mentioned that our techniques work only in dimension 2, one of the reasons being that the fundamental solution for the Laplacian is in BMO only in the plane (see below for a more detailed discussion of this fact). Our main result reads as follows.

Theorem. *Let $X \subset \mathbb{C}$ be compact and let $f \in VMO(\mathbb{C})$ be harmonic on $\overset{\circ}{X}$. Then we can find a sequence (f_n) in $VMO(\mathbb{C})$, each f_n being harmonic on some neighbourhood of X, such that*

$$f_n \to f \quad in \quad BMO(\mathbb{C}).$$

There is a more conventional way of stating the above theorem, which we now proceed to explain. To each measurable subset E of $\mathbb{R}^n$ one associates the restriction spaces

$$BMO(E) = BMO(\mathbb{R}^n)|_E \simeq BMO(\mathbb{R}^n)/I(E)$$

and

$$VMO(E) = VMO(\mathbb{R}^n)|_E \simeq VMO(\mathbb{R}^n)/J(E),$$

where

$$I(E) = \{f \in BMO(\mathbb{R}^n): f = 0 \text{ a.e. on } E\}$$

and

$$J(E) = \{f \in VMO(\mathbb{R}^n): f = 0 \text{ a.e. on } E\}.$$

Endowed with the quotient norm, $BMO(E)$ and $VMO(E)$ become Banach spaces (modulo constants). T. Wolff has identified $BMO(E)$ with the set of measurable functions on E which satisfy the conclusions of the John-Nirenberg Theorem [4, p. 440]. As a result, functions in $BMO(E)$ (and their norms) can be described in a very concrete way in terms of their values on E. The analogous task for $VMO(E)$ has been achieved by Holden [8].

Consider now a compact $X \subset \mathbb{C}$ and let $H(X)$ be the closure in $BMO(X)$ or those functions which are harmonic on neighbourhoods of X. Holden's extension theorem implies that $H(X)$ is contained in $VMO(X)$, and consequently in

$$h(X) = VMO(X) \cap \{f: \Delta f = 0 \text{ on } \overset{\circ}{X}\}.$$

We can now restate our main result.

Theorem 1. *$H(X) = h(X)$, for any compact $X \subset \mathbb{C}$.*

The main point is, of course, than the compact X is arbitrary. This is in strong contrast with the well-known results on uniform harmonic approximation due to Keldysh [10] and Deny [3]. Using them, one can easily construct examples of compacta X such that not all continuous functions on X which are harmonic on $\overset{\circ}{X}$ can be uniformly approximated on X by functions harmonic on neigbourhoods of X. One of the reasons for this difference is that the fundamental solution for the Laplacian in the plane is in BMO, but not in L^∞. But this fact alone cannot provide a full explanation for Theorem 1, as becomes clear when one considers the L^p harmonic approximation problem: $\log|z|$ is still (locally) in L^p, but the natural L^p version of Theorem 1 does

not hold for $2 \leqslant p$, according to a clever example of Hedberg [6, p. 77]. It turns out that the gradient of the fundamental solution plays a role too. As shown in [17], the capacity associated to the kernel $1/z = \nabla(\log|z|)$ and to *VMO* vanishes on sets of σ-finite length, and this is the reason why Hedberg's construction fails to give a counter-example in the *BMO* case.

As far as techniques are concerned, we use a refinement of Vitushkin localization and matching coefficients methods and a final duality argument, which takes full advantage of an old result of Kolmogorov and Vercenko in geometric measure theory, via a differentiability theorem for functions with second derivatives in $H^1(\mathbb{R}^2)$ due to J. Dorronsoro.

We come now to spectral synthesis. We wish to consider functions on $\mathbb{R}^n$ with all derivatives of order s in $H^1(\mathbb{R}^n)$. A convenient way of doing that is to define, for $0 < s \leqslant n$,

$$I_s H^1(\mathbb{R}^n) = \{I_s * h: h \in H^1(\mathbb{R}^n)\},$$

where I_s is the Riesz potential of order s, that is, $I_s(x) = |x|^{-n+s}$ if $0 < s < n$ and $I_s(x) = \log|x|$ if $s = n$. We endow $I_s H^1(\mathbb{R}^n)$ with the Banach space norm $\|f\| = \|h\|_{H^1}$, $f = I_s * h$. Recent results of Adams [1] show that functions in $I_s H^1(\mathbb{R}^n)$ can be well defined except for sets of zero $n - s$ dimensional Hausdorff measure (they are continuous for $s = n$), and when so strictly defined they enjoy some continuity properties measured by the $n - s$ dimensional Hausdorff content M^{n-s}. Assume now that a closed set E in $\mathbb{R}^n$ is fixed and a function $f \in I_s H^1(\mathbb{R}^n)$ is given with the property that for some $\varphi_j \in C_0^\infty(E^c)$, $\varphi_j \to f$ in $I_s H^1(\mathbb{R}^n)$. Then it is not difficult to show that

$$\partial^\alpha f(x) = 0, \quad \text{for} \quad M^{n-s+|\alpha|} - \text{almost all } x \in E, \ |\alpha| \leqslant s - 1. \tag{1}$$

We believe that the following spectral synthesis theorem is true.

Conjecture. Let $f \in I_s H^1(\mathbb{R}^n)$ safisfy (1). Then there exist $\varphi_j \in C_0^\infty(E^c)$ with

$$\varphi_j \to f \quad \text{in} \quad I_s H^1(\mathbb{R}^n).$$

For $s = 1$ and any n the above statement has been shown to be true very recently by J. Orobitg [11] using truncation results of S. Janson for $I_1 H^1(\mathbb{R}^n)$. Our contribution here is to verify the conjecture for the first case in which truncation is not available, that is, for $n = s = 2$.

Theorem 2. *Let $E \subset \mathbb{R}^2$ be closed and let $f \in I_2 H^1(\mathbb{R}^2)$ satisfy*

$$f(x) = 0, \qquad x \in E,$$

and

$$\nabla f(x) = 0, \quad \textit{for } M^1 - \textit{almost all } x \in E.$$

Then there is a sequence (φ_j) *in* $C_0^\infty(E^c)$ *such that*

$$\varphi_j \to f \quad in \quad I_2H^1(\mathbb{R}^2).$$

Again our techniques are a combination of the constructive methods of Vitushkin with duality arguments exploiting ideas coming from the proof of Theorem 1.

Section 1 contains the constructive part of the proof of Theorem 1: after some preliminary lemmas an approximation theorem by potentials of measures is proven. In Section 2 the duality argument completing the proof of Theorem 1 is presented. Theorem 2 is dealt with in Section 3.

We close this introductory section by establishing some notation and recalling a few well-known facts which will be used throughout the paper.

1. *Hausdorff content.* Let $h(t)$, $t > 0$, be a non-decreasing function. For $E \subset \mathbb{R}^n$ we set

$$M^h(E) = \inf \sum_j h(\sigma_j), \tag{2}$$

where the infimum is taken over all countable coverings of E by squares with sides of length σ_j and parallel to the coordinate axis. When $h(t) = t^\alpha$, $0 < \alpha$, one writes M^α instead of M^h. We will need also the dyadic version M_d^h of M^h, which is defined by the right-hand side of (2) where now the infimum is over all coverings of E by dyadic squares of side length σ_j. It is clear that for some constant C depending on n,

$$M^h(E) \leqslant M_d^h(E) \leqslant CM^h(E), \quad \text{for any} \quad E.$$

2. *Cauchy and Beurling transforms.* They are the operators defined respectively by

$$Cf = \frac{1}{z} * f \quad \text{and} \quad Bf = P.V.\frac{1}{z^2} * f,$$

where f is an appropriate function on the plane.

Then we have $B = -\partial C$, where ∂ is the conjugate of the Cauchy-Riemann operator $\bar{\partial}$.

3. For $f \in BMO(\mathbb{R}^2)$ we shall write

$$\|f\|_* = \sup \frac{1}{|\Delta|} \int_\Delta |f - f_\Delta|, \tag{3}$$

where $|\Delta|$ is the area of Δ, f_Δ is the mean value of f on Δ and the supremum

is taken over all discs Δ. If D is a disc, we let $\|f\|_{*,D}$ stand for the right-hand side of (3), now the supremum being taken just on the discs Δ contained in D.

4. The letter C will denote a constant, which may be different at each occurence and which is independent of the relevant variables under consideration.

1. Approximation by Potentials of Measures

A family (E_j) of subsets of the plane is said to be *almost disjoint with constant* N whenever each $z \in \mathbb{C}$ belongs to at most N sets E_j.

Our first lemma was discovered in [2]. However, a (slightly different) proof of it will be presented here for the reader's convenience.

Lemma 1.1. *Let (Δ_j) be a finite family of open discs such that for some $\lambda > 1$ $(\lambda\Delta_j)$ is almost disjoint with constant N. Let $h_j \in BMO(\mathbb{C})$ be harmonic outside a compact subset of Δ_j and assume that*

$$h_j(z) = O(|z|^{-2}) \quad \text{as} \quad z \to \infty.$$

Then

$$\left\| \sum_j h_j \right\|_* \leqslant C \max_j \|h_j\|_* \tag{4}$$

for some positive constant $C = C(\lambda, N)$.

PROOF. We claim that for each j we can find $\chi_j \in BMO(\mathbb{C})$, with compact support contained in $\lambda\Delta_j$, such that

$$h_j = B(\chi_j) + \overline{B(\chi_j)} \quad \text{on} \quad (\lambda\Delta)^c$$

and

$$\|\chi_j\|_* \leqslant C\|h_j\|_*.$$

To prove the claim fix j and set $h = h_j$, $\Delta = \Delta_j$. Let us assume that Δ is centered at the origin at let δ be its radius. On $(\lambda\Delta)^c$ h has the expansion

$$h(z) = \sum_{n=2}^{\infty} a_n z^{-n} + \sum_{n=2}^{\infty} \bar{a}_n \bar{z}^{-n},$$

where $a_n = n^{-1}\langle \Delta h, z^n \rangle$, the bracket meaning the action of the compactly supported distribution Δh on the function z^n. Set

$$\lambda_1 = 1 + \frac{1-\lambda}{3} \quad \text{and} \quad \lambda_2 = 1 + \frac{2(1-\lambda)}{3}.$$

We need now the estimate

$$|a_n| \leqslant C\lambda_1^n \delta^n \|h\|_*. \tag{5}$$

A quick argument to get (5) is the following. Let $\varphi \in C_0^\infty(\lambda_1\Delta)$, $\varphi = 1$ on Δ, $|\partial^\alpha \varphi| \leqslant C\delta^{-|\alpha|}$, $0 \leqslant |\alpha| \leqslant 2$.

We have

$$\begin{aligned} n|a_n| &= |\langle \Delta h, \varphi(z)z^n\rangle| \\ &= |\langle h, \Delta(\varphi(z)z^n)\rangle| \\ &= |\langle h - h_{\lambda_1\Delta}, \Delta(\varphi(z)z^n)\rangle| \\ &\leqslant \int_{\lambda_1\Delta} |h - h_{\lambda_1\Delta}|\, |\Delta(\varphi(z)z^n)| \\ &\leqslant Cn\lambda_1^n\delta^n\|h\|_*, \end{aligned}$$

which is (5).

Take now $\varphi \in C^\infty(\mathbb{C})$, $\varphi = 0$ on $\lambda_2\Delta$, $\varphi = 1$ on $(\lambda\Delta)^c$, $|\nabla\varphi| \leqslant C\delta^{-1}$. Then

$$\frac{\varphi(z)}{z^{n-1}} = C\left(\frac{\bar\partial\varphi}{z^{n-1}}\right), \qquad n = 2, 3, \ldots,$$

and so

$$\frac{1}{z^{n-1}} = C\left(\frac{\bar\partial\varphi}{z^{n-1}}\right) \quad \text{on} \quad (\lambda\Delta)^c. \tag{6}$$

Differentiating with respect to z both sides of (6) we obtain

$$\frac{1}{z^n} = B(\psi_n) \quad \text{on} \quad (\lambda\Delta)^c,$$

where $\psi_n = (n-1)^{-1} z^{-(n-1)}\bar\partial\varphi$. It is clear that

$$\|\psi_n\|_* \leqslant 2\|\psi_n\|_\infty \leqslant Cn^{-1}(\lambda_2\delta)^{-n}.$$

Set

$$\chi = \sum_{n=2}^{\infty} a_n\psi_n,$$

so that

$$\|\chi\|_* \leqslant \sum_{n=2}^{\infty} C(\lambda_1\delta)^n\|h\|_* n^{-1}(\lambda_2\delta)^{-n} = C\|h\|_*,$$

and $h(z) = B(\chi) + \overline{B(\chi)}$ on $(\lambda\Delta)^c$. The claim is then proven.

Define

$$b_j = h_j - B(\chi_j) - \overline{B(\chi_j)}.$$

Therefore $\|b_j\|_* \leqslant C\|h_j\|_*$ and $\operatorname{supp}(b_j) \subset \lambda\Delta_j$. Since

$$\begin{aligned}\Big\|\sum h_j\Big\|_* &\leqslant \Big\|\sum b_j\Big\|_* + 2\Big\|B\Big(\sum \chi_j\Big)\Big\|_* \\ &\leqslant \Big\|\sum b_j\Big\|_* + C\Big\|\sum \chi_j\Big\|_*,\end{aligned}$$

it suffices to prove Lemma 1.1 under the additional assumption that $\operatorname{supp}(h_j) \subset \lambda\Delta_j$ for each j, which is not difficult (see [2] for details). □

The next lemma is a variant of Lemma 1.1, in which we require less decay of the h_j at infinity but we have a packing condition on the family of discs Δ_j.

Lemma 1.2. *Let $\omega(t)$, $t > 0$, a non-decreasing function satisfying $\omega(2t) \leqslant C\omega(t)$, $t > 0$.* Let (Δ_j) *be a finite family of open discs of radii δ_j with the following properties.*

(i) *For some $\lambda > 1$ $(\lambda\Delta_j)$ is an almost disjoint family.*
(ii) *For any disc Δ of radius δ*

$$\sum_{\Delta_j \subset \Delta} \delta_j \omega(\delta_j) \leqslant C\delta\omega(\delta).$$

Let $h_j \in BMO(\mathbb{C})$ be harmonic outside a compact subset of Δ_j, $\|h_j\|_* \leqslant \omega(\delta_j)$ and $h_j(z) = O(|z|^{-1})$ as $z \to \infty$.

Then

$$\Big\|\sum_j h_j\Big\|_* \leqslant C\omega(d),$$

d being the diameter of $\bigcup_j \Delta_j$.

Remark. Notice that when $\omega \equiv 1$ the conclusion of the lemma is essentially (4) and the hypothesis (ii) can be regarded as a linearity condition on the family (Δ_j).

PROOF. We are going to perform a reduction to Lemma 1.1. To this aim expand h_j at ∞

$$h_j(z) = \frac{a_j}{z} + \frac{\bar{a}_j}{\bar{z}} + O(|z|^{-2}),$$

and notice that

$$|a_j| \leqslant C\delta_j\|h_j\|_* \leqslant C\delta_j\omega(\delta_j).$$

Set

$$\mu_j = a_j \left| \frac{1}{2}\Delta_j \right|^{-1} \chi_{\frac{1}{2}\Delta_j}$$

and

$$P_j = C(\mu_j) + \overline{C(\mu_j)},$$

so that

$$h_j = P_j + O(|z|^{-2}).$$

Since

$$\Big\| \sum h_j \Big\|_* \leqslant \Big\| \sum_j h_j - P_j \Big\|_* + \Big\| \sum P_j \Big\|_*,$$

by Lemma 1.1 it is clearly enough to show that

$$\|P_j\|_* \leqslant C \|h_j\|_* \tag{7}$$

and

$$\Big\| \sum P_j \Big\|_* \leqslant C\omega(d). \tag{8}$$

Now, (7) is a consequence of the well known inequality [9]

$$\|C(\mu)\|_* \leqslant C \sup \{ |\mu|(\Delta(z, r) r^{-1}) : z \in \mathbb{C}, r > 0 \}, \tag{9}$$

μ being any locally finite measure, and (8) follows also from (9) provided we ascertain that for each disc $\Delta(z, r)$

$$\sum_j |\mu_j|(\Delta(z, r)) \leqslant C\omega(d) r.$$

This can be done as follows.

$$\sum |\mu_j|(\Delta(z, r)) = \sum_{r \leq \delta_j} + \sum_{r > \delta_j} \leqslant C\omega(d) \sum_{r \leq \delta_j} r^{-1} |\Delta(z, r) \cap \Delta_j| + C \sum_{\Delta_j \subset \Delta(z, 3r)} \delta_j \omega(\delta_j)$$

$$= I + II.$$

Clearly $I \leqslant C\omega(d)r$. If $r \leqslant d$ we estimate II by $Cr\omega(r) \leqslant Cr\omega(d)$. Otherwise, given any $w \in \bigcup \Delta_j$, II can be estimated by

$$\sum_j \delta_j \omega(\delta_j) = \sum_{\Delta_j \subset \Delta(w, d)} \delta_j \omega(\delta_j) \leqslant C\, d\omega\,(d) \leqslant Cr\omega(d). \quad \square$$

We come now to a result which gives a simple but useful device to produce functions with a double zero at ∞, as required by Lemma 1.1.

Lemma 1.3. *Let $\delta > 0$ and $0 < p \in \mathbb{Z}$. Let z_1, z_2, z_3 be three points satisfying*

$$|z_j - z_k| \leqslant C\delta, \qquad j, k = 1, 2, 3,$$

and

$$\min\{|z_1 - z_2|, d(z_3, l(z_1, z_2))\} \geqslant \delta/p,$$

where $l(z_1, z_2)$ is the straight line joining z_1 and z_2. Assume $a \in \mathbb{R}$, $b \in \mathbb{C}$ and

$$|a| \leqslant CN, \qquad |b| \leqslant C\delta N$$

for some $N > 0$. Then there exists a function g satisfying $\Delta g = 0$ on $\mathbb{C} \setminus \{z_1, z_2, z_3\}$,

$$g(z) = a \log|z| + \frac{b}{z} + \frac{\bar{b}}{z} + O(|z|^{-2}), \quad \text{as} \quad z \to \infty$$

and

$$\|g\|_* \leqslant Cp^2 N.$$

PROOF. Without loss of generality we can assume $z_1 = 0$, $z_2 = r > 0$. Set

$$g = \log|z| * \mu, \qquad \mu = \lambda_1 \delta_0 + \lambda_2 \delta_r + \lambda_3 \delta_{z_3},$$

where δ_z is the Dirac measure at the point z and the λ_j are complex numbers to be determined. The function g has the required expansion at ∞ if

$$\begin{aligned} a &= \lambda_1 + \lambda_2 + \lambda_3 \\ -b &= \lambda_2 r + \lambda_3 z_3 \\ -\bar{b} &= \lambda_2 r + \lambda_3 \bar{z}_3. \end{aligned}$$

Solving for the λ_j, we get the estimates

$$|\lambda_j| \leqslant Cp^2 N, \qquad j = 1, 2, 3,$$

which give the desired bound for $\|g\|_*$. □

The next lemma is the bulk of our technical arsenal. It deals with a covering property which seems to be of interest by itself (see the remark below).

Main Lemma 1.4. *Let $h(t) = t\omega(t)$ be a measure function with ω non-decreasing and satisfying $\omega(2t) \leqslant C\omega(t)$. Then for any compact set $K \subset \mathbb{C}$ there exists a finite family of discs (Δ_j) which can be divided into two subfamilies (Δ_j^g) and (Δ_j^b) (the superscripts g and b stand for good and bad) in such a way that the following holds.*

(*a*) $K \subset \bigcup_j \Delta_j$.

(*b*) *For some* $\lambda = \lambda(h) > 1$, $(\lambda\Delta_j)$ *is an almost disjoint family with constant depending on h but not on K.*

(*c*) *Each* Δ_j^g *has the three points property, that is, there exist* $z_1, z_2, z_3 \in \Delta_j^g \cap K$ *such that*

$$\min\{|z_1 - z_2|, d(z_3, l(z_1, z_2))\} > \eta\delta_j^g,$$

where δ_j^g *is the radius of* Δ_j^g, $\eta = \eta(h) > 0$ *and* $l(z_1, z_2)$ *is the straight line through* z_1 *and* z_2.

(*d*) $\sum_j h(\delta_j^b) \leqslant CM^h(K)$, *where C depends on h but not on K.*

(*e*) *For each disc* Δ *of radius* δ

$$\sum_{\Delta_j^b \subset \Delta} h(\delta_j^b) \leqslant Ch(\delta),$$

where C depends on h but not on K.

Remark. We do not know, even for $h(t) = t$, when a family (Δ_j) can be constructed satisfying (*a*), (*b*) and (*d*), (*e*) with Δ_j^b replaced by Δ_j. This is true when $h(t)$ tends to zero fast enough, for example for $h(t) = t^{1+\epsilon}$, $\epsilon > 0$.

PROOF OF THE MAIN LEMMA. The first step consists in finding a family of discs D_j of radius r_j such that

(i) $K \subset \bigcup_j D_j$

(ii) $\sum_j h(r_j) \leqslant CM^h(K)$

(iii) $\sum_{D_j \subset D} h(r_j) \leqslant Ch(r)$, for each disc D of radius r.

To construct the D_j we consider a family (Q_j) of dyadic squares of side length σ_j satisfying $K \subset \bigcup_j Q_j$ and $\sum_j h(\sigma_j) \leqslant 2M_d^h(K)$. It is easy to modify that family so that in addition one has

$$\sum_{Q_j \subset Q} h(\sigma_j) \leqslant h(\sigma), \tag{10}$$

for each dyadic square Q of side length σ. In fact, if for such a Q (10) fails then we remove the Q_j contained in Q from our family and we put Q in it. Take now as D_j the disc circumscribed to Q_j.

Unfortunately property (*b*) with Δ_j replaced by D_j does not necessarily hold. The idea is to construct a «halo» around each D_j, consisting of smaller discs which will be called «defenses» because their effect is to prevent too much overlapping. The difficulty will be then to preserve properties (ii) and (iii).

Generation of defenses: step 1. Let ϵ be a small positive number to be determined later, such that ϵ^{-1} is an integer. Choose one of the discs D_j of maximal size. After relabelling we can assume it is D_1. We distinguish two cases.

Case 1. D_1 has the three points property, that is, (*c*) holds with Δ_j^b replaced by D_1 and $\eta = \epsilon^4/2$. D_1 will be a member of the family (Δ_j^g) and it will be surrounded by some defenses which now we proceed to define. Consider ϵ^{-3} points equally spaced on the boundary of D_1, and define the defenses of D_1 to be the discs of radii $\epsilon^2 r_1$ centered at these points. Then the defenses of D_1 form an almost disjoint family with constant comparable to ϵ^{-1}.

We now define two new families of discs, $\mathcal{F}_1^d$ (the superscript d is for definitive) and $\mathcal{F}_1^p$ (the superscript p is for process). Only the family $\mathcal{F}_1^p$ will be subject to a new process of generation of defenses. $\mathcal{F}_1^d$ consists of the single disc D_1 and $\mathcal{F}_1^p$ consists of the defenses of D_1 plus the discs D_j which are not contained in the union of D_1 and its defenses.

Case 2. D_1 has not the three points property. In this case $K \cap D_1$ is contained in a strip S of width $\epsilon^4 r_1$.

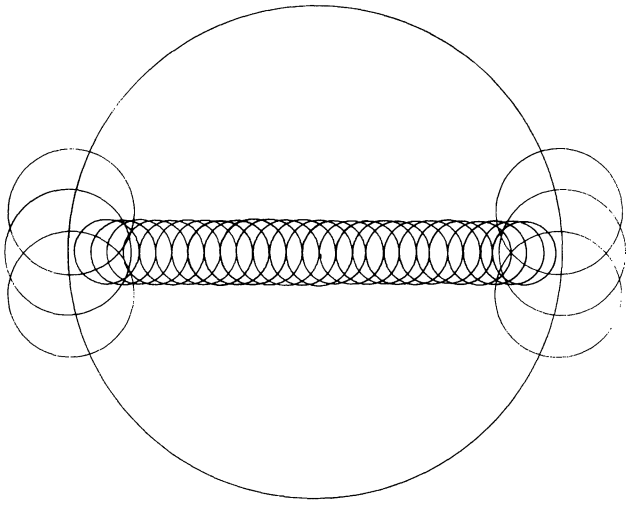

Figure 1

Observe that the length of $S \cap \partial D_1$ is at most $C\epsilon^2 r_1$. Cover $S \cap D_1$ by discs of radii $\epsilon^4 r_1$ in such a way that they form an almost disjoint family with some numerical constant. These discs will form the family $\mathcal{F}_1^q$. Consider, as in Case 1, ϵ^{-3} equally spaced discs of radii $\epsilon^2 r_1$ centered on ∂D_1, and take as defenses of D_1 those that intersect $S \cap \partial D_1$. In the family $\mathcal{F}_1^p$ we put the

defenses of D_1 plus the discs D_j different from D_1 which are not contained in the union of the defenses of D_1 and of the discs in $\mathcal{F}_1^d$.

Generation of defenses: step 2. We now examine the family $\mathcal{F}_1^p$. We take a disc D in $\mathcal{F}_1^p$ of maximal size and distinguish two cases.

Case 1. D has the three points property. If D is one of the discs D_j we proceed as in step 1, but this time, we only use those defenses which are not already contained in the union of the discs in $\mathcal{F}_1^d \cup \mathcal{F}_1^p$.

If D is not a D_j then D is a defense of D_1. In this case let C be the circumference passing through the points of intersection of the boundaries of pairs of defenses of D_1 (see figure 2). If r is the radius of D we consider a family $\mathfrak{D}$ of discs of radii $\epsilon^2 r$ with equally spaced centers on C, which is almost disjoint with constant comparable to ϵ^{-1}. As shown in figure 2 just an arc γ of ∂D of length comparable to ϵr is in the exterior of the disc bounded by C.

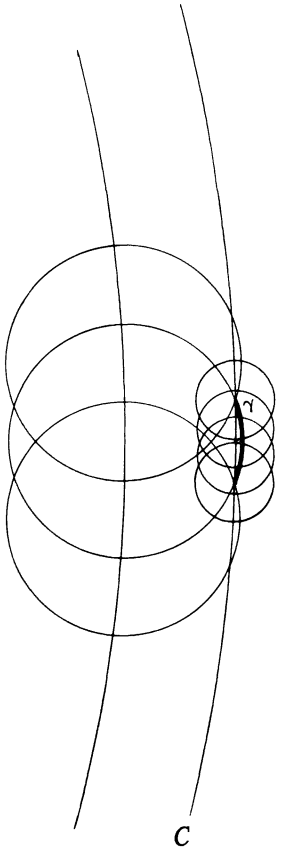

Figure 2

The defenses of D are the discs in the family $\mathfrak{D}$ which intersect γ. The family $\mathfrak{F}_2^d$ consists of $\mathfrak{F}_1^d$ plus the disc D, and the family $\mathfrak{F}_2^p$ consists of the defenses of D and the discs in $\mathfrak{F}_1^p$ which are not contained in the union of $\mathfrak{F}_2^d$ and the defenses of D.

Case 2. D has not the three points property. We proceed as in step 1, the only difference being that the defenses of D are now centered on C when D itself is a defense of D_1 (see figure 3).

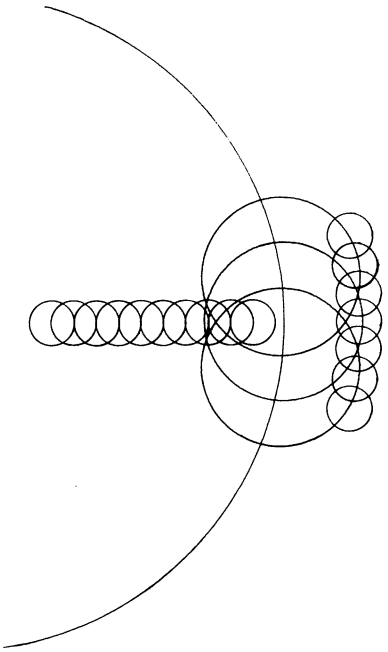

Figure 3

Proceeding inductively we produce at the n-th step families $\mathfrak{F}_n^d$ and $\mathfrak{F}_n^p$. We stop when all discs D_j have been either rejected or subject to a process of generation of defenses, that is, we stop at the N-th step provided $\mathfrak{F}_N^p$ does not include any D_j. We claim that $\mathfrak{F}_N^d = \{\Delta_j\}$ is the family we are looking for. We say that a disc $\Delta_j \in \mathfrak{F}_N^d$ is a good disc, that is, a disc Δ_j^g, if it has the three points property. Otherwise it is a bad disc Δ_j^b. Therefore, properties (*a*) and (*c*) clearly hold. Let us proceed to prove (*b*).

Claim. Let $\lambda = 1 + \epsilon^2/4$. Then for ϵ small enough $\{\lambda\Delta_j\}$ is an almost disjoint family with constant depending on ϵ.

PROOF OF THE CLAIM. In the typical case of generation of defenses we start with a disc D_j of radius δ. At the n-th step we will produce defenses of radii $\epsilon^{2n}\delta$ centered on a circumference C_{n-1} concentric with D_j. Let L_n be the boundary of the smallest disc concentric with D_j containing the dilates with factor λ of the defenses in the n-th generation.

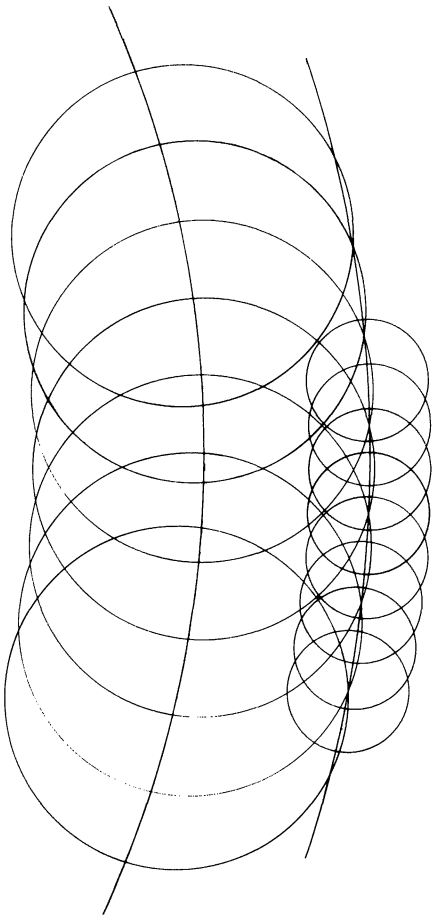

Figure 4

An easy but cumbersome computation shows that for ϵ small enough

$$\text{radius } L_n < \text{radius } C_n + \delta\epsilon^{2n+2}/2.$$

This inequality guarantees that to travel from the exterior of the disc bounded by C_{n+1} to a dilate of a disc D in the n-th generation one has to cover a distance not less than $C\epsilon^2$ radius D.

Let $z \in \mathbb{C}$ and assume that $z \in \lambda\Delta_j$ and that Δ_j is of maximal size among the discs with this property. The choice of ϵ (precisely, the just mentioned travelling property) shows that if $z \in \lambda\Delta_k$ then

$$\text{radius } \Delta_k \geqslant C\epsilon^2 \text{ radius } \Delta_j.$$

Thus there is a finite number $p = p(\epsilon)$ of possible values for the radii of the Δ_k such that $z \in \lambda\Delta_k$. Since the discs $\lambda\Delta_k$ of a fixed size form an almost disjoint family with constant depending on ϵ, the claim is proven.

We turn now to the proof of (d). We wish to show that

$$\sum_j h(\delta_j^b) \leqslant C \sum_k h(r_k), \tag{11}$$

but before it will be convenient to introduce some terminology to simplify the exposition of the argument.

Assume that the process of generation of defenses has been applied to some disc D. The first generation descendants of D are either the defenses of D or the discs which have been used to cover the strip S in the case that D has not the three points property. We define inductively the i-th generation descendants of D, $2 \leqslant i$, in the obvious way. The halo of D, Halo (D), is the set of all descendants of D. A disc D is an ancestor of a disc D^* if $D^* \in \text{Halo}(D)$. For $D \in \{\Delta_j\}$ and a non-negative integer n, $A_n(D)$ is the set of bad discs Δ_j^b which have got exactly n bad ancestors (a bad ancestor is an ancestor without the three points property) different from D in the halo of D.

Fix now $D \in \{\Delta_j\}$ and let r be its radius. We claim that

$$\sum_{\Delta_j^b \in A_0(D)} h(\delta_j^b) \leqslant C\epsilon^{-1} h(r), \quad \text{if } D \text{ is good,} \tag{12}$$

and

$$\sum_{\Delta_j^b \in A_0(D)} h(\delta_j^b) \leqslant C\epsilon h(r), \quad \text{if } D \text{ is bad.} \tag{13}$$

To show (12) observe first that the halo of D is contained in $2D$. Consider a disc Δ_j^b in $A_0(D)$. Project the part of $\partial\Delta_j^b$ which will be eventually covered by defenses of Δ_j^b onto an interval I_j on $\partial(2D)$ as shown in figure 5, and notice that the length of I_j is comparable to $\epsilon\delta_j^b$.

The intervals I_j are disjoint because any disc Δ_j intersecting the shaded region R is not in $A_0(D)$. Thus

$$\sum_{\Delta_j^b \in A_0(D)} h(\delta_j^b) \leqslant C\epsilon^{-1}\omega(r) \sum_j \text{length } (I_j) \leqslant C\epsilon^{-1} r\omega(r).$$

The proof of (13) is similar, the only difference being that if D is bad then the halo of D emanates from an arc of ∂D of length comparable to $\epsilon^2 r$, and so $\sum_j$ length $(I_j) \leqslant C\epsilon^2 r$.

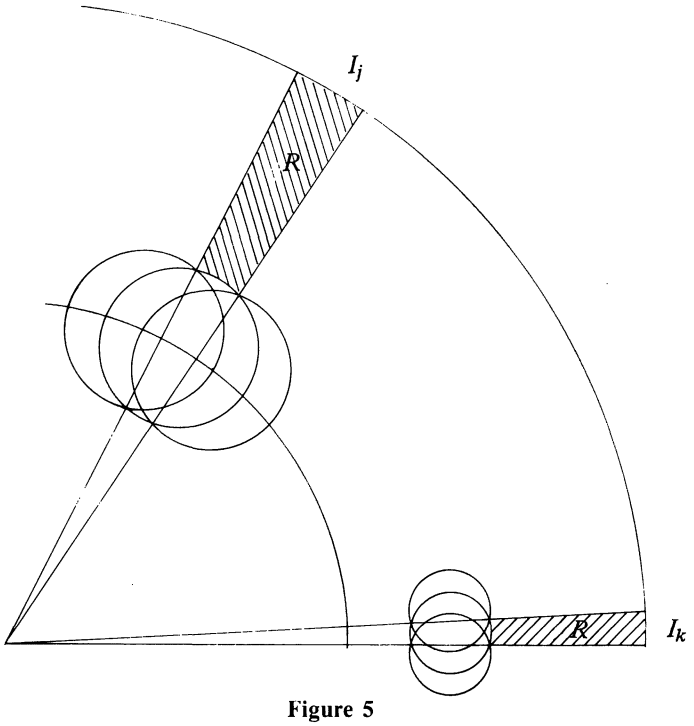

Figure 5

Fix now a D_k. One has

$$\sum_{\Delta_j^b \in \text{Halo}(D_k)} h(\delta_j^b) = \sum_{n=0}^{\infty} \sum_{\Delta_j^b \in A_n(D_k)} h(\delta_j^b) \leqslant \sum_{n=0}^{\infty} (C\epsilon)^n \sum_{\Delta_j^b \in A_0(D_k)} h(\delta_j^b)$$
$$\leqslant C\epsilon^{-1}(1 - C\epsilon)^{-1} h(r_k),$$

where in the first inequality we applied (13) n times and in the second (12) once. The above inequality completes the proof of (11) provided ϵ is small enough so that $C\epsilon < 1$.

We are left with property (*e*). Fix a Δ_j^b. We wish to associate to it a certain ancestor $A(\Delta_j^b)$. We distinguish two cases.

Case 1. All ancestors of Δ_j^b are contained in 2Δ, where Δ is the test disc in (*e*). Then $A(\Delta_j^b)$ is, by definition, the oldest ancestor of Δ_j^b.

Case 2. At least one ancestor of Δ_j^b is not contained in 2Δ. In this case there is a youngest ancestor not contained in Δ, say D. If the center of D is not in Δ, then we set $A(\Delta_j^b) = D$. Otherwise the «father» D^* of D intersects Δ and its center is outside. We set $A(\Delta_j^b) = D^*$.

Let I and II be the families of ancestors so constructed which are in cases 1 and 2 respectively. We get then

$$\sum_{\Delta_j^b \subset \Delta} h(\delta_j^b) = \Big(\sum_{A \in I} \sum_{A(\Delta_j^b) = A}\Big) h(\delta_j^b) + \sum_{A \in II} \sum_{A(\Delta_j^b) = A} h(\delta_j^b) = S_1 + S_2.$$

As we know, S_1 can be estimated by

$$\sum_{A \in I} Ch(\text{radius } A) \leqslant C \sum_{D_k \subset 2\Delta} h(r_k) \leqslant Ch(\delta).$$

To estimate S_2 fix $A \in II$ and let $\sigma(A) = \partial A \cap \Delta$ and $\sigma(\Delta) = \partial\Delta \cap A$. The halo of A contained in Δ emanates from $\sigma(A)$, and so

$$S_2 \leqslant C\omega(\delta) \sum_{A \in II} \text{lenght } \sigma(A) \leqslant C\omega(\delta) \sum_{A \in II} \text{length } \sigma(\Delta) \leqslant C\delta\omega(\delta),$$

where in the last inequality we used the fact that the family (Δ_j) is almost disjoint, and in the next to the last that $\sigma(A)$ and $\sigma(\Delta)$ have comparable lengths owing to the relative positions of A and Δ. □

We still need another lemma. To any $\varphi \in C_0^\infty(\mathbb{C})$ one associates the Vitushkin localization operator (see [18, p. 168])

$$V_\varphi f = \frac{1}{2\pi} \log|z| * \varphi\Delta f,$$

where f is any distribution in the plane.

Lemma 1.5. *Let $\varphi \in C_0^\infty(\Delta)$, Δ a disc of radius δ. We then have for any $f \in BMO(\mathbb{C})$*

$$\| V_\varphi f \|_* \leqslant C(\varphi) \| f \|_{*, 3\Delta},$$

where

$$C(\varphi) = C \sum_{|\alpha| \leq 2} \delta^{|\alpha|} \| \partial^\alpha \varphi \|_\infty.$$

PROOF. Using

$$\varphi\partial\bar{\partial} f = \partial\bar{\partial}(\varphi f) - (f\partial\bar{\partial}\varphi + \partial\varphi\bar{\partial} f + \bar{\partial}\varphi\partial f)$$

we readily get

$$V_\varphi f = \varphi f - \frac{1}{\pi}\frac{1}{z} * f\bar{\partial}\varphi - \frac{1}{\pi}\frac{1}{\bar{z}} * f\partial\varphi + \frac{1}{2\pi}\log|z| * f\Delta\varphi.$$

The *BMO*-norm estimate for the first three terms can be found in [17, 3.2. p. 290]. The *BMO*-norm of the last one is less than

$$C\int |f|\,|\Delta\varphi| \leqslant C\,\frac{1}{|\Delta|}\int |f|.$$

Since $V_\varphi f = V_\varphi(f - f_\Delta)$, we can replace f by $f - f_\Delta$ in the previous inequality, and this completes the proof of the lemma.

Given a compact $X \subset \mathbb{C}$ let $P(X)$ be the linear span in $BMO(X)$ of the functions of the type $\log|z-a|$, $a \notin X$, and $Re\,((1/z) * \mu)$, where μ is a complex Borel measure with compact support disjoint from $\overset{\circ}{X}$, and satisfying the growth condition

$$|\mu|(\Delta(z,r)) \leqslant \epsilon(r)r, \qquad z \in \mathbb{C}, r > 0,$$

for some $\epsilon(r) \to 0$ as $r \to 0$. Since $(1/z) * \mu$ is in $VMO(\mathbb{C})$ [9], $P(X)$ is a subspace of $h(X)$.

Theorem 1.6. *For each compact $X \subset \mathbb{C}$, $P(X)$ is dense in $h(X)$.*

PROOF. Let $f \in h(X)$. We can think that in fact $f \in VMO(\mathbb{C})$ and $\Delta f = 0$ on $\overset{\circ}{X}$. Replacing f by φf where $\varphi \in C^\infty(\mathbb{C})$ takes the value 1 on a neighbourhood of X, we can assume also that f is compactly supported. Fix a $\delta > 0$ and let (D_k, φ_k, f_k) be a δ-Vitushkin scheme for the approximation of f. This means (cf. [18, p. 168]) the following.

1. Each D_k is an open disc of radius δ and (D_k) is an almost disjoint covering of the plane.
2. $\varphi_k \in C_0^\infty(D_k)$, $\sum \varphi_k = 1$ on $\mathbb{C}$ and $|\partial^\alpha \varphi_k| \leqslant C\delta^{-|\alpha|}$, $0 \leqslant |\alpha| \leqslant 2$.
3. $f_k = \dfrac{1}{2\pi}\log|z| * \varphi_k \Delta f$, and so $f = \sum_k f_k$.

Notice that $f_k \equiv 0$ whenever $D_k \cap \operatorname{supp}(f) = \varnothing$, and thus only finitely many f_k do not vanish identically.

Choose a point $z_k \in D_k$ and expand f_k at ∞.

$$f_k(z) = a_k \log|z - z_k| + \frac{b_k}{z - z_k} + \frac{\bar{b}_k}{\bar{z} - \bar{z}_k} + O(|z|^{-2}).$$

Then we have $|a_k| \leqslant C\omega(\delta)$ and $|b_k| \leqslant C\delta\omega(\delta)$ where

$$\omega(\delta) = \sup\left\{\frac{1}{|\Delta|}\int_\Delta |f - f_\Delta| : \text{radius}\,\Delta \leqslant \delta\right\}$$

If we could improve the estimate for b_k to

$$|b_k| \leqslant CM^h(D_k \setminus \overset{\circ}{X}), \tag{14}$$

where $h(t) = t\omega(t)$, then the proof of Theorem 1.6 would be much simpler, as the reader will realize later. As a matter of fact (14) is not true. We get around this difficulty using a second localization for each f_k, whose purpose is to exploit the basic estimate $|b_k| \leqslant Ch(\delta)$ at a lower level, forcing in this way M^h to come on the scene.

To begin with the above program we claim that it suffices to construct $p_k \in P(X)$, p_k harmonic outside $5D_k$,

$$f_k = p_k + O(|z|^{-2}) \quad \text{as} \quad z \to \infty,$$

and

$$\| p_k \|_* \leqslant C\omega(\delta).$$

In fact, setting $p = \sum_k p_k \in P(X)$ we get by Lemma 1.1.

$$\| f - p \|_* = \left\| \sum f_k - p_k \right\| \leqslant C \max \| f_k - p_k \|_* \leqslant C\omega(\delta),$$

where the last inequality is a consequence of Lemma 1.5 and the well known fact that $\omega(2\delta) \leqslant C\omega(\delta)$ [14, p. 596]. To construct the p_k we proceed as follows. Since $\Delta f_k = \varphi_k \Delta f$, f_k is harmonic outside a compact subset K of $D_k \setminus \overset{\circ}{X}$. Let $\{\Delta_j\} = \{\Delta_j^g\} \cup \{\Delta_j^b\}$ be the family of discs given by Lemma 1.4 applied to K and $h(t) = t\omega(t)$. It is then clear that $\Delta_j \subset 5D_k$ for each j.

Observe now that, by construction, the sequence of different values of the radii of the Δ_j decreases as a geometric progression with ratio less than one. It is then clear that an appropriate variant of a Lemma of Harvey and Polking [5, 3.1, p. 43] can be applied to construct functions $\psi_j \in C_0^\infty(\lambda_1 \Delta_j)$, $\lambda_1 = (1 + \lambda)/2$, with $\sum_j \psi_j = 1$ on $\bigcup_j \Delta_j$ and $|\partial^\alpha \psi_j| \leqslant C\delta_j^{-|\alpha|}$, $0 \leqslant |\alpha| \leqslant 2$, where δ_j is the radius of Δ_j.

Set

$$f_{kj} = \frac{1}{2\pi} \log |z| * \psi_j \Delta f_k,$$

so that $f_k = \sum_j f_{kj}$. Fix now a j and distinguish two cases.

Case 1. $\Delta_j = \Delta_j^g$ is a good disc. Then one can find tree points as in (*c*) of Lemma 1.4. These points lie in $\Delta_j \setminus \overset{\circ}{X}$, but by a density argument they can be taken in $\Delta_j \setminus X$. Lemma 1.3. now produces a function $h_{kj}^g \in P(X)$, h_{kj}^g harmonic outside Δ_j^g, and such that $f_{kj}^g = h_{kj}^g + O(|z|^{-2})$, $\| h_{kj}^g \|_* \leqslant C \| f_{kj}^g \|_*$, where we have written f_{kj}^g instead to f_{kj}.

Case 2. $\Delta_j = \Delta_j^b$ is a bad disc. In this case we can suppose that $\Delta_j \setminus X$ is non-empty. Otherwise $\Delta_j \setminus \mathring{X}$ is also empty, and so Δ_j is not really necessary to cover K. Let $z_{kj} \in \Delta_j \setminus X$ and let f_{kj}^b stand for f_{kj}.

If

$$f_{kj}^b(z) = a_{kj} \log |z - z_{kj}| + O(|z|^{-1})$$

then we define

$$h_{kj}^b(z) = a_{kj} \log |z - z_{kj}|,$$

so that $h_{kj}^b \in P(X)$, h_{kj}^b is harmonic outside Δ_j^b, $f_{kj}^b = h_{kj}^b + O(|z|^{-1})$ and

$$\|h_{kj}^b\|_* \leqslant C \|f_{kj}^b\|_* \leqslant C\omega(\delta_j^b),$$

δ_j^b being the radius of $\Delta_j = \Delta_j^b$. Let

$$F_k = \sum_j h_{kj}^g + \sum_j h_{kj}^b.$$

Then $F_k \in P(X)$, F_k is harmonic outside $5D_k$, and $f_k = F_k + O(|z|^{-1})$ as $z \to \infty$.

The *BMO*-norm of F_k is estimated by $\|F_k - f_k\|_* + \|f_k\|_*$ and

$$\|F_k - f_k\|_* \leqslant \left\| \sum_j h_{kj}^g - f_{kj}^g \right\|_* + \left\| \sum_j h_{kj}^b - f_{kj}^b \right\|_* = I + II.$$

Applying Lemma 1.1 we get $I \leqslant C \max_j \|f_{kj}^g\|_* \leqslant C\|f_k\|_*$. As for the term *II* we must resort to Lemma 1.2. We obtain

$$II \leqslant C\omega(\delta),$$

because the diameter of the union of the Δ_j is less than 10δ. Collecting the above estimates we finally get

$$\|F_k\|_* \leqslant C\omega(\delta).$$

Consider the expansions

$$f_k(z) - F_k(z) = \frac{b}{z - z_k} + \frac{\bar{b}}{\bar{z} - \bar{z}_k} + O(|z|^{-2}),$$

$$f_{kj}^b(z) - h_{kj}^b(z) = \frac{b_{kj}}{z - z_{kj}} + \frac{\bar{b}_{kj}}{\bar{z} - \bar{z}_{kj}} + O(|z|^{-2}).$$

Then

$$|b| = \left| \sum_j b_{kj} \right| \leqslant C \sum \delta_j^b \omega(\delta_j^b) \leqslant CM^h(K).$$

Let now μ be a positive measure supported on K satisfying

$$\mu(\Delta(z, r)) \leqslant h(r), \qquad z \in \mathbb{C}, \qquad r > 0,$$

and

$$\|\mu\| \geqslant CM^h(K).$$

If we set

$$p_k = F_k + \operatorname{Re}\left(\frac{1}{z} * b\|\mu\|^{-1}\mu\right),$$

it is clear that p_k satisfies all required properties, and so the proof of Theorem 1.6 is complete.

2. End of the Proof of Theorem 1

The proof of Theorem 1 will be completed by a duality argument combined with a result of Kolmogorov and Vercenko in geometric measure theory (see Lemma 2.2. below). The link between our problem and geometry is provided by the next lemma, due to J. Dorronsoro who proved it in answer to a question of the authors.

Lemma 2.1. (Dorronsoro.) *Let* $F = \log|z| * h$, *with* $h \in H^1(\mathbb{R}^2)$. *Then* f *has an ordinary differential at* M^1*-almost all* $z \in \mathbb{R}^2$.

PROOF. Clearly $\nabla f = (1/z) * h$, and so ∇f can be defined M^1 a.e., being a Riesz potential of order 1 of a function in $H^1(\mathbb{R}^2)$ [1]. We will show that, for M^1-almost all $a \in \mathbb{R}^2$ and for all $z \neq a$, we have

$$|f(z) - f(a) - \nabla f(a) \cdot (z - a)|\,|z - a|^{-1} \leqslant Tf(a), \tag{15}$$

where Tf is an operator satisfying the weak-type estimate

$$M^1(\{z: Tf(z) > \lambda\}) \leqslant C\lambda^{-1}\|h\|_{H^1}.$$

Since functions in $C_0^\infty(\mathbb{R}^2)$ with zero integral are dense in $H^1(\mathbb{R}^2)$ a standard argument will then complete the proof of the Lemma.

To prove (15) fix $z \neq a$ and set $\delta = |z - a|$, $\Delta = \Delta(a, \delta)$. The left hand-side of (15) is estimated by

$$\delta^{-1}|f(z) - f_\Delta| + \delta^{-1}|f(a) - f_\Delta| + |\nabla f(a)|.$$

An elementary argument involving a regularization of f now gives (cf. [16, p. 125-126])

$$|f(z) - f_\Delta| \leqslant C \int_{|\zeta|<2\delta} |\zeta|^{-1} |\nabla f(z+\zeta)|\, dm(\zeta),$$

where m is planar Lebesgue measure. Let $0 \leqslant \varphi \in C_0^\infty(\Delta(z, 3\delta))$, $\varphi \equiv 1$ on $\Delta(z, 2\delta)$ and $|\nabla\varphi| \leqslant C\delta^{-1}$. Then

$$|f(z) - f_\Delta| \leqslant C(|\zeta|^{-1} * \varphi|\nabla f|)(z).$$

According to a remark of Stein [15, p. 385] the right hand-side of the above inequality is less than or equal to a constant times the norm of $\varphi|\nabla f|$ in the Lorenz space $L^{2,1}(\mathbb{R}^2)$, which in turn can be estimated by $C\|\nabla(\varphi\nabla f)\|_{L^1(\mathbb{R}^2)}$, owing to an imbedding theorem due to Poornima [12, Th. 1.4, p. 163]. Therefore

$$|f(z) - f_\Delta| \leqslant C|\Delta|^{-1} \int_{6\Delta} |\nabla f| + C \int_{6\Delta} |\nabla^2 f|.$$

The same inequality is, of course, true for $z = a$, and so the left hand-side of (15) is estimated by

$$CM(\nabla f)(a) + CM_1(\nabla^2 f)(a) + |\nabla f(a)| = I + II + III,$$

where M is the Hardy-Littlewood maximal operator and

$$M_1 g(a) = \sup_\delta \delta^{-1} \int_{|\zeta|<\delta} |g(a+\zeta)|\, dm(\zeta).$$

The required weak-type estimates for I and III follow from results of Adams [1, Theorem B] and they are obviously satisfied by II. □

To state the result of Kolmogorov and Vercenko mentioned above we need a definition. Let E be any subset of the plane and $a \in E$. A half-line $\{a + \rho v : \rho > 0\}$, $|v| = 1$, issuing from a is called an intermediate half-tangent of E at a if there exists a sequence $z_n \in E$, $z_n \neq a$, such that $z_n \to a$ and $|z_n - a|^{-1}(z_n - a) \to v$.

We have [13, 3.6, p. 266]

Lemma 2.2. (Kolmogorov and Vercenko.) *Let E be a subset of the plane with the property that for each $a \in E$ there exists some half-line issuing from a which is not an intermediate half-tangent of E at a. Then E has σ-finite length.*

Lemmas 2.1 and 2.2 will give now the result we really need.

Lemma 2.3. *Let $f = \log|z| * h$, $h \in H^1(\mathbb{R}^2)$. Let E be the set of points at which f vanishes and f has a non-zero ordinary differential. Then E has σ-finite length.*

PROOF. We are going to check that E satisfies the hypothesis of 2.2. Let $a \in E$ and assume without loss of generality that $a = 0$, $\partial f(0)/\partial x = 0$ and $\partial f(0)/\partial y = \rho > 0$. The definition of differential at 0 gives ($z = x + iy$)

$$f(z) = \rho y + \epsilon(z)|z|,$$

with $\epsilon(z) \to 0$ as $z \to 0$. If $\delta > 0$ is small enough, then $|\epsilon(z)| < \rho/2$ if $|z| < \delta$. Since f vanishes on E, we get

$$|y| < |x|, \quad \text{for} \quad 0 \neq z \in E \quad \text{and} \quad |z| < \delta.$$

Therefore any of the two imaginary semi-axis is not an intermediate half-tangent of E at 0. □

END OF THE PROOF OF THEOREM 1. The argument at the beginning of the proof of Theorem 1.6 shows that

$$VMO(X) = CMO(\mathbb{C})/K(X),$$

where $CMO(\mathbb{C})$ is the closure in $BMO(\mathbb{C})$ of $C_0^\infty(\mathbb{C})$ and

$$K(X) = \{f \in CMO(\mathbb{C}): f = 0 \text{ a.e. on } X\}.$$

Since $CMO(\mathbb{C})^* = H^1(\mathbb{C})$, it is clear that

$$VMO(X)^* = \{h \in H^1(\mathbb{C}): h = 0 \text{ a.e. on } X^c\}.$$

Let $h \in VMO(X)^*$ and assume that h annihillates $H(X)$. We must show that h annihillates $h(X)$. Set $f = \log|z| * h$, so that f vanishes on X^c. Since f is continuous, it vanishes also on $\overset{\circ}{X}{}^c$. Because of Theorem 1.6 we are just left with the task of showing that $\langle h, (1/z) * \mu\rangle = 0$, whenever μ is a complex Borel measure supported on $(\overset{\circ}{X})^c$ satisfying the growth condition

$$|\mu|(\Delta(z, r)) \leqslant \epsilon(r)r, \; z \in \mathbb{C}, \; r > 0, \tag{16}$$

for some $\epsilon(r) \to 0$ as $r \to 0$.

We have

$$2\pi\left\langle h, \frac{1}{z} * \mu \right\rangle = 4\left\langle \bar{\partial}\partial f, \frac{1}{z} * \mu \right\rangle = -4\pi\langle \partial f, \mu\rangle = -4\pi \int \partial f \, d\mu.$$

Let D be the set of points at which f has an ordinary differential. Now, μ vanishes on sets of σ-finite length because of the growth condition (16) and

the complement of D has zero length because of Lemma 2.1. Hence

$$\int \partial f \, d\mu = \int_{D \cap (\mathring{X})^c} \partial f \, d\mu.$$

On the other hand f vanishes on $(\mathring{X})^c$, and thus

$$\int \partial f \, d\mu = \int_E \partial f \, d\mu,$$

where E is the set of points at which f vanishes and f has a non-zero ordinary differential. But E has σ-finite length because of Lemma 2.3 and consequently $\mu(E) = 0$. Then $\int \partial f \, d\mu = 0$, as desired. □

3. Spectral Synthesis

We begin by observing that the mapping

$$I_2(\mathbb{R}^2) \to H^1(\mathbb{R}^2)$$
$$f \to \Delta f$$

is an onto isomorphism. Hence the dual of $I_2(\mathbb{R}^2)$ is $BMO(\mathbb{R}^2)$ and the action of a $b \in BMO(\mathbb{R}^2)$ on a $f \in I_2(\mathbb{R}^2)$ is given by

$$b(f) = \langle b, \Delta f \rangle$$

where $\langle \, , \rangle$ is the $BMO - H^1$ duality.

Let $b \in BMO(\mathbb{R}^2)$ and assume

$$b(\varphi) = 0, \qquad \varphi \in C_0^\infty(U), \qquad U = E^c. \tag{17}$$

We must show that $b(f) = 0$, f being the function in the statement of Theorem 2. Obviously (17) is equivalent to the harmonicity of b on U. The strategy of the proof consists in using Vitushkin's constructive scheme to approximate b in the weak* topology of $BMO(\mathbb{R}^2)$ by simpler functions β for which we know that $\beta(f)$ is either zero or small.

It will be first shown that, without loss of generality, one can suppose b to be harmonic outside a compact subset of E. Let $\varphi \in C_0^\infty(\mathbb{C})$, $\operatorname{supp}(\varphi) \subset \{|z| < 2\}$ and $\varphi = 1$ on $|z| < 1$. Set $\varphi_n(z) = \varphi(z/n)$, $n = 1, 2, \ldots$ $b_n = V_{\varphi_n} b$ and $R_n = b - b_n$. Then b_n is harmonic on $U \cup \{|z| > 2n\}$ and $\|b_n\|_* \leqslant C\|b\|_*$. Passing to a subsequence we get $R_{n_j} \to R$ weak* in $BMO(\mathbb{R}^2)$ as $j \to \infty$, and it turns out that R is harmonic on $\mathbb{R}^2$ because R_n is harmonic on $\{|z| < n\}$. Hence R is constant, or, in other words, $b_n \to b$ weak* in $BMO(\mathbb{R}^2)$. Replacing b by b_n it becomes clear that we can assume that b is harmonic on the complement V of a compact subset K of E.

Fix $\delta > 0$ and let $(\Delta_j, \varphi_j, b_j)$ a δ-Vitushkin scheme for the approximation of b. Then $b_j = (1/2\pi)\log|z| * \varphi_j \Delta b$ and $b = \sum_j b_j$, where the sum is over those j such that Δ_j intersects K. We proceed, as in the proof of Theorem 1.6, to localize again each b_j, this time using Lemma 1.4 with $h(t) = t$. We get functions

$$P_j(z) = \sum_k c_k \log|z - w_k| + 2\,\mathrm{Re}\left(\frac{1}{z} * \mu_j\right),$$

where $c_k \in \mathbb{C}$, $w_k \in K$ and μ_j is a complex Borel measure with compact support contained in K, satisfying $|\mu_j|(\Delta(z, r)) \leqslant Cr$, $z \in \mathbb{C}$, $r > 0$. Moreover

$$b_j = P_j + O(|z|^{-2}) \quad \text{as} \quad z \to \infty,$$

and

$$\|P_j\|_* \leqslant C\|b_j\|_* \leqslant C\|b\|_*.$$

We need a triple zero at ∞, so we set

$$b_j - P_j = \frac{c_j}{(z - z_j)^2} + \frac{\bar{c}_j}{(\bar{z} - \bar{z}_j)^2} + O(|z|^{-3}) \quad \text{as} \quad z \to \infty,$$

z_j being the center of Δ_j. We recall that

$$|c_j| \leqslant C\delta^2\|b_j - P_j\|_* \leqslant C\delta^2\|b\|_*.$$

Let $\psi_j \in C_0^\infty(\Delta_j)$, $\int \psi_j = 1$ and $\|\psi_j\|_\infty \leqslant C\delta^{-2}$. Define $B_j = 2\,\mathrm{Re}\,B(c_j\psi_j)$ (B is the Beurling transform).

Thus

$$b_j = P_j + B_j + O(|z|^{-3}) \quad \text{as} \quad z \to \infty$$

and

$$\|B_j\|_* \leqslant C\|b\|_*.$$

Set

$$\beta_\delta = \sum_j P_j + B_j \quad \text{and} \quad D_\delta = b - \beta_\delta.$$

We claim that $\beta_\delta \to b$ as $\delta \to 0$, uniformly on compact subsets of V. To prove the claim consider a compact $H \subset V$ and choose δ so that $3\delta < d$, d being the distance from H to K. We have the decay estimate

$$|b_j(z) - P_j(z) - B_j(z)| \leqslant C\delta^3|z - z_j|^{-3}\|b\|_*, \qquad |z - z_j| > 2\delta,$$

and therefore, for $z \in H$ and any positive integer N,

$$\sum_j |b_j(z) - P_j(z) - B_j(z)| \leqslant \Big(\sum_{|z-z_j| \leq N\delta} + \sum_{|z-z_j| > N\delta} \Big) C\delta^3 |z - z_j|^{-3} \|b\|_*$$

$$\leqslant CN^2\delta^3 d^{-3}\|b\|_* + C\|b\|_* + C\|b\|_* \sum_{n=N}^{\infty} n^{-2}.$$

One concludes the proof of the claim by taking N big enough and then δ small enough.

Because of Lemma 1.1 $\|D_\delta\|_* \leqslant C\|b\|_*$. Take now a sequence $\delta_n \to 0$ and set $\beta_n = \beta_{\delta_n}$, $D_n = D_{\delta_n}$. Passing to a subsequence we can assume $\beta_n \to \beta$ and $D_n \to D$ weak* in $BMO(\mathbb{R}^2)$. Since $D_n \to 0$ uniformly on compact subsets of V, $D = 0$ on V. Hence $\langle D, \Delta f \rangle = 0$ because D lives on V^c and Δf on $U \subset V$. Thus

$$b(f) = \langle \beta, \Delta f \rangle.$$

Now, $\langle \beta, \Delta f \rangle = 0$ will follow from

$$\langle P_j, \Delta f \rangle = 0 \tag{18}$$

and

$$\left\langle \sum_j B_j, \Delta f \right\rangle \to 0 \quad \text{as} \quad \delta \to 0. \tag{19}$$

To show (18) write

$$\langle \log|z - w_k|, \Delta f \rangle = 2\pi f(w_k),$$

which vanishes because $w_k \in F$, and

$$-\left\langle \frac{1}{z} * \mu_j, \Delta f \right\rangle = 4\pi \langle \mu_j, \partial f \rangle = 4 \int \partial f \, d\mu_j,$$

which is zero because μ_j lives on F, μ_j vanishes on sets of zero length and ∇f vanishes M^1 – a.e. on F.

We turn now to the proof of (19). Set $\chi_j = c_j \psi_j$. Then

$$-\left\langle \sum_j B(\chi_j), \Delta f \right\rangle = 4\left\langle \sum_j B(\bar\partial \chi_j), \partial f \right\rangle$$

$$= 4\left\langle \sum_j C(\bar\partial \chi_j), \partial^2 f \right\rangle$$

$$= 4\pi \left\langle \sum_j \chi_j, \partial^2 f \right\rangle,$$

and so

$$\left| \left\langle \sum_j B_j, \Delta f \right\rangle \right| \leqslant C \int_{\cup_j \Delta_j} |\partial^2 f| \to C \int_E |\partial^2 f| \quad \text{as} \quad \delta \to 0.$$

But the last integral vanishes because $\partial^2 f = 0$ a.e. on E, which follows from the fact that $\partial f = 0$ a.e. on E.

Acknowledgements. The idea of the three points Lemma 1.3 was suggested four years ago by J. Garnett to the second author in connection with a similar problem. We are grateful also to J. Dorronsoro for supplying a proof of Lemma 2.1 and to J. J. Carmona who brought to our attention the result of Kolmogorov and Vercenko used in Section 2.

References

[1] Adams, D. R. A note on the Choguet integrals with respect to Hausdorff capacity, Proceedings of the US-Swedish Conference of Function Spaces and Interpolation Theory (1986), Lund, Sweden.
[2] Boivin, A. and Verdera, J. Approximation analytique sur les fermés du plan. Preprint (1988).
[3] Deny, J. Systèmes totaux de fonctions harmoniques, *Ann. Inst. Fourier,* **1**(1949), 103-113.
[4] García-Cuerva, J. and Rubio de Francia, J. L. *Weighted norm inequalities and related topics*, North-Holland Mathematical Studies, Amsterdam, 1985.
[5] Harvey, R. and Polking, J. Removable singularities for solutions of linear partial differential equations, *Acta Math.* **125**(1970), 39-56.
[6] Hedberg, L. I. Two approximation problems in function spaces, *Ark. Mat.* **16**(1978), 51-81.
[7] Hedberg, L. I. and Wolff, T. H. Thin sets in non-linear potential theory. *Ann. Inst. Fourier,* **33**(1983), 161-187.
[8] Holden, P. J. Extension theorems for functions of vanishing mean oscillation, PhD Thesis (1987), California Inst. of Technology.
[9] Kaufman, R. Hausdorff measure, BMO, and analytic functions, *Pacific J. Math.* **102**(1982), 369-371.
[10] Keldysh, M. V. On the solubility and stability of the Dirichlet problem, *Amer. Math. Soc. Transl.* **(2) 51**(1966), 1-73.
[11] Orobitg, J. Spectral synthesis in spaces of functions with derivatives in H^1. Preprint (1988).
[12] Poornima, S. On the Sobolev spaces $W^{k,1}(\mathbb{R}^n)$, *Lecture Notes Math.*, 992, Springer-Verlag, 1983.
[13] Saks, S. *Theory of the Integral*, Dover Publ., New York, 1964.
[14] Spanne, S. Some function spaces defined using the mean oscillation over cubes, *Ann. Scuola Norm. Sup. Pisa* **19**(1965), 593-608.
[15] Stein, E. M. The differentiability of functions in $\mathbb{R}^n$, *Annals of Math.* **113**(1981), 383-385.
[16] Stein, E. M. *Singular integrals and differentiability properties of functions*, Princeton Univ. Press, 1970.
[17] Verdera, J. BMO rational approximation and one dimensional Hausdorff content, *Trans. Amer. Math. Soc.* **297**(1986), 283-304.

[18] Verdera, J. C^m approximation by solutions of elliptic equations, and Calderón-Zygmund operators, *Duke Math. J.* **55**(1987), 157-187.

Joan Mateu
Departament de Matemàtiques
Universitat Politècnica de Catalunya
Diagonal 647, 08028 Barcelona.
SPAIN

Joan Verdera
Departament de Matemàtiques
Universitat Autónoma de Barcelona
08193 Bellaterra, Barcelona
SPAIN

Both authors were partially supported by the grant CAYCIT PR 85-0374.

REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.º 2, 1988

Almost-Everywhere Convergence of Fourier Integrals for Functions in Sobolev Spaces, and an L^2-Localisation Principle

Anthony Carbery and Fernando Soria

In memory of José Luis Rubio de Francia

1. Introduction

For a function f defined on $\mathbb{R}^n$, we associate its Fourier transform

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i x \cdot \xi} \, dx,$$

and the partial sum operators

$$S_R f(x) = \int_{|\xi| \leq R} \hat{f}(\xi) e^{2\pi i x \cdot \xi} \, d\xi \tag{1.1}$$

whenever these make sense. One of the most classical problems in Fourier analysis is that of pointwise convergence of $S_R f(x)$ to $f(x)$ as $R \to \infty$. That is, for which spaces of functions f do we have $S_R f(x) \to f(x)$ at a given point, almost-everywhere, everywhere, or uniformly? One may consider also the corresponding lacunary problem, *i.e.* when does $S_{2^k} f(x) \to f(x)$ as $k \to \infty$, $k \in \mathbb{N}$?

The natural spaces to consider for uniform convergence are (spaces of) continuous functions, and the natural spaces to consider for almost-everywhere convergence are the L^p-spaces, $1 \leqslant p < \infty$. Even for $n = 1$ and $p = 2$, the

almost-everywhere question has only relatively recently been answered by Carleson [4], and historically progress has been made by considering easier problems. One such easier problem arises when we replace the rough Fourier multiplier $\chi_{|\xi| \leq R}$ in (1.1) by a smoother cousin, for example $(1 - |\xi|^2/R^2)_+^\alpha$ with $\alpha > 0$. This family of multipliers, the so-called Bochner-Riesz multipliers, gives rise to a summability method which has been much studied in the last 20 years-see for example [1, 2, 5, 7, 8, 9, 10, 12]. As an alternative way to make things simpler, one may, instead of insisting upon results, say, for L^p spaces, look for results for (large) subspaces of L^p which consist of functions having some smoothness. Of course, the classical tests for pointwise convergence of Fourier series are stated in terms of smoothness hypotheses on the functions, but this point of view seems not yet to have been treated from a modern perspective. The aim of this article is to lay down a framework within which this question may be considered. This turns out to be possible due to the progress made in the theory of Bochner-Riesz multipliers, although there is not apparently any formal link between smoothing the mutliplier in (1.1) and multiplying $\hat{f}(\xi)$ by a power of $|\xi|$.

To place our results in context, let us recall some classical results on the unit circle $\boldsymbol{T}$, where S_N denotes the N^{th} partial sum operator.

(i) For some $f \in L^1(\boldsymbol{T})$, $S_{2^k} f \to f$ a.e. fails (Kolmogorov's example).
(ii) If, however, $f \in L^1(\boldsymbol{T})$ and f vanishes in some open set Ω, then $S_N f \to 0$ on Ω (Riemann's localisation principle).
(iii) If $f \in L^p(\boldsymbol{T})$, $1 < p \leqslant \infty$, then $S_{2^k} f \to f$ a.e. (Littlewood-Paley).
(iv) For some continuous f, $S_N f \to f$ uniformly on $\boldsymbol{T}$ fails (du Bois Reymond).
(v) If, however, $\omega(t)$ denotes the L^∞ modulus of continuity of f, and

$$\int_0^1 \omega(t) \frac{dt}{t} < \infty,$$

then $S_N f \to f$ uniformly on $\boldsymbol{T}$ (Dini's test).
(vi) If the Fourier coefficients of f satisfy $\sum |a_k|^2 \log |k| < \infty$, then $S_N f \to f$ a.e. (Kolmogorov-Seliverstov-Plessner).

(For all these results, see Zygmund [19].) Only in the 1960's was the situation on $\boldsymbol{T}$ finalised, the result *par excellence* being due to Carleson [4] and Hunt [13]:

(vii) If $f \in L^p(\boldsymbol{T})$, $1 < p \leqslant \infty$, then $S_N f \to f$ a.e.

In this paper we concentrate on the higher-dimensional case, and so from now on we work in $\mathbb{R}^n$, with $n \geqslant 2$.

Let $1 \leqslant p < \infty$, and let us consider the almost-everywhere convergence problem for $L^p(\mathbb{R}^n)$. If almost-everywhere convergence is to hold for $f \in L^p$, then $S_1 f$ should at the very least be defined as a tempered distribution. Hence by duality, $S_1 : \mathcal{S} \to L^{p'}(\mathbb{R}^n)$, ($\mathcal{S}$ denoting the Schwartz class), and, by choosing $\hat{f} \equiv 1$ on $\{|\xi| \leqslant 1\}$ we see that unless

$$K_1(x) = \int_{|\xi| \leq 1} e^{2\pi i x \cdot \xi} d\xi$$

belongs to $L^{p'}(\mathbb{R}^n)$, we cannot make sense of the problem. But of course

$$K_1(x) = \frac{J_{n/2}(2\pi|x|)}{|x|^{n/2}} \sim \frac{e^{2\pi i|x|}}{|x|^{(n+1)/2}} + \text{lower order terms at } \infty,$$

(see [17]). Thus we see that it is reasonable to consider only the almost-everywhere convergence problem for $f \in L^p(\mathbb{R}^n)$, $p < 2n/(n-1)$. (We owe this observation to José L. Rubio de Francia.) On the other hand, if $p < 2$, any sort of almost-everywhere convergence of $S_R f(x)$ for $f \in L^p$ would imply that the corresponding maximal operator would be weak type p, due to the abstract principle of Calderón and Stein-see for example [11]. This would then contradict C. Fefferman's disc multiplier theorem [9] or one of its variants.

Of course it will not be an easy task to obtain a positive a.e. convergence result for $S_R f$, $f \in L^p$, $2 \leqslant p < 2n/(n-1)$, since such a result would logically imply Carleson's Theorem. However, we can obtain a form of the localisation principle in this range.

Theorem 1. *If $n \geqslant 2$, $f \in L^p(\mathbb{R}^n)$ with $2 \leqslant p < 2n/(n-1)$ and $f \equiv 0$ on an open set Ω, then $S_R f \to 0$ a.e. on Ω.*

An earlier variant of Theorem 1 where f was assumed to have compact support was obtainded by Sjölin, [14]. It may be worth mentioning that the analogue of the Littlewood-Paley theorem for $\mathbb{R}^n$ has been obtained by Rubio de Francia, Vega and the first author in [2]:

Theorem A. *If $n \geqslant 2$ and $f \in L^p(\mathbb{R}^n)$ with $2 \leqslant p < 2n/(n-1)$, then $S_{2^k} f \to f$ a.e.*

The proof of Theorem 1 uses many of the ideas from [2].

We now turn to see how matters may be improved if we add a little smoothness to our L^p-spaces. For $1 \leqslant p < \infty$ and $\alpha > 0$, we let $L^p_\alpha(\mathbb{R}^n)$ be the Sobolev space of Bessel potentials of order α, *i.e.* the closure of $\{f \in \mathcal{S}: [(1 + |\xi|^2)^{\alpha/2} \hat{f}(\xi)]^\vee \in L^p\}$ under the obvious norm. (When $\alpha \in N$ and $1 < p < \infty$ this coincides with the usual Sobolev space of functions whose derivatives of order up to α are in L^p. For this and all the elementary facts about Sobolev spaces, see [16, Chap. 5].) If we now consider the a.e. convergence problem for L^p_α, we observe that inserting a power of $(1 + |\xi|^2)^{1/2}$ in (1.1) does not affect the nature of the singularity of $\chi_{|\xi| \leq R}$; and so the argument above shows us that *no matter how large α is*, we should only consider the problem for $p < 2n/(n-1)$.

Theorem 2. *If $n \geqslant 2$, $f \in L^p_\alpha(\mathbb{R}^n)$ with $\alpha > 0$ and $2 \leqslant p < 2n/(n-1)$, then $S_R f \to f$ a.e.*

In fact the proof of Theorem 2 will show us that a stronger statement holds.

Theorem 3. *Let $n \geqslant 2$, $2 \leqslant p < 2n/(n-1)$ and $f \in L^p(\mathbb{R}^n)$. If for all $\phi \in C_c^\infty(\mathbb{R}^n)$ the function $g = \phi f$ satisfies*

$$\int |\hat{g}(\xi)|^2 (1 + \log^+ |\xi|)^2 \, d\xi < \infty,$$

then $S_R f \to f$ a.e.

Theorem 3 is a weak form of a Kolmogorov-Seliverstov-Plessner-type theorem, and is formally similar to the Rademacher-Menšov theorem (see [19]) for general orthogonal series. It would therefore be interesting to be able to replace the factor $(1 + \log^+ |\xi|)^2$ by $(1 + \log^+ |\xi|)$ (or something smaller!) in the statement of this theorem.

To conclude the discussion of the case $p \geqslant 2$, let us state a result which might be regarded as an analogue for $\mathbb{R}^n$ of Dini's test:

Theorem 4. *Let $n \geqslant 2$. If $f \in L^{2n/(n-1),1}_{(n-1)/2}(\mathbb{R}^n)$, then $S_R f \to f$ uniformly.*

(Here, $L^{p,1}_\alpha$ is the space of functions whose derivatives up to order α are in the Lorentz space $L^{p,1}$; $L^{2n/(n-1),1}_{(n-1)/2}(\mathbb{R}^n)$ embeds in $C_0(\mathbb{R}^n)$.)

We now consider what may happen when $1 \leqslant p < 2$. In this situation, any positive result for L^p_α, $p > 1$ implies a corresponding positive result for Bochner-Riesz means of order $\beta > \alpha$ on L^p. For example, a «lacunary» localisation principle for L^p_α would imply that $(1 - |\xi|^2)^\beta_+$ is a Fourier multiplier of L^p for each $\beta > \alpha$, and a «full» localisation principle would imply that

$$\left\| \sup_{1 \leqslant t \leqslant 2} |[(1 - |\xi|^2/t^2)^\beta_+ \hat{f}(\xi)]^\vee| \right\|_p \leqslant c \|f\|_p \tag{1.2}$$

for each $\beta > \alpha$. Now a necessary condition for $(1 - |\xi|^2)^\beta_+$ to be a Fourier multiplier of L^p is

$$\frac{1}{p} < \frac{n + 1 + 2\beta}{2n}$$

(take $\hat{f} \in \mathcal{S}$, $\hat{f} \equiv 1$ on $|\xi| \leqslant 1$) and this has been shown to be sufficient (for $p \leqslant 2$) only when

$$\beta > \frac{n-1}{2(n+1)}$$

(see [10], [18]) or when $n = 2$ and $\beta > 0$ (see [5], [12], [10], [7]). The state of affairs for the maximal Bochner-Riesz operator is even worse: (1.2) has only been shown to hold in the cases corresponding to interpolation with the «trivial» L^1 and L^2 estimates: that is, (1.2) is only known to hold for $1 < p \leqslant 2$ and $\alpha > 0$ when

$$\frac{1}{p} < \frac{\alpha}{n-1} + \frac{1}{2}.$$

For the a.e. convergence problem, we have

Theorem 5. *Let $n \geqslant 2$. If $f \in L^1_{(n-1)/2}(\mathbb{R}^n)$, then $S_R f \to f$ a.e.*

Notice that Theorem 5 is false when $n = 1$, by Kolmogorov's example. Moreover the corresponding statement for Bochner-Riesz means of order $(n-1)/2$ is false, since Stein has shown (see [17]) that localisation fails for L^1 in this case. However Theorem 5 is very easy to prove, and is closely related to Theorem 4. A immediate consequence of Theorems 2 and 5 is

Theorem 6. *Let $n \geqslant 2$. If $f \in L^p_\alpha(\mathbb{R}^n)$ with*

$$1 < p \leqslant 2, \ \alpha > 0 \quad \text{and} \quad \frac{1}{p} < \frac{\alpha}{n-1} + \frac{1}{2},$$

then $S_R f \to f$ a.e.

Finally, for lacunary convergence, we can improve Theorem 6 in some special cases.

Theorem 7. *If $n = 2$ and $\alpha > 0$, or if $n \geqslant 3$ and $\alpha > \dfrac{n-1}{2(n+1)}$, then $S_{2^k} f \to f$ a.e. for $f \in L^p_\alpha$, $\dfrac{1}{2} \leqslant \dfrac{1}{p} < \dfrac{n+1+2\alpha}{2n}$.*

In view of the recent work of M. Christ on weak type estimates for Bochner-Riesz operators (see [6]) it would be interesting to establish Theorem 7 for $1/p = (n+1+2\alpha)/2n$. We hope to return to this matter in a forthcoming article.

The paper is organized as follows. Theorem 1 is proved in Section 2, Theorems 3 and 4 in Section 3, and Theorems 5 and 7 in Section 4. The proof of Theorem 7 requires some $L^p - L^q$ estimates for the operator S_1, and more generally one can consider $L^p - L^q$ estimates for Bochner-Riesz means of order $\alpha \leqslant 0$. These have previously been studied by Sogge, [15], and we make

some improvements on his estimates in Section 5. (Although for $\alpha < 0$ these estimates have no relation with the rest of the paper, it seems convenient to include them here.)

We were privileged to discuss all stages of this work with José Luis Rubio de Francia. As always, he freely shared his ideas with us, and we feel this work owes much to his insight. The first author enjoyed the generous and gracious hospitality of José Luis and his family during several delightful visits to Madrid. We will miss him a great deal, and so it is with great sadness that we dedicate this work to his memory.

2. A Localisation Principle

We begin with our localisation principle.

Theorem 2.1. *Suppose* $2 \leqslant p < 2n/(n-1)$ *and suppose* $f \in L^p(\mathbb{R}^n)$. *Then, almost everywhere off* $\operatorname{supp} f$ *we have* $S_R f(x) \to 0$.

PROOF. By translation and dilation invariance, we may assume without loss of generality that f is supported in $\{|x| \geqslant 3\}$. We must show that for every open ball B contained in $\{|x| < 3\}$ we have $S_R f(x) \to 0$ a.e. in B. We shall do this for $B = \{x: |x| < 1\}$; for any other ball $\{|x| < r\}$ with $r < 3$ the proof is similar.

It suffices, as usual, to obtain an appropriate estimate for the maximal operator

$$S_* f(x) = \sup_{1 < R < \infty} |S_R f(x)|;$$

in fact, for the L^2 case we shall prove

$$\int_{|x| \leq 1} |S_* f(x)|^2 \, dx \leqslant c \int_{|x| \geq 3} |f|^2. \tag{2.1}$$

In order to make use of the support restrictions in (2.1) we split up both f and

$$K_R(x) = \int_{|\xi| \leq R} e^{2\pi i x \cdot \xi} \, d\xi = R^n K_1(Rx)$$

into dyadic pieces. Let ϕ be a smooth radial function with $\chi_{|x| \leq 1} \leqslant \phi \leqslant \chi_{|x| \leq 2}$. Then

$$1 = \phi(x) + \sum_{j=1}^{\infty} [\phi(x/2^j) - \phi(x/2^{j-1})] = \phi(x) + \sum_{j=1}^{\infty} \psi(x/2^j) = \phi(x) + \sum_{j=1}^{\infty} \psi_j(x),$$

where $\psi(x) = \phi(x) - \phi(2x)$. Let $K_R^j = K_R \psi_j$, and $f_j = f\psi_j$. Then

$$S_R f = K_R * f = \sum_{j,k \geq 1} K_R^j * f_k + (K_R \phi) * f,$$

since $\phi f \equiv 0$. Now if $|j-k| \geq 3$, then $\operatorname{supp}(K_R^j * f_k) \cap \{|x| \leq 1\} = \varnothing$, and $\operatorname{supp}(K_R \phi * f) \cap \{|x| \leq 1\} = \varnothing$ since $\operatorname{supp} f \subseteq \{|x| \geq 3\}$. Thus to obtain (2.1) it suffices to show

$$\int \sup_{R>1} \left| \sum_{j \geq 1} K_R^j * f_{j+s} \right|^2 dx \leq c \int |f|^2 \tag{2.2}$$

for $|s| < 3$; we shall deal only with the diagonal case $s = 0$, the other cases being exactly similar. The inequality (2.2) is now the special case $0 < \alpha = \gamma < 1/2$ of the following theorem:

Theorem 2.2. *Let $\{g_j\}_{j=1}^{\infty}$ be an arbitrary sequence of test functions. Suppose $0 \leq \alpha < 1/2$ and $\gamma > 0$. Then*

$$\int_{\mathbb{R}^n} \sup_{R>1} \left| \sum_{j \geq 1} K_R^j * g_j \right|^2 \frac{dx}{|x|^{2\alpha}} \leq c_{\alpha,\gamma} \sum_{j \geq 1} 2^{2j\gamma} \int_{\mathbb{R}^n} |g_j|^2 \frac{dx}{|x|^{2\alpha}}.$$

Remark 1. If we set $g_j = f_j = f\psi_j$ with $\operatorname{supp} f \subseteq \{|x| \geq 3\}$, Theorem 2.2 shows that

$$\int_{|x| \leq 1} |S_* f(x)|^2 dx \leq c_{\alpha,\gamma} \int_{|x| \geq 3} |f|^2 \frac{dx}{|x|^{\eta}}$$

with $\eta = 2(\alpha - \gamma) < 1$ but arbitrary. In particular since $L^p(|x| \geq 3, dx) \subseteq L^2(dx/|x|^{\eta})$ for $n(1 - 2/p) < \eta$, we see that Theorem 2.1 follows from Theorem 2.2. It is easy to see that Theorem 2.2 with parameters $\alpha = \gamma = 0$ is false, and thus even to obtain the L^2 case of Theorem 2.1 one is forced to consider weighted L^2 inequalities.

Remark 2. The results of this paper were announced in [3], where the statement of Theorem 2 is slightly incorrect. The correct version is the present Theorem 2.2.

The proof of Theorem 2.2 requires several lemmas. Let H_α be the (homogeneous) Sobolev space $\left\{ f \in \mathcal{S}'(\mathbb{R}^n) : \int |\hat{f}(\xi)|^2 |\xi|^{2\alpha} d\xi < \infty \right\}$.

Lemma 2.3. *If $0 \leq \alpha < 1/2$, $t > 0$ and $\delta < 2t$, then*

$$\int_{||\xi| - t| \leq \delta} |h(\xi)|^2 d\xi \leq c_\alpha \delta^{2\alpha} \|h\|_{H_\alpha}^2$$

with c_α independent of t and δ.

PROOF. If $\delta < t/2$, the inequality is merely a rescaled version of Lemma 3 in [2]; if $t/2 \leqslant \delta < 2t$,

$$\int_{||\xi| - t| \leq \delta} |h|^2 \leqslant \int_{|\xi| \leq 3t} |h|^2 \leqslant c_\alpha t^{2\alpha} \|h\|_{H_\alpha}^2,$$

as may be seen by chopping up the ball $\{|\xi| \leqslant 3t\}$ into dyadic annuli.

Let $m_t^j(\xi) = (K_t^j)^\wedge(\xi)$ for $j \geqslant 1$ and $t > 0$.

Lemma 2.4. *Let $\beta \in \mathbb{N}$, $\lambda \geqslant 0$, and $t > 0$. Then*

$$\left| \left(\frac{d}{dt} \right)^\beta m_t^j(\xi) \right| \leqslant c_{\lambda,\beta} 2^{j\beta} / (1 + 2^j ||\xi| - t|)^\lambda$$

with $c_{\lambda,\beta}$ independent of t and j.

PROOF. By a change of variables it suffices to consider the case $t = 1$ and we shall consider only the case $\beta = 0$, the cases $\beta \geqslant 1$ being handled similarly.

First suppose $|\xi| < 1$. Then, since $\psi(0) = 0$,

$$\begin{aligned} |m_1^j(\xi)| &= |\chi_B * (\hat{\psi})_{2^j}(\xi)| \\ &= \left| \int (\hat{\psi})_{2^j}(\eta)[\chi_B(\xi - \eta) - 1]\, d\eta \right| \\ &= \left| \int_{|\xi - \eta| \geq 1} (\hat{\psi})_{2^j}(\eta)\, d\eta \right| \\ &\leqslant \int_{|\eta| \leq ||\xi| - 1|} |(\hat{\psi})_{2^j}(\eta)|\, d\eta \\ &\leqslant \int_{|u| \geq 2^j ||\xi| - 1|} |\hat{\psi}(u)|\, du \\ &\leqslant c_\lambda / (1 + 2^j ||\xi| - 1|)^\lambda \quad \text{for any} \quad \lambda \geqslant 0. \end{aligned}$$

Similarly, if $|\xi| > 1$,

$$\begin{aligned} |m_1^j(\xi)| &= \left| \int_{|\xi - \eta| \leq 1} (\hat{\psi})_{2^j}(\eta)\, d\eta \right| \\ &\leqslant \int_{|\eta| \geq ||\xi| - 1|} |(\hat{\psi})_{2^j}(\eta)|\, d\eta \\ &\leqslant c_\lambda / (1 + 2^j ||\xi| - 1|)^\lambda \quad \text{for any} \quad \lambda \geqslant 0. \end{aligned}$$

Fix a smooth function $\gamma\colon \mathbb{R} \to \mathbb{R}$ such that $\gamma(t) \equiv 0$ for $t \leqslant 1/2$ and $\gamma(t) \equiv 1$ for $t \geqslant 1$. Let $L_t^j = \gamma(t) K_t^j$ for $t \in \mathbb{R}$, and let D^β be the fractional differentiation operator of order β on $\mathbb{R}$, that is $(D^\beta h)^\wedge(v) = |v|^\beta \hat{h}(v)$.

Lemma 2.5. *For $\beta \geqslant 0$ and $0 \leqslant \alpha < 1/2$ we have*

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}} |D^\beta L_t^j * f|^2\, dt\, \frac{dx}{|x|^{2\alpha}} \leqslant c 2^{-j} 2^{2j\beta} \int |f|^2 \frac{dx}{|x|^{2\alpha}} \tag{2.4}$$

(where D^β acts with respect to the t-variable).

PROOF. The Lemma says the operator $f \mapsto L_t^j * f$ is bounded into certain (Hilbert space valued) L^2-Sobolev spaces on $\mathbb{R}$. If we can establish this boundedness (with the correct constants) for integral values of the derivative parameter β, then we shall have the lemma in general by interpolation. By Plancherel's theorem on $\mathbb{R}$ (in the Hilbert space valued setting), it suffices to replace D^β in (2.4) by the usual derivative $(d/dt)^\beta$ when $\beta \in \mathbb{N}$. Thus it suffices to prove

$$\int_{\mathbb{R}^n} \int_0^\infty \left| \left(\frac{d}{dt}\right)^\beta K_t^j * f \right|^2 dt \frac{dx}{|x|^{2\alpha}} \leqslant c 2^{-j} 2^{2j\beta} \int |f|^2 \frac{dx}{|x|^{2\alpha}} \tag{2.5}$$

for $0 \leqslant \alpha < 1/2$, $\beta \in \mathbb{N}$, and we shall do so only in the case $\beta = 0$, the other cases being exactly the same.

We first notice that the operator $f \to K_t^j * f$ is a local one (in x-space); that is, if f is supported in a cube Q of side 2^j, then $K_t^j * f$ is supported in a larger cube Q^* of comparable volume. Thus it suffices to prove (2.5) for f supported in such a cube, and we consider two cases, that when $\operatorname{dist}(Q^*, 0) \geqslant 2^j$ and when $\operatorname{dist}(Q^*, 0) < 2^j$.

When $\operatorname{dist}(Q^*, 0) \geqslant 2^j$, the factors $|x|^{-2\alpha}$ in the integrals in (2.5) are essentially constants, so we may ignore them. Hence it is enough to show that

$$\int_{\mathbb{R}^n} \int_0^\infty |K_t^j * f|^2 \, dt \, dx \leqslant c 2^{-j} \int |f|^2 \, dx,$$

or, by Plancherel's theorem, that

$$\sup_{\xi \in \mathbb{R}^n} \int_0^\infty |m_t^j(\xi)|^2 \, dt \leqslant c 2^{-j}.$$

But this is an easy consequence of Lemma 2.4.

When $\operatorname{dist}(Q^*, 0) \leqslant 2^j$, it suffices to show

$$\int_{\mathbb{R}^n} \int_0^\infty |K_t^j * f|^2 \, dt \frac{dx}{|x|^{2\alpha}} \leqslant c 2^{-2j\alpha} 2^{-j} \int_Q |f|^2 \, dx \tag{2.6}$$

since

$$\int_Q |f|^2 \, dx \leqslant c 2^{2j\alpha} \int_Q |f|^2 \frac{dx}{|x|^{2\alpha}}.$$

By duality, (2.6) is equivalent to

$$\int_{\mathbb{R}^n} \left| \int_0^\infty K_t^j * f_t \, dt \right|^2 dx \leqslant c 2^{-2j\alpha} 2^{-j} \int \int_0^\infty |f_t(x)|^2 \, dt \, |x|^{2\alpha} \, dx,$$

or, by Plancherel's theorem, to

$$\int_{\mathbb{R}^n} \left| \int_0^\infty m_t^j(\xi) g_t(\xi) \, dt \right|^2 d\xi \leqslant c 2^{-2j\alpha} 2^{-j} \int_0^\infty \| g_t \|_{H_\alpha}^2 \, dt. \tag{2.7}$$

Now, by Lemma 2.4 we may write, for any $\lambda \geqslant 0$,

$$|m_t^j(\xi)| \leqslant c_\lambda \sum_{k=0}^{\infty} 2^{-k\lambda} \chi_{\{\xi: 2^j||\xi|-t| \leq 2^k\}},$$

and so we consider

$$\begin{aligned}
&\int_{\mathbb{R}^n} \left| \int_0^\infty g_t(\xi) \chi_{\{t: ||\xi|-t| \leq 2^{k-j}\}}\, dt \right|^2 d\xi \\
&\qquad \leqslant \int_0^\infty \int_{\mathbb{R}^n} |g_t(\xi)|^2 \chi_{\{\xi: ||\xi|-t| \leq 2^{k-j}\}}\, d\xi\, dt \sup_\xi |\{t>0: ||\xi|-t| \leqslant 2^{k-j}\}| \\
&\qquad \leqslant c 2^{k-j} \int_0^\infty \int_{\{\xi: ||\xi|-t| \leq 2^{k-j}\}} |g_t(\xi)|^2\, d\xi\, dt \\
&\qquad \leqslant c 2^{k-j} 2^{2(k-j)\alpha} \int_0^\infty \|g_t\|_{H_\alpha}^2\, dt
\end{aligned}$$

if $0 \leqslant \alpha < 1/2$, by Lemma 2.3. Thus, if $\lambda > \alpha + 1/2$,

$$\begin{aligned}
\left(\int_{\mathbb{R}^n} \left| \int_0^\infty m_t^j(\xi) g_t(\xi)\, dt \right|^2 d\xi \right)^{1/2} &\leqslant c_\lambda \sum_{k=0}^{\infty} 2^{-k\lambda} 2^{(k-j)/2} 2^{(k-j)\alpha} \left(\int_0^\infty \|g_t\|_{H_\alpha}^2\, dt \right)^{1/2} \\
&\leqslant c 2^{-j\alpha} 2^{-j/2} \left(\int_0^\infty \|g_t\|_{H_\alpha}^2\, dt \right)^{1/2},
\end{aligned}$$

which is (2.7).

PROOF OF THEOREM 2.2. By the Sobolev embedding theorem on $\mathbb{R}$,

$$\begin{aligned}
\sup_{R \geq 1} \left| \sum_{j \geq 1} K_R^j * g_j \right|^2 &\leqslant \sup_{t \in \mathbb{R}} \left| \sum_{j \geq 1} L_t^j * g_j \right|^2 \\
&\leqslant c_\beta \int_{\mathbb{R}} \left| \sum_{j \geq 1} D^\beta L_t^j * g_j \right|^2 dt
\end{aligned}$$

provided $\beta > 1/2$, and so

$$\begin{aligned}
\left(\int \sup_{R \geq 1} \left| \sum_{j \geq 1} K_R^j * g_j \right|^2 \frac{dx}{|x|^{2\alpha}} \right)^{1/2} &\leqslant \left(c_\beta \int_{\mathbb{R}^n} \int_{\mathbb{R}} \left| \sum_{j \geq 1} D^\beta L_t^j * g_j \right|^2 dt \frac{dx}{|x|^{2\alpha}} \right)^{1/2} \\
&\leqslant c_\beta \sum_{j \geq 1} \left(\int_{\mathbb{R}^n} \int_{\mathbb{R}} |D^\beta L_t^j * g_j|^2\, dt \frac{dx}{|x|^{2\alpha}} \right)^{1/2} \\
&\leqslant c_\beta \sum_{j \geq 1} 2^{-j/2} 2^{j\beta} \left(\int |g_j|^2 \frac{dx}{|x|^{2\alpha}} \right)^{1/2}
\end{aligned}$$

(by Lemma 2.5 if $0 \leqslant \alpha < 1/2$)

$$\leqslant c_\beta \left(\sum_{j \geq 1} 2^{-j} 2^{2j\beta} 2^{-2j\gamma} \right)^{1/2} \left(\sum_{j \geq 1} 2^{2j\gamma} \int |g_j|^2 \frac{dx}{|x|^{2\alpha}} \right)^{1/2}$$

$$\leqslant c_{\alpha,\gamma}\left(\sum_{j\geqslant 1} 2^{2j\gamma}\int_{\mathbb{R}^n} |g_j|^2 \frac{dx}{|x|^{2\alpha}}\right)^{1/2}$$

provided $\gamma > \beta - 1/2$ for some $\beta > 1/2$; that is, provided $\gamma > 0$. This concludes the proof of Theorems 2.2 and 2.1.

3. L^p_α Estimates, $p \geqslant 2$

In this section we first use the result of Section 2 to prove Theorem 3 of the introduction, and then establish an analogue of Dini's test for $\mathbb{R}^n$.

Theorem 3.1. *Let $n \geqslant 2$, $2 \leqslant p < 2n/(n-1)$ and $f \in L^p(\mathbb{R}^n)$. If for each $\phi \in C_c^\infty$ the function $g = \phi f$ satisfies*

$$\int |\hat{g}(\xi)|^2 (1 + \log^+ |\xi|)^2 \, d\xi < \infty,$$

then $S_R f \to f$ a.e.

PROOF. We shall show $S_R f \to f$ a.e. on $\{|x| < 1\}$; by the localisation principle (Theorem 2.1) it suffices to assume that f is supported in $\{|x| \leqslant 2\}$, and satisfies

$$\int |\hat{f}(\xi)|^2 (1 + \log^+ |\xi|)^2 \, d\xi < \infty.$$

By the Hardy-Littlewood maximal theorem we may assume that we are dealing with modified partial sum operators $\tilde{S}_R$ with multipliers $\tilde{m}_R = \chi_{|\xi| \leqslant R} - \tau(|\xi|/R)$ with $\tau \in C^\infty$ and $\chi_{|\xi| \leqslant 1/2} \leqslant \tau \leqslant \chi_{|\xi| \leqslant 3/4}$.

We shall again consider a maximal operator

$$\tilde{S}_* f(x) = \sup_{R>1} |\tilde{S}_R f(x)| \quad \text{for} \quad |x| \leqslant 1,$$

and for the moment fix an $N \in \mathbb{N}$ and consider $\tilde{S}_R f(x)$ with $R \sim 2^N$. Now

$$\begin{aligned}\tilde{S}_R f(x) &= \tilde{K}_R * f(x)\\ &= (\tilde{K}_R \phi) * f(x) \quad \text{if} \quad |x| \leqslant 1,\end{aligned}$$

where $\phi \in \mathbb{S}$, and $\phi \equiv 1$ on $\{|x| \leqslant 30\}$. Thus

$$\begin{aligned}\tilde{S}_R f(x) &= [\tilde{K}_1 \phi(\cdot/2^N)]_R * f(x)\\ &= \left[\sum_{j=1}^{N} \tilde{K}_1(\cdot)\psi(\cdot/2^j) + \tilde{K}_1(\cdot)\phi(\cdot)\right]_R * f(x)\end{aligned}$$

where $\psi(x) = \phi(x) - \phi(2x)$. Hence,

$$\sup_{R \sim 2^N} |\tilde{S}_R f(x)| \leqslant CMf(x) + \sup_{R \sim 2^N} \left| \sum_{j=1}^{N} (\tilde{K}^j)_R * f(x) \right|$$

$$\leqslant CMf(x) + \sum_{j=1}^{N} \sup_{R \sim 2^N} |(\tilde{K}^j)_R * f(x)|,$$

(where M denotes the Hardy-Littlewood maximal function and

$$\tilde{K}^j(\cdot) = \tilde{K}_1(\cdot)\psi(\cdot/2^j)).$$

Lemma 3.2. *There is a c independent of j so that*

$$\int \sup_{1<t<2} |(\tilde{K}^j)_t * h|^2 \, dx \leqslant c \int |h|^2.$$

Accepting the truth of this lemma for the moment, we see that

$$\left(\int_{|x| \leq 1} \sup_{R \sim 2^N} |\tilde{S}_R f|^2 \right)^{1/2} \leqslant C\|f\|_2 + \sum_{j=1}^{N} \left(\int_{|x| \leq 1} \sup_{t \sim 2^N} |(\tilde{K}^j)_t * f|^2 \, dx \right)^{1/2}$$

$$\leqslant C(1+N)\|f\|_2.$$

Observe that this estimate is true, by the localisation Principle, for any function in L^2. This fact will be used later. We are indebted to Pascal Auscher, M.ª Jesús Carro and Guido Weiss, from Washington University, for the observation leading to this estimate. Our original arguments for the proof of the theorem were incorrect.

Now, if we set $\hat{g}(\xi) = (1 + \log^+ |\xi|)\hat{f}(\xi)$, $g \in L^2$, and we have

$$[(\tilde{K})_R * f]\hat{}(\xi) = (\tilde{K})\hat{}(\xi/R)\hat{f}(\xi)$$

$$= (\tilde{K}_1)\hat{}(\xi/R)\hat{g}(\xi)(1 + \log^+ |\xi|)^{-1}$$

$$= \frac{1}{N} (\tilde{K}_1)\hat{}(\xi/R)\widehat{P_N g}(\xi)$$

where

$$\widehat{P_N g}(\xi) = N(1 + \log^+ |\xi|)^{-1}\sigma(\xi/2^N)\hat{g}(\xi),$$

with σ a suitable smooth bump function of compact support away from 0 around the unit annulus. We can do this because the modified Fourier multiplier $(\tilde{K}_1)\hat{}$ has support in an annulus not containing the origin. Hence,

$$\int_{|x|\le 1} \sup_{R\ge 1} |\tilde{S}_R f|^2 \le \sum_{N=1}^{\infty} \frac{1}{N^2} \int_{|x|\le 1} \sup_{R\sim 2^N} |\tilde{S}_R p_N g|^2$$

$$\le C \sum_{N=1}^{\infty} \| p_N g\|_2^2$$

$$\le C\| g\|_2^2,$$

which finishes the proof of the theorem.

PROOF OF LEMMA 3.2. This lemma is entirely standard. We have

$$\sup_{1<t<2} |(\tilde{K}^j)_t * h(x)|^2 \le 2 \int_1^2 |(\tilde{K}^j)_t * h(x)| \left| \left(\frac{d}{dt}\right)(\tilde{K}^j)_t * h(x) \right| dt + O.K.$$

$$\le 2\left(\int_1^2 |(\tilde{K}^j)_t * h|^2 \frac{dt}{t} \right)^{1/2} \left(\int_1^2 \left| t\left(\frac{d}{dt}\right)(\tilde{K}^j)_t * h \right|^2 \frac{dt}{t} \right)^{1/2}$$

and so

$$\int \sup_{1<t<2} |(\tilde{K}^j)_t * h(x)|^2 \, dx$$

$$\le c\left(\int_1^2 \int |(\tilde{K}^j)_t * h|^2 \, dx \frac{dt}{t} \right)^{1/2} \left(\int_1^2 \int \left| t \frac{d}{dt} (\tilde{K}^j)_t * h \right|^2 dx \frac{dt}{t} \right)^{1/2}$$

$$\le c\| h\|_2^2$$

provided we can show that

$$\left(\sup_{\xi} \int_1^2 |(\tilde{K}^j)\hat{}(\xi/t)|^2 \frac{dt}{t} \right) \left(\sup_{\xi} \int_1^2 \left| t \frac{d}{dt} (\tilde{K}^j)\hat{}(\xi/t) \right|^2 \frac{dt}{t} \right) \le c \tag{3.1}$$

independently of j. By a calculation similar to that of Lemma 2.4 we have

$$\left| \left(\frac{d}{dt} \right)^{\beta}_{t=1} (\tilde{K}^j)\hat{}(\xi/t) \right| \le c_{\lambda,\beta} 2^{j\beta}/(1 + 2^j||\xi| - 1|)^{\lambda},$$

so the first term in the product in (3.1) is dominated by 2^{-j}, while the second is dominated by 2^j. Thus (3.1) holds and so does Lemma 3.2.

If we assume significantly more smoothness of our function f, we can get uniform convergence of $S_R f$ to f.

Proposition 3.3. *Let $n \ge 2$. If $f \in L^{2n/(n-1),1}_{(n-1)/2}(\mathbb{R}^n)$, then $S_R f$ converges to f uniformly on $\mathbb{R}^n$.*

PROOF. Suppose $g \in L^2_\alpha(\mathbb{R}^n)$ with $\alpha > n/2$; then $S_R g \to g$ uniformly since

$$|[S_R g - S_T g](x)| = \left| \int_{R \le |\xi| \le T} \hat{g}(\xi) e^{2\pi i x \cdot \xi} \, d\xi \right|$$

$$\le \left(\int_{R \le |\xi| \le T} |\hat{g}(\xi)|^2 (1 + |\xi|^2)^\alpha \, d\xi \right)^{1/2} \left(\int \frac{d\xi}{(1 + |\xi|^2)^\alpha} \right)^{1/2}$$

$$\to 0 \quad \text{as} \quad R, T \to \infty.$$

Note that L^2_α is a dense subclass of $L^{2n/(n-1),1}_{(n-1)/2}$ for $\alpha > n/2$, so if $f \in L^{2n/(n-1),1}_{(n-1)/2}$ and $\epsilon > 0$ we may pick a $g \in L^2_\alpha$, $\alpha > n/2$, with $\|f - g\|_{L^{2n/(n-1),1}_{(n-1)/2}} < \epsilon$. Hence

$$\|S_R f - f\|_\infty \le \|S_R(f - g)\|_\infty + \|S_R g - g\|_\infty + \|g - f\|_\infty$$

$$\le \|S_R\|_{L^{2n/(n-1),1}_{(n-1)/2}, L^\infty} \|f - g\|_{L^{2n/(n-1),1}_{(n-1)/2}} + c\epsilon$$

(if R is sufficiently large, since $\|g - f\|_\infty \le c\|g - f\|_{L^{2n/(n-1),1}_{(n-1)/2}}$). It thus suffices to obtain the norm estimate

$$\|S_R f\|_\infty \le c\|f\|_{L^{2n/(n-1),1}_{(n-1)/2}}. \tag{3.2}$$

We may again assume we are working with the modified operators $\tilde{S}_R$ (since $\|(S_R - \tilde{S}_R)f\|_\infty \le c\|f\|_\infty$) and let $h \in L^{2n/(n-1),1}$ be such that

$$\hat{f}(\xi)(1 + |\xi|^2)^{(n-1)/4} = \hat{h}.$$

Then

$$(\tilde{S}_R f)\hat{}(\xi) = \frac{\tilde{m}(\xi/R)}{(1 + |\xi|^2)^{(n-1)/4}} \hat{h}(\xi)$$

$$= R^{-(n-1)/2} \tilde{m}(\xi/R) \widehat{\rho^R}(\xi) \hat{h}(\xi)$$

with $\widehat{\rho^R}$ a smooth bump function adapted to the annulus $|\xi| \sim R$. So

$$\|\tilde{S}_R f\|_\infty = \|R^{-(n-1)/2} \tilde{K}_R * \rho^R * h\|_\infty$$

$$\le R^{-(n-1)/2} \|\tilde{K}_R\|_{L^{2n/(n+1),\infty}} \|\rho^R * h\|_{L^{2n/(n-1),1}}$$

$$\le CR^{-(n-1)/2} \|\tilde{K}_R\|_{L^{2n/(n+1),\infty}} \|f\|_{L^{2n/(n-1),1}_{(n-1)/2}}.$$

Now as is well-known,

$$K_1(x) \sim \frac{e^{2\pi i|x|}}{|x|^{(n+1)/2}} + \text{lower order terms},$$

so $\|\tilde{K}_R\|_{L^{2n/(n+1),\infty}} = O(R^{(n-1)/2})$.

4. L^p_α Estimates, $p < 2$

As noted in the introduction, the only general theorem available when $p < 2$ is Theorem 6, which is an immediate consequence of the results of the previous section and this next result (which is itself closely allied to Proposition 3.3).

Proposition 4.1. *Let $f \in L^1_{(n-1)/2}(\mathbb{R}^n)$ for $n \geqslant 2$. Then $S_R f \to f$ a.e.*

PROOF. It is enough to see that the modified maximal operator $\tilde{S}_* f$ is finite almost everywhere for $f \in L^1_{(n-1)/2}$. As in the previous section,

$$\tilde{S}_R f(x) = R^{-(n-1)/2} \tilde{K}_R * \rho^R * h,$$

with $\|h\|_1 = \|f\|_{L^1_{(n-1)/2}}$, and, since $L^{2n/(n+1),\infty}$ is a Banach space when $n \geqslant 2$, we see that

$$\|\tilde{S}_* f\|_{L^{2n/(n+1),\infty}} \leqslant c \left\| \sup_{R>1} R^{-(n-1)/2} |\tilde{K}_R * \rho^R| \right\|_{L^{2n/(n+1),\infty}} \|f\|_{L^1_{(n-1)/2}}.$$

A moment's thought will convince the reader that $|\tilde{K}_R * \rho^R(x)| \leqslant c|\tilde{K}_R(x)|$ (since ρ^R is an averaging operator on a scale $1/R$), and, from the form of K_1, we see that indeed

$$\sup_{R>1} R^{-(n-1)/2} |\tilde{K}_R| \in L^{2n/(n+1),\infty}(\mathbb{R}^n).$$

If we instead consider the lacunary problem, we can say a little more.

Proposition 4.2. *If $n = 2$ and $\alpha > 0$, or if $n \geqslant 3$ and $\alpha > (n-1)/2(n+1)$, then $S_{2^k} f \to f$ a.e. for $f \in L^p_\alpha(\mathbb{R}^n)$, if $1/2 \leqslant 1/p < (n+1+2\alpha)/2n$.*

PROOF. As in previous proofs, it suffices to see that

$$\left\| \sup_{k \geqslant 1} |\tilde{S}_{2^k} f| \right\|_q \leqslant c \|f\|_{L^p_\alpha},$$

for some $1/q < (n+1)/2n$ since

$$\left\| \sup_k |(S_{2^k} - \tilde{S}_{2^k}) f| \right\|_q \leqslant \|Mf\|_q \leqslant c\|f\|_q \leqslant c\|f\|_{L^p_\alpha}$$

if $1/q \geqslant 1/p - \alpha/n$. To do this, it suffices to show that

$$\left\| \sup_{k \geqslant 1} |2^{-k\alpha} \tilde{K}_{2^k} * \rho^{2^k} * h| \right\|_q \leqslant c\|h\|_p$$

(since $(\tilde{K})^\wedge$ has compact support in an annulus). Now the case $\alpha = 0$ of

Theorem 5.2 of Section 5 below states that

$$\|K_1 * f\|_q \leqslant c\|f\|_p$$

for $1/q < (n+1)/2n$ if $(n+3)/2(n+1) < 1/p \leqslant 1$, or if $n = 2$ and $3/4 < 1/p \leqslant 1$. By Young's inequality, the same holds for $\tilde{K}_1$ in place of K_1, and this now scales to

$$\|\tilde{K}_t * f\|_q \leqslant ct^{n(1/p - 1/q)}\|f\|_p.$$

Finally,

$$\left\| \sup_{k \geqslant 1} |2^{-k\alpha}\tilde{K}_{2^k} * \rho_{2^k} * h| \right\|_q \leqslant \sum_{k \geqslant 1} 2^{-k\alpha} 2^{kn(1/p - 1/q)} \|\rho_{2^k} * h\|_p$$

$$\leqslant c\|h\|_p$$

if $1/q < (n+1)/2n$, $\alpha > n(1/p - 1/q)$ and $(n+3)/2(n+1) < 1/p \leqslant 1$ ($3/4 < 1/p \leqslant 1$ when $n = 2$). Thus we obtain a.e. convergence of $S_{2^k} f$ for $1/2 \leqslant 1/p < (n+1+2\alpha)/2n$ and $\alpha > (n-1)/2(n+1)$ or, when $n = 2$, $\alpha > 0$.

5. Bochner-Riesz Means of Negative Order

For $-(n+1)/2 \leqslant \alpha \leqslant 0$, we define for $f \in \mathcal{S}$

$$(T^\alpha f)^\wedge(\xi) = \Gamma(\alpha+1)^{-1}(1 - |\xi|^2)_+^\alpha \hat{f}(\xi),$$

so that T^α is convolution with $\pi^{-\alpha}|x|^{-n/2-\alpha} J_{n/2+\alpha}(2\pi|x|)$. In this section we examine for which $1 \leqslant p \leqslant q \leqslant \infty$, α and n we have

$$\|T^\alpha f\|_{L^q(\mathbb{R}^n)} \leqslant c\|f\|_{L^p(\mathbb{R}^n)}. \tag{5.1}$$

The case $\alpha = 0$ of this problem occurred in the preceding section, and the problem has also been previously studied by Börjeson and Sogge. By taking $\hat{f} \in \mathcal{S}$, $\hat{f} \equiv 1$ on $|\xi| \leqslant 1$, we see that necessary conditions for (5.1) to hold are $1/q < (n+1+2\alpha)/2n$, and, by duality, $1/p > (n-1-2\alpha)/2n$ (except when $\alpha = -(n+1)/2$, when we trivially have $\|T^{-(n+1)/2} f\|_\infty \leqslant c\|f\|_1$). Also, we can not expect, in general, (5.1) to hold unless $1/q \leqslant 1/p + 2\alpha/(n+1)$ since the optimal $L^p - L^{p'}$ theorem for T^{-1} is for $(n+3)/2(n+1) \leqslant 1/p \leqslant 1$-anything stronger would give a false $L^p - L^2$ restriction theorem for the Fourier transform.

For $-(n+1)/2 \leqslant \alpha \leqslant 0$, let

$$A_\alpha = \{(1/p, 1/q) \in [0,1]^2 : 1/q < (n+1+2\alpha)/2n,$$
$$1/p > (n-1-2\alpha)/2n,$$
$$1/q < 1/p + 2\alpha/(n+1)\}.$$

Sogge's results may be summarised as follows

Theorem [15].

(*a*) *If* $n = 2$ *and* $(1/p, 1/q) \in A_\alpha$, *then*

$$\|T^\alpha f\|_{L^q(\mathbb{R}^2)} \leqslant c\|f\|_{L^p(\mathbb{R}^2)}.$$

(*b*) *If* $n \geqslant 3$, $(1/p, 1/q) \in A_\alpha$ *and if* $\alpha \leqslant -1/2$, *then*

$$\|T^\alpha f\|_q \leqslant c\|f\|_p.$$

Thus a natural conjecture is that in all dimensions $(1/p, 1/q) \in A_\alpha$ is sufficient to imply (5.1) if $-(n+1)/2 < \alpha \leqslant 0$. It would also seem reasonable to conjecture that T^α is weak type $(p, 2n/(n+1+2\alpha))$ for $[(n+1+2\alpha)/2n] - [2\alpha/(n+1)] < 1/p \leqslant 1$ (true for $p = 1$!) and perhaps even weak type $(([n+1+2\alpha)/2n] - [2\alpha/(n+1)])^{-1}, 2n/(n+1+2\alpha))$ if $\alpha > 0$. This last conjecture would imply that T^α is strong type (p, q) on the line $1/q = 1/p + 2\alpha/(n+1)$, which is true for $q = p'$ (but false by Fefferman's theorem [8] if $\alpha = 0$ and $p \neq 2$).

Part (*b*) of Sogge's theorem is implied by the following lemma

Lemma [15]. *Let* $K^j(\cdot) = K_1(\cdot)\psi(\cdot/2^j)$ (*with* ψ *smooth of compact support in an annulus*). *Then*

$$\|K^j * f\|_q \leqslant c2^{-j/2}2^{jn(1/q-1/2)}\|f\|_p$$

if $1/q = (n+1)(n-1)^{-1}(1-1/p)$ *and* $0 \leqslant 1/q \leqslant 1/2$.

Here we notice that a stronger statement is true.

Lemma 5.1. *If* $(n+1)(n-1)^{-1}(1-1/p) \leqslant 1/q \leqslant 1/p$ *and* $(n+3)/2(n+1) \leqslant 1/p \leqslant 1$, *then*

$$\|K^j * f\|_q \leqslant c2^{-j/2}2^{jn(1/q-1/2)}\|f\|_p.$$

This lemma together with the observation that $T^\alpha: L^p \to L^{p'}$ if $1/p = 1/2 - \alpha/(n+1)$ immediately gives us:

Theorem 5.2. *If* $n \geqslant 3$, $(1/p, 1/q) \in A_\alpha$, *and* $(1/p, 1/q)$ *is strictly below both the lines joining*

$(2^{-1} - \alpha(n+1)^{-1}, 2^{-1} + \alpha(n+1)^{-1})$ *to* $((n+3)/2(n+1), (n+1+2\alpha)/2n)$ and $(2^{-1} - \alpha(n+1)^{-1}, 2^{-1} + \alpha(n+1)^{-1})$ *to* $((n-1-2\alpha)/2n, (n-1)/2(n+1))$, *then* $\|T^\alpha f\|_q \leqslant c\|f\|_p$.

Curiously, Lemma 5.1 is a consequence of the (trivial) observation that $\|K^j * f\|_{p'} \leqslant c2^{-j(n+1)(1/p-1/2)}\|f\|_p$ although it is not obtainable by interpolation from this estimate. This came to light in conversations with Luis Vega.

PROOF OF LEMMA 5.1. Only the case $q \leqslant 2$ is of interest, since we can interpolate this with the trivial $L^1 - L^\infty$ estimate to get the case $q > 2$. The case $q = 2$ follows immediately from the Stein-Tomas restriction theorem (and hence from $\|K^j * f\|_{p'} \leqslant c2^{-j(n+1)(1/p-1/2)}\|f\|_p$)-see Stein's argument in [10]. If now $p \leqslant q < 2$, we may assume that f is supported in a ball of radius 2^j; now

$$\|K^j * f\|_q \leqslant c\|K^j * f\|_2 2^{jn(1/q-1/2)}$$

by Hölder's inequality, and the case $q = 2$ gives the desired result.

For Section 4, we need part (a) of Sogge's theorem and Theorem 5.2, both in the case $\alpha = 0$,

$$1/q = (n+1)/2n - \epsilon,\ 1 \leqslant p \leqslant 4/3\ (n = 2),\ 1 \leqslant p \leqslant 2(n+1)/(n+3)\ (n \geqslant 3).$$

Theorem 5.2 has been obtained independently by A. Seeger, (personal communication).

We should like to thank C. Kenig and C. Meaney for pointing out to us, after this work was completed, that the analogue of Theorem 2 on the n-torus has been obtained by B. I. Golubov (Math USSR Sbornik **25**(1974), 177-197).

However, his methods do not apply to give our Theorem 3.

References

[1] Carbery, A. The boundedness of the maximal Bochner-Riesz operator on $L^4(\mathbb{R}^2)$, *Duke Math. J.* **2**(1983), 409-416.

[2] Carbery, A., Rubio de Francia J. L. and Vega, L. Almost everywhere summability of Fourier integrals, *J. Lond. Math. Soc.* (2) **38**(1988), 513-524.

[3] Carbery, A. and Soria, F. Almost everywere convergence of Fourier integrals in $\mathbb{R}^n$. *Extracta Mathematicae* 3, **2**(1988), 55-57.

[4] Carleson, L. On convergence and growth of partial sums of Fourier series, *Acta Math.* **116** (1966), 135-157.

[5] Carleson, L. and Sjölin, P. Oscillatory integrals and a multiplier problem for the disc, *Studia Math.* **44** (1972), 287-299.

[6] Christ, M. Weak type endpoint bounds for Bochner-Riesz multipliers, *Rev. Mat. Iberoamericana,* 3, **1**(1987), 25-31.

[7] Córdoba, A. The Kakeya maximal function and the spherical summation multipliers, *Amer. J. Math.* **99** (1977), 1-22.

[8] Fefferman, C. Inequalities for strongly singular convolution operators, *Acta Math.* **124**(1970), 9-36.

[9] Fefferman, C. The multiplier problem for the ball, *Ann. of Math.* **94**(1971), 330-336.
[10] —. A note on spherical summation multipliers, *Israel J. Math.* **15** (1973), 44-52.
[11] García-Cuerva, J. and Rubio de Francia, J. L. *Weighted Norm Inequalites and Related Topics*, North-Holland, 1985.
[12] Hörmander, L. Oscillatory integrals and multipliers on FL^p, *Ark. Mat.* **11**(1971), 1-11.
[13] Hunt, R. On the convergence of Fourier Series, *Proc. Conf. Southern Illinois Univ.*, Carbondale (1967).
[14] Sjölin, P. Two Theorems on Fourier Integrals and Fourier Series, Upsala University, Report No. 3, (1986).
[15] Sogge, C. Oscillatory integrals and spherical harmonics, *Duke. Math. J.* **53**(1986), 43-65.
[16] Stein, E. M. *Singular Integrals and Differentiability Properties of Functions.* Princeton Univ. Press, 1970.
[17] Stein, E. M. and Weiss, G. *An Introduction to Fourier Analysis on Euclidean Spaces.* Princeton Univ. Press, 1971.
[18] Tomas, P. A. A restriction theorem for the Fourier transform, *Bull. Amer. Math. Soc.* **81**(1975), 477-478.
[19] Zygmund, A. *Trigonometric Series.* Cambridge Univ. Press, 1959.

A. Carbery
Mathematics Division,
University of Sussex
Falmer, Brighton BN19QH
ENGLAND

F. Soria
Departamento de Matemáticas
Universidad Autónoma de Madrid
28049 Madrid
SPAIN

Both authors were partially supported by grants from the British Council and the Spanish Ministry of Education under the Acciones Integradas programmes, 1986-1987.

Research at MSRI supported in part by NSF Grant DMS-812079-05.

REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.º 2, 1988

Fundamental Solutions and Asymptotic Behaviour for the p-Laplacian Equation

S. Kamin and J. L. Vázquez

In memory of José Luis Rubio de Francia

Abstract

We establish the uniqueness of fundamental solutions to the p-Laplacian equation

(PLE) $$u_t = \operatorname{div}(|Du|^{p-2}Du), \qquad p > 2,$$

defined for $x \in \mathbb{R}^N$, $0 < t < T$. We derive from this result the asymptotic behaviour of nonnegative solutions with finite mass, *i.e.* such that $u(\cdot, t) \in L^1(\mathbb{R}^N)$. Our methods also apply to the porous medium equation

(PME) $$u_t = \Delta(u^m), \qquad m > 1,$$

giving new and simpler proofs of known results. We finally introduce yet another method of proving asymptotic results based on the idea of asymptotic radial symmetry. This method can be useful in dealing with more general equations.

1. Introduction

The p-Laplacian equation

(1.1) $$u_t = \operatorname{div}(|Du|^{p-2}Du), \qquad p > 1$$

admits a fundamental solution, *i.e.* a function $w \geqslant 0$ which solves (1.1) in a weak sense in $Q = \mathbb{R}^N \times (0, \infty)$ and takes on the initial data

$$u(x, 0) = M\delta(x), \qquad M > 0. \tag{1.2}$$

It is given by the following selfsimilar expression (the «Barenblatt» solution [Ba])

$$w(x, t) = t^{-k}(C - q|\xi|^{p/(p-1)})_+^{(p-1)/(p-2)} \tag{1.3}$$

where

$$\xi = xt^{-k/N}, \qquad k = \left(p - 2 + \frac{p}{N}\right)^{-1}, \qquad q = \frac{p-2}{p}\left(\frac{k}{N}\right)^{1/(p-1)}$$

and C is related to the mass M by $C = cM^{\alpha}$, with $\alpha = p(p-2)k/N(p-1)$ and $c = c(p, N)$ determined from the condition $\int w(x, t)\, dx = M$.

In this paper we prove that $w = w_M$ is the unique solution of (1.1), (1.2) which is nonnegative (and satisfies the natural growth conditions as $|x| \to \infty$). Moreover, these are the only nonnegative solutions of (1.1) which take on the initial value $u(x, 0) = 0$ for $x \neq 0$, apart from the trivial solution $u = 0$ (*i.e.* they have a positive isolated singularity at $(0, 0)$, see Theorem 1, Section 3).

In the course of our proof we estimate the rate of expansion of supports of nonnegative solutions and give necessary and sufficient conditions for the existence of a waiting time.

We then use Theorem 1 to prove that w represents the asymptotic behaviour as $t \to \infty$ of any nonnegative solution with mass $\int u_0(x)\, dx = M$. (Theorem 2, Section 4.) This result was known in one space dimension. In fact, for $N = 1$ Esteban and Vázquez [EV] establish a detailed description of the asymptotic behaviour of solutions and interfaces for a class of equations which generalizes both (1.1) and the porous medium equation

$$u_t = \Delta(u^m), \qquad m > 1, \tag{1.4}$$

under the assumptions: $u_0 \in L^1(\mathbb{R})$, $u_0 \geqslant 0$, u_0 has compact support.

Our method uses in a strong way the similarity properties attached to the power-like nonlinearity of the equation.

Corresponding results are also known for the porous medium equation for $N \geqslant 1$. Uniqueness of solutions with a bounded measure as initial datum was proved by Pierre [P]. Asymptotic behaviour is due to Friedman and Kamin [FK]. Of course, the results are well-known for the heat equation, which corresponds to the case $p = 2$ or $m = 1$ above.

Our proofs of Theorem 1 and 2 apply without any essential modification to equation (1.4), thus establishing the uniqueness of fundamental solutions (which is a part of Pierre's result) and the asymptotic behaviour of [FK] by different methods. We will give detailed references to the background

material as the proof proceeds. The corresponding self-similar solutions w are explicitly stated in Section 5.

In the last section we give yet another proof of the asymptotic behaviour both for the porous medium and p-Laplacian equations which uses completely different methods. Since the proof is analogous in both cases we have chosen to present it only for the porous medium equation. It is based on the idea of asymptotic radial symmetry.

Finally, we recall that in the stationary case

$$-\operatorname{div}(|Du|^{p-2}Du) = 0, \qquad x \neq 0 \tag{1.5}$$

the description of solutions with a singularity at the origin is due to [FV] and [KV]. In this work Kichenassamy and Veron establish a classification of solutions with an isolated singularity at the origin, generalizing classical results by Serrin. For equation

$$-\operatorname{div}(|Du|^{p-2}Du) + |u|^{q-1}u = 0 \tag{1.6}$$

a similar study is done in [FV].

2. Preliminaries

Since equation (1.1) is degenerate at the points where $Du = (u_{x_1}, \ldots, u_{x_N})$ vanishes we need to introduce a suitable concept of solution. A measurable function $u = u(x, t)$ defined in $Q_T = \mathbb{R}^N \times (0, T)$, $T > 0$, is a weak solution of (1.1) if

$$u \in C((0, T]; L^1_{\mathrm{loc}}(\mathbb{R}^N)) \cap L^1(0, T; W^{1,p-1}_{\mathrm{loc}}(\mathbb{R}^N))$$

and for every test function

$$\phi \in W^{1,\infty}(0, T; L^\infty(\mathbb{R}^N)) \cap L^\infty(0, T; W^{1,\infty}(\mathbb{R}^N))$$

having compact support, the following identity holds

$$\iint_{Q_T} (-u\phi_t + |Du|^{p-2}Du \cdot D\phi)\, dx\, dt = 0 \tag{2.1}$$

($Du \cdot D\phi$ means scalar product). Di Benedetto and Herrero [dBH] study the existence and uniqueness of weak solutions. They define the norm

$$|||f|||_r = \sup_{R \geq r} R^{-\lambda} \int_{B_R(0)} |f|\, dx, \qquad \lambda = N + \frac{p}{p-2} \tag{2.2}$$

for $f \in L^1_{\mathrm{loc}}(\mathbb{R}^N)$ and $r > 0$; $B_s(x) = \{y \in \mathbb{R}^N : |y - x| < s\}$. They prove the following results in particular.

Result 1. *For every* $u_0 \in L^1_{\mathrm{loc}}(\mathbb{R}^N)$ *with* $|||u_0|||_r < \infty$, $r > 0$, *there exists a time* $T = T(u_0)$ *and a weak solution* $u(x, t)$ *of* (1.1) *in* Q_T *which takes on the initial data* u_0: $u \in C([0, T]; L^1_{\mathrm{loc}}(\mathbb{R}^N))$ *and* $u(\cdot, t) \to u_0$ in $L^1_{\mathrm{loc}}(\mathbb{R}^N)$. *Moreover for* $0 < t \leqslant T_r(u_0)$, *where*

$$T_r(u_0) = C_0 |||u_0|||_r^{-(p-2)}, \tag{2.3}$$

we have

$$|||u(\cdot, t)|||_r \leqslant C_1 |||u_0|||_r \tag{2.4}$$

$$|u(x, t)| \leqslant C_2 t^{-k} R^{p(p-2)} |||u_0|||_r^{pk/N} \quad \text{if} \quad |x| \leqslant R, \qquad r \leqslant R \tag{2.5}$$

$$|Du(x, t)| \leqslant C_3 t^{-(N+1)k/N} R^{2/(p-2)} |||u_0|||_r^{2k/N} \quad \text{if} \quad |x| < R, \qquad r \leqslant R \tag{2.6}$$

where $C_i = C_i(N, p)$, $i = 0, \ldots, 3$. *Finally* u *and* Du *are Hölder continuous in* $\mathbb{R}^N \times [\tau, T(u_0) - \tau]$ *with Hölder constants and exponents depending upon* N, p, C_i, τ *and* $|||u_0|||_r$.

Result 2. *Two weak solutions* u, v *of* (1.1) *defined in* Q_T, *such that* $|||u(\cdot, t)|||_r$ *and* $|||v(\cdot, t)|||_r$ *are uniformly bounded for* $0 \leqslant t \leqslant T$ *and some* $r > 0$, *and such that*

$$u(\cdot, t) - v(\cdot, t) \to 0 \quad \text{in} \quad L^1_{\mathrm{loc}}(\mathbb{R}^N) \tag{2.7}$$

as $t \downarrow 0$, *are equal in* Q_T, $u \equiv v$.

Result 3. (Maximum Principle) *If under the above conditions* $u_0 \leqslant v_0$ *a.e. in* $\mathbb{R}^N$, *then* $u \leqslant v$ *in their common strip of definition* Q_T.

Result 4. *A nonnegative solution of* (1.1) *in a domain* Q_T *admits a unique initial trace* u_0 *which is a* σ*-finite Borel measure.*

Result 5. (Harnack inequality) *Let* u *be a nonnegative solution of* (1.1) *in* Q_T. *Then for every* $R > 0$; $\tau, \theta > 0$ *with* $\tau + \theta \leqslant T$.

$$R^{-N} \int_{B_R} u(x, \tau)\, dx \leqslant C_4(p, N) \left\{ \left(\frac{R^p}{\theta} \right)^{\frac{1}{p-2}} + \left(\frac{\theta}{R^p} \right)^{\frac{N}{p}} u(0, \tau + \theta)^{1 + N\frac{p-2}{p}} \right\} \tag{2.8}$$

As a matter of fact, existence is proved in [dBH] for initial data which are σ-finite Borel measures under the restriction of finite $|||\cdot|||_r$ norm (suitably defined for measures). One of the open problems mentioned in the paper is

the uniqueness of such solutions. A partial answer to this question will be provided by our Theorem 2.

Corresponding results are well known for the porous medium equation. Analogues of Results 1, 2 and 3 were proved by Bénilan, Crandall and Pierre [BCP], while Results 4 and 5 are due to Aronson and Caffarelli [AC]. As mentioned above the uniqueness of solutions with bounded measures as initial data was settled by Pierre [P]. This result was later extended to measures with a growth rate as $|x| \to \infty$ by Dahlberg and Kenig [DK].

We will need in our study some precise information about how supports propagate. It is well-known that for $p > 2$ equation (1.1) enjoys the *finite propagation property*, by which we mean that the support of a solution travels a finite distance in a finite time; in particular a compact support at time $t = 0$ stays compact for $t > 0$, cf. [DH], [DV]. We will deal from this moment on with nonnegative solutions, $u \geqslant 0$.

Let $\Omega(t) = \{x \in \mathbb{R}^N : u(x, t) > 0\}$ the positivity set of u at time t, $0 < t \leqslant T$. We know that the family $\{\Omega(t)\}_{t>0}$ is expanding in time (essentially a consequence of the Maximum Principle and the property of expanding supports for the solutions w_M). We begin by characterizing the points at which u stays zero for a certain time.

Proposition 2.1. *Given $x \in \mathbb{R}^N$ we have $u(x, t) > 0$ for every $t > 0$ if and only if*

$$B(x) = \sup_{R>0} R^{-\lambda} \int_{B_R(x)} |u_0| = \infty. \tag{2.9}$$

Moreover if $B(x) < \infty$ then $u(x, t) = 0$ for

$$0 < t < C_0(p, N) B(x)^{-(p-2)}. \tag{2.10}$$

PROOF. (2.10) is a simple consequence of (2.5) and (2.3) above taking x as the origin, so that $B(x) = \lim_{r \to 0} |||u_0|||_r$ and letting $R = r \to 0$.

For the positivity when $B(x) = \infty$ we use the Harnack inequality with $x = 0$, $\tau + \theta = t$, divide both members by $R^{p/(p-2)}$ and let $R, \tau \to 0$ to conclude that $u(0, t) > 0$ for every $t > 0$. Remark that $B(0) = \infty$ is equivalent to

$$\lim_{R \to 0} R^{-\lambda} \int_{B_R(0)} u_0(x)\, dx = \infty. \quad \square$$

In case u_0 does not vanish a.e. in any neighborhood of x, and $B(x) < \infty$, then for some positive time $u(x, t) = 0$ and x will belong to the free boundary separating the regions $[u > 0]$ and $[u = 0]$ in Q_T. Therefore we have a stationary interface and

$$t^*(x) = \sup \{t \geqslant 0 : u(x, t) = 0\} \tag{2.11}$$

is called the *waiting time* at x. Proposition 2.1 gives a necessary and sufficient condition for a waiting time to occur.

On the contrary, if u_0 vanishes in a neighborhood of x, then we can estimate the time it takes for the solution to reach x. This estimate can be precised if $u_0 \in L^1(\mathbb{R}^N)$ as follows. Let

$$d(x) = \sup \{R: u_0(y) = 0 \text{ a.e. in } B_R(x)\} \tag{2.12}$$

be the distance from x to the support of u_0 and let us define for a set $S \subset \mathbb{R}^N$ the ρ-neighborhood as $N_\rho(S) = \{y \in \mathbb{R}^N: d(y, S) \leqslant \rho\}$.

Proposition 2.2. *Let $x \in \mathbb{R}^N$ with $d(x) > 0$ and let $u_0 \in L^1(\mathbb{R}^N)$, then $u(x, t) = 0$ for*

$$0 \leqslant t \leqslant C_0(p, N) d^{\lambda(p-2)} \|u_0\|_1^{-(p-2)} \tag{2.13}$$

therefore for $T > t_2 > t_1 \geqslant 0$, $\Omega(t_2)$ is contained in the ρ-neighborhood of $\Omega(t_1)$, with

$$\rho = C_5(p, N)(\|u_0\|_1^{p-2}(t_2 - t_1))^{k/N}. \tag{2.14}$$

PROOF. (2.13) follows from the observation that under our assumptions the supremum in the definition of $B(x)$ is taken for $R \geqslant d(x)$ and choosing the worst possibility. (2.14) is a consequence of (2.13) changing the origin of time to $t = t_1$. □

It will be of great importance in the next section that, but for the constant C_5, formula (2.14) is the exact rate of propagation of the explicit solution w.

The above development follows closely the similar study by Caffarelli, Vázquez and Wolanski [CVW] for the porous medium equation.

3. Uniqueness of the Fundamental Solution

Let u be a nonnegative weak solution of (1.1) defined in Q_T, $T > 0$ and such that it takes continuously the initial data

$$u(x, 0) = 0 \quad \text{for} \quad x \neq 0. \tag{3.1}$$

This solution can exhibit a *singular* behaviour as $(x, t) \to (0, 0)$. In fact, by Result 4 there exists a σ-finite Borel measure μ such that

$$\lim_{t \to 0} \int_{\mathbb{R}^N} u(x, t)\phi(x)\, dx = \int_{\mathbb{R}^N} \phi\, d\mu$$

for every ϕ continuous and compactly supported in $\mathbb{R}^N$. Because of (3.1) μ is supported in $\{0\}$, therefore $\mu = M\delta$ and we have

$$\lim_{t\to 0}\int_{B_R} u(x,t)\,dx = M \tag{3.2}$$

for every $R > 0$. We say that u is a *fundamental* solution if it is a nonnegative weak solution and satisfies (3.1), (3.2) for some $M > 0$. The functions w defined in (1.3) are fundamental solutions of (1.1). We prove the following result.

Theorem 1. *The only nonnegative solutions of* (1.1) *satisfying* (3.1) *are the fundamental solutions* w_M *and the trivial solution.*

The proof is divided into several steps. We begin by estimating the support of any such solution u.

Lemma 3.1. $u(\cdot, t)$ *has compact support for every* $t > 0$. *Moreover* $\operatorname{supp}(u(\cdot, t)) \subset B_{\rho(t)}(0)$ *with* $\rho(t) \to 0$ *as* $t \to 0$.

PROOF. Let $\hat{u}(x,t) = w_{M_1}(x, t+\tau)$ with some M_1 and $\tau > 0$. For R and θ small enough $\hat{u}(x,t) \geqslant u(x,t)$ on $\Gamma_1 = \{(x,t): |x| = R, 0 \leqslant t \leqslant \theta\}$. The inequality is true if θ is small enough by (3.1). Consider the domain

$$D_R = \{(x,t): |x| > R, 0 < t < \theta\}. \tag{3.3}$$

Since $u = 0 \leqslant \hat{u}$ on the bottom: $|x| \geqslant R$ the Maximum Principle (a variation of Result 3 above) implies that

$$u \leqslant \hat{u} \tag{3.4}$$

in D_R. In particular $u(\cdot, t)$ has compact support, since $\hat{u}$ does for $0 < t < \theta$. Since we have the finite propagation property, $u(\cdot, t)$ is compactly supported for all t.

By (3.4), for $0 < t < \theta$

$$\operatorname{supp}(u(\cdot, t)) \subset B_{r(t)}(0) \tag{3.5}$$

with $r(t) = c(t+\tau)^{k/N}$, c depends on M_1. Letting $\tau \to 0$ we conclude that the support of $u(\cdot, t)$ shrinks to $\{0\}$ as $t \to 0$. □

A simple consequence of Lemma 3.1 is the fact that $u(\cdot, t) \in L^1(\mathbb{R}^N)$. By conservation of mass and (3.2) we have

Corollary 3.2. $\int u(x,t)\,dx = M$ *for every* $M > 0$.

We may now exclude the case $M = 0$, where necessarily $u = 0$. In the sequel we assume that $M > 0$. We also know that u exists for all times $t > 0$ since it has finite mass (cf. [dBH]).

We need also L^∞ bounds. Such bounds have been proved by [Ve], [B], [HV] and are known as «smoothing effects». We need here the *sharp* version that is contained in [V2]. See [V1] for the porous medium case.

Lemma 3.3. *For every $t > 0$ and every solution u with initial data u_0 we have*

$$|u(x,t)| \leqslant w_M(0,t) = c_*(p,N)M^{pk/N}t^{-k} \tag{3.6}$$

where $M = \|u_0\|_1$ and $k = (p - 2 + p/N)^{-1}$.

We shall use later the sharp constant c_* in an essential way.

Our next step consists in comparing u with a Barenblatt solution. In order to simplify the calculation we set $M = 1$, what implies no loss of generality using a rescalling of our solution (see (3.18) below).

Lemma 3.4. *There exists $M' \geqslant 1$ such that*

$$u(x,t) \leqslant w_{M'}(x,t) \quad \text{in} \quad Q_T \tag{3.7}$$

PROOF. We first improve our control on the support of $u(x,t)$. Let $\tau > 0$ be small so that $\operatorname{supp}(u(\cdot,\tau)) \subset B_\epsilon(0)$ with $\epsilon \approx 0$. Then by (2.14) the support of $u(\cdot,t)$, $t > \tau$, is contained in a ball of radius

$$C_5(p,N)(t-\tau)^{k/N} + \epsilon$$

(recall that we are assuming that $\|u(\cdot,\tau)\|_1 = 1$). Letting $\epsilon \to 0$, we obtain $C_5 t^{k/N}$ which corresponds to a Barenblatt solution with a certain mass M_1.

For M' large enough $w_{M'}(x,t) \geqslant c_* t^{-k}$ for $\xi = xt^{-k/N} \leqslant C_5$. By (3.6) we know that $u(x,t) \leqslant c_* t^{-k}$ for every $x \in \mathbb{R}^N$. On the other hand $u(x,t) = 0$ if $\xi \geqslant C_5$, therefore $u(x,t) \leqslant w_{M'}(x,t)$. Notice that $\int u(x,t)\,dx = 1$ implies that $M' \geqslant 1$. □

We now take the minimum of these bounds

$$M = \inf\{M' : w_{M'}(x,t) \geqslant u(x,t)\}. \tag{3.8}$$

By Lemma 3.4 such M exists, $M \geqslant 1$ and $w_M \geqslant u$ in Q. If we prove that $M = 1$ the theorem is complete since $\int u = \int w$, hence $u = w$. Therefore we are left with the task of excluding the possibility $M > 1$. This we do next.

Lemma 3.5. $M = 1$.

PROOF. We assume that M is larger than 1 and arrive at a contradiction by showing that it can be decreased further.

Step 1. Take $t = 1$. We know that $u(x, 1)$ is bounded above by $w_M(x, 1)$ and also by c_*, therefore if we consider the solution U which at time $t = 1$ takes on the data

$$U(x, 1) = \min \{c_*, w_M(x, 1)\} \tag{3.9}$$

it will be clear that for $t \geqslant 1$

$$u(x, t) \leqslant U(x, t) \leqslant w_M(x, t) \tag{3.10}$$

In fact, since $U(0, 1) \leqslant c_*$ and c_* is *strictly less* than $w_M(0, 1) = c_* M^{pk/N}$, by continuity, there exist $t_1 > 1$ and $r > 0$ such that

$$U(x, t_1) < w_M(x, t_1) \quad \text{for} \quad |x| < r. \tag{3.11}$$

We claim that $U(\cdot, t_1)$ and $w_M(\cdot, t_1)$ do not touch at any point inside the support of $w_M(\cdot, t_1)$, *i.e.* for

$$|x| < a(p, N)(M^{p/N} t_1)^k = \rho_M(t_1). \tag{3.12}$$

This is so because of the Strong Maximum Principle: since for $0 < |x| < \rho$, w_M is a classical solution of (1.1) with $|Dw_M| > 0$, if they were to touch at a point (x_0, t_1), then $Du(x_0, t_1) = Dw_M(x_0, t_1) \neq 0$. By the continuity of Du, in a neighborhood of (x_0, t_1), $u - w_M$ is also a solution of a linear strictly parabolic equation with continuous coefficients, so the Strong Maximum Principle applies and forces $u \equiv w_M$, a contradiction.

The solutions U and w_M could touch at the free boundary $|x| = \rho(t_1)$. To tackle this situation we slightly change w_M by putting in a small delay and consider $w_M(x, t_1 + \tau)$ instead of $w_M(x, t_1)$ for a small $\tau > 0$. We claim that $U(\cdot, t_1)$ is strictly below $w_M(x, t_1 + \tau)$ inside the support of the latter function, $\{x: |x| \leqslant \rho_M(t_1 + \tau)\}$.

For this we need a small technical diversion.

Lemma 3.5. *$w_M(x, t) > w_M(x, t + \tau)$ in a region $|x| \leqslant c(\tau, p, N)\rho_M(t)$ and $w_M(x, t + \tau) > w_M(x, t)$ for $c(\tau, p, N)\rho_M(t) < |x| < \rho_M(t + \tau)$. Moreover as $\tau \to 0$*

$$c(\tau, p, N) \to c_\# = ((p - 2)k)^{(p-1)/p} < 1. \tag{3.13}$$

PROOF. It is reduced to knowing the sign of $\partial w_M(x, t)/\partial t$. After a typically cumbersome computation we find that

$$\frac{\partial}{\partial t} w_M(x, t) = 0 \quad \text{for} \quad |x| = \rho_M(t) c_\#,$$

being negative for smaller $|x|$ and positive otherwise. □

Now it is clear how to prove the pending claim. Take a small $\delta > 0$. If $\tau > 0$ is small enough we have

$$U(x, t_1) < w_M(x, t_1 + \tau) \tag{3.14}$$

for $|x| \leqslant c_{\#}\rho_M(t_1) + \delta$ because U is strictly separated in this compact region from $w_M(x, t_1)$ and $\tau \approx 0$. On the other hand, for $c_{\#}\rho_M(t_1) + \delta \leqslant |x| \leqslant \rho_M(t_1 + \tau)$

$$U(x, t_1) \leqslant w_M(x, t_1) < w_M(x, t_1 + \tau). \tag{3.15}$$

Now we come to the crucial part. Since $U(\cdot, t_1)$ is strictly separated from $w_M(\cdot, t_1 + \tau)$ inside their supports, which are also strictly separated, we may slightly reduce the mass of w and have still the relatioship

$$U(x, t_1) \leqslant w_{M-\epsilon}(x, t_1 + \tau), \qquad x \in \mathbb{R}^N. \tag{3.16}$$

By the Maximum Principle the same holds for $t \geqslant t_1$. Moreover, since $u \leqslant U$ for $t \geqslant 1$ we get

$$u(x, t) \leqslant w_{M-\epsilon}(x, t + \tau) \quad \text{for} \quad x \in \mathbb{R}^N, \qquad t \geqslant t_1. \tag{3.17}$$

Step 2. We apply to U the rescaling operator

$$T_\theta U = \frac{1}{\theta^k} U\left(\frac{x}{\theta^{k/N}}, \frac{t}{\theta}\right). \tag{3.18}$$

Let us call $U_\theta = T_\theta U$. U_θ is a solution of (1.1) for $t \geqslant \theta$ which takes on the initial data

$$U_\theta(x, \theta) = \min\{c_* \theta^{-k}, w_M(x, \theta)\}, \tag{3.19}$$

since w_M is invariant under T_θ. Moreover (3.16) translates into

$$U_\theta(x, t) \leqslant w_{M-\epsilon}(x, t + \theta\tau) \quad \text{for} \quad t \geqslant \theta t_1, x \in \mathbb{R}^N. \tag{3.20}$$

The initial conditions (3.19) and our estimates on u imply again that $U_\theta \geqslant u$ for $t \geqslant \theta t_1$. Hence

$$u(x, t) \leqslant w_{M-\epsilon}(x, t + \theta\tau) \quad \text{for} \quad t \geqslant \theta t_1. \tag{3.21}$$

At last we see the end. We let $\theta \to 0$ and obtain

$$u(x, t) \leqslant w_{M-\epsilon}(x, t) \quad \text{for} \quad x \in \mathbb{R}^N \quad \text{and} \quad t > 0. \tag{3.22}$$

We are done with Lemma 3.4 and with Theorem 1. □

Remark. When we apply the above method to the porous medium equation, everything is similar. Even an important step is simpler. In our proof of Lemma 3.5 we have to be careful to eliminate the possibility that U touches w at $x = 0$, since the equation degenerates ($Dw_M = DU = 0$) and the Strong Maximum Principle does not apply. This difficulty does not occur for equation (1.4).

4. Asymptotic Behaviour

We discuss in this section the asymptotic behaviour of nonnegative solutions with finite mass. We prove the following result.

Theorem 2. *Let u be a nonnegative solution of* (1.1) *with $u_0 \in L^1(\mathbb{R}^N)$. Then*

$$\lim_{t\to\infty} t^k |u(x,t) - w_M(x,t)| = 0 \tag{4.1}$$

uniformly in $x \in \mathbb{R}^N$. Here $M = \int u_0(x)\,dx$.

PROOF. (i). To begin with, we assume that u_0 has compact support. We consider the family $\{u_\lambda = T_\lambda u\}$ of rescaled solutions, where T_λ is defined by (3.18). Since T_λ is mass preserving, $\int u_{\lambda,0}(x)\,dx = M$ for every $\lambda > 0$, therefore the family $\{u_\lambda\}$ is uniformly bounded in $\mathbb{R}^N \times (\tau, \infty)$ for any $\tau > 0$ (by estimate (3.6)). Also u_λ and Du_λ are uniformly equicontinuous on compact subsets of Q, (by Result 1, Section 2).

Therefore for every sequence $\lambda_n \to 0$ there exists a subsequence (which we also denote by λ_n) such that u_λ converges uniformly to a continuous function $\bar{u}$ and $Du_\lambda \to D\bar{u}$, also uniformly. It is clear that $\bar{u}$ will be a weak solution of (1.1).

We claim that $\bar{u} = w_M$. In order to prove this we recall that by (2.14) the support of $u(\cdot, t)$ is contained in a ball of radius $R(t)$ with

$$R(t) \leqslant c\rho_M(t) + R_0, \tag{4.2}$$

where $c = c(p, N) \geqslant 1$ and R_0 is radius of a ball $B_{R_0}(0) \supset \operatorname{supp}(u_0)$. This implies that for any fixed $t > 0$ the supports of the family $\{\bar{u}_\lambda(\cdot, t)\}_{\lambda > 1}$ are uniformly bounded by $R(t)$. Therefore $u_\lambda(\cdot, t)$ converges in $L^1(\mathbb{R}^N)$ to $\bar{u}(\cdot, t)$ so that

$$\int \bar{u}(\cdot, t)\,dx = M. \tag{4.3}$$

Every solution has an initial trace, μ. Since the support of $\bar{u}(\cdot, t)$ shrinks as $t \to 0$ its initial trace is a Dirac mass. Hence $\bar{u}$ is a fundamental solution and, by Theorem 1, $\bar{u} = w_M$.

By the uniqueness of the limit the whole family converges as $\lambda \to \infty$. Now observe that (4.1) is equivalent to $u_\lambda(x, 1) \to w_M(x, 1)$ uniformly in x as $\lambda \to \infty$.

(ii) For general u_0 we consider an approximating sequence $\{u_{0n}\}$ such that u_{0n} has compact support and $0 \leqslant u_{0n} \leqslant u_{0,n+1} \leqslant u_0$.

We use the L^1-dependence of solutions on initial data (contraction property in $L^1(\mathbb{R}^N)$, cf. [Be]): if $u, \hat{u}$ are two solutions with initial data $u_0, \hat{u}_0 \in L^1(\mathbb{R}^N)$, then for every $t > 0$

$$\|u(\cdot, t) - \hat{u}(\cdot, t)\|_1 \leqslant \|u_0 - \hat{u}_0\|_1. \tag{4.1}$$

Now assume that u_0 has mass M and u_{0n} has M_n with $M_n \uparrow M$. We repeat the rescaling above both on u and u_n to obtain families $\{T_\lambda u\}$, $\{T_\lambda u_n\}$. Along a subsequence $\lambda_n \to \infty$ $T_\lambda u \to \bar{u}$ while $T_\lambda u_n \to w_{M_n}$ by step (i). But since

$$\|T_\lambda u(\cdot, t) - T_\lambda u_n(\cdot, t)\|_1 = \|u(\cdot, \lambda t) - u_n(\cdot, \lambda t)\|_1 \leqslant \|u_0 - u_{0n}\|_1$$

in the limit we will get

$$\|\bar{u} - w_{M_n}\|_1 \leqslant \|u_0 - u_{0n}\|_1$$

which as $n \to \infty$ gives $\bar{u} = w_M$. This ends the proof. □

Remark. Our proof applies literally to the porous medium equation, thus giving a new proof of the asymptotic behaviour of [FK], easily derived from the uniqueness result for fundamental solutions.

5. Another Approach to Asymptotic Behaviour

We give another method of proof of Theorem 2 which does not rely on the uniqueness of fundamental solutions. It uses instead the idea that solutions with compact support become asymptotically radially symmetric and the principle of concentration comparison introduced by Vázquez in [V2] for the porous medium equation, and valid also for the p-Laplacian equation. Since this is of interest also for the porous medium equation we give the proof for this equation.

We consider a solution $u \geqslant 0$ of the problem

$$u_t = \Delta(u^m) \quad \text{for} \quad (x, t) \in Q, \qquad m > 1 \tag{5.1}$$

$$u(x, 0) = u_0(x) \quad \text{for} \quad x \in \mathbb{R}^N \tag{5.2}$$

with $u_0 \in L^1(\mathbb{R}^N)$, $\int u_0(x)\,dx = M$. We want to prove

Theorem 3. *As* $t \to \infty$

$$t^k |u(x, t) - w_M(x, t)| \to 0$$

uniformly in $x \in \mathbb{R}^N$, *where now* $k = (m - 1 + 2/N)^{-1}$ *and* w_M *is Barenblatt's solution*

$$w_M(x,t) = t^{-k}\left(C - \frac{(m-1)k}{2mN}|\xi|^2\right)_+^{1/(m-1)} \tag{5.3}$$

with $\xi = xt^{-k/N}$ *and* C *is determined in terms of* M, m *and* N.

PROOF. (i) Using the L^1-dependence of solutions on initial data which is also true for (5.1), (5.2) and arguing as in (ii) of Theorem 2 we may assume that u_0 has compact support,

$$\operatorname{supp}(u_0) \subset B_{R_0}(0). \tag{5.4}$$

(ii) The asymptotic symmetry of the solution will be a consequence of the following Lemma, which is a variant of Proposition 2.1 of [AC] (see also [CVW], [BA]).

Lemma 5.1. *For any* $|x_1| \geqslant R_0$ *and any* $t > 0$ *we have*

$$u(x_1,t) \geqslant u(x_2,t) \tag{5.5}$$

if $|x_2| \geqslant |x_1| + 2R_0$.

PROOF. We draw the hyperplane Π which bisects the segment x_1x_2. Assume that this hyperplane leaves the support of u_0 in the same half-space Ω_1 as x. We consider the solutions

$$u_1(x,t) = u(x,t),$$
$$u_2(x,t) = u(\pi(x),t),$$

where $\pi(x)$ is the symmetric of $x \in S = \Omega_1 \times (0,\infty)$ with respect to Π. Since $u_1 = u_2$ on the lateral boundary $\Pi \times [0,\infty)$ and $u_1 \geqslant u_2 = 0$ at the bottom $\bar{\Omega}$ we conclude that $u_1 \geqslant u_2$ in S, *i.e.* $u(x,t) \geqslant u(\pi(x),t)$. Since $x_1 \in \Omega$ and $x_2 = \pi(x_1)$, $u(x_1,t) \geqslant u(x_2,t)$ for $t > 0$.

Finally, the condition for Π to leave $B_{R_0}(0)$ on the same side as x_1 is just $|x_2| \geqslant |x_1| + 2R_0$. □

Suppose now that we consider the rescaled solutions $T_\lambda u$ given again by (3.18) with our present value of k. If happens that (5.5) holds if

$$|x_2| \geqslant |x_1| + 2R_0\lambda^{k/N}, \qquad |x_1| \geqslant R_0\lambda^{-k/N}. \tag{5.6}$$

As in Section 4 we pass to the limit $\lambda_n \to \infty$ thanks to the boundedness and equicontinuity of the solutions which are well-known (cf. [FK]) and we obtain

$T_{\lambda_n} u \to \bar{u}$ uniformly on compact subsets of Q. Moreover the supports are uniformly bounded (cf. [CVW]) hence $T_{\lambda_n} u \to \bar{u}$ uniformly in $x \in \mathbb{R}^N$, $0 < \tau < t < 1/\tau$. Moreover $\bar{u}(\cdot, t) \in L^1(\mathbb{R}^N)$ with $\|\bar{u}(\cdot, t)\|_1 = M$. Passing to the limit $\lambda \to \infty$ in (5.5), (5.6) we deduce that

$$\bar{u}(x_1, t) \geqslant \bar{u}(x_2, t) \quad \text{if} \quad |x_2| \geqslant |x_1|, \qquad t > 0, \tag{5.7}$$

i.e. $\bar{u}$ is *radially symmetric* and *decreasing*.

(iii) We now recall the concept of concentration for nonnegative, radially symmetric functions in $L^1(\mathbb{R}^N)$ as defined in [V1]. We say that *f is less concentrated than g*, $f \prec g$ if for every $r > 0$.

$$\int_0^r f(\rho)\rho^{N-1}\, dp \leqslant \int_0^r g(\rho)\rho^{N-1}\, d\rho. \tag{5.8}$$

We will use the property that the relation $\prec$ is hereditary for problem (5.1), (5.2).

Lemma 5.2 [V1]. *Let u_1, u_2 be radially symmetric and nonnegative solutions of* (5.1), (5.2) *with initial data u_{01}, u_{02}. Then if $u_{01} \prec u_{02}$, we have $u_1(\cdot, t) \prec u_2(\cdot, t)$ for all $t > 0$.*

Obviously the most concentrated initial data are $u_0 = M\delta(x)$ corresponding to a Dirac mass. Therefore we have for every $t > 0$ and every u with mass M,

$$\bar{u}(\cdot, t) \prec w_M(\cdot, t). \tag{5.9}$$

It easily follows from (5.9) that $\|u(\cdot, t)\|_\infty \leqslant \|w(\cdot, t)\|_\infty$, *i.e.* the smoothing effect. (The above argument needs a justification since δ is a measure, not a function. This is however done easily by approximation; in particular replace $M\delta$ with $w_M(\cdot, \tau)$, $\tau \approx 0$ and approximate u_{0n} from below.)

(iv) We need now an estimate converse to (5.9). This is done by first remarking that arguing as in Sections 3 and 4 with the estimates for the support of [CVW] we conclude that

$$\operatorname{supp}(\bar{u}(\cdot, t)) \subset B_{R(t)}(0) \tag{5.10}$$

with

$$R(t) \leqslant c\rho_M(t) \tag{5.11}$$

where $\rho_M(t)$ is the radius of the support of w_M. Take $\tau > 0$, τ small. Then $\operatorname{supp} \bar{u}(\cdot, \tau)$ is contained in a very small ball. Since also $\bar{u}$ is radially symmetric and decreasing, it is easy to see that for some θ depending on τ

$$w_M(\cdot, \tau + \theta) \prec \bar{u}(\cdot, \tau). \tag{5.12}$$

This property will continue to be valid for $t \geqslant \tau$. Letting $\tau \to 0$, hence $\theta \to 0$, we obtain

$$w_M(\cdot, t) \prec \bar{u}(\cdot, t) \quad \text{for} \quad t > 0. \tag{5.13}$$

(5.13) and (5.9) imply that

$$\int_0^r w_M(\rho, t)\rho^{N-1}\, d\rho = \int_0^r \bar{u}(\rho, t)\rho^{N-1}\, d\rho, \tag{5.14}$$

for every $r > 0$, hence $w_M \equiv \bar{u}$.

(v) Since the limit $\bar{u}$ is unique, the whole family $T_\lambda u$ converges to $\bar{u}$. □

Remark. This proof generalizes easily to the p-Laplacian equation with $p > 1$. It can be also applied to $u_t = \Delta\phi(u)$, with assumptions on the nondecreasing function ϕ which make it resemble a power, as the ones imposed in [dPV].

References

[AC] Aronson, D. G., Caffarelli, L. A. The initial trace of a solution of the porous medium equation, *Trans. Amer. Math. Soc.* **380**(1983), 351-366.

[Ba] Barenblatt, G. I. On self-similar motions of compressible fluids in porous media, *Prikl. Mat. Mech.* **16**(1952), 679-698. (Russian).

[Ba] Bardi, M. Asymptotical spherical symmetry of the free boundary in degenerate diffusion equations, to appear.

[Be] Bénilan, Ph. Operateurs accrétifs et semi-groupes dans les espaces L_p $(1 \leqslant p \leqslant \infty)$, France-Japan Seminar, Tokyo, 1976.

[BCP] Bénilan, P. H., Grandall, M. G. and Pierre, M. Solutions of the porous medium equation in $\mathbb{R}^n$ under optimal conditions on initial values, *Indiana Univ. Math. J.* **33**(1984), 51-87.

[CVW] Caffarelli, L. A., Vázquez, J. L. and Wolanski, N. E. Lipschitz continuity of solutions and interfaces to the N-dimensional porous medium equation, *Indiana Univ. Math. J.* **36**(1987), 32-401.

[DK] Dahlberg, B. E. J. and Kenig, C. E. Nonnegative solutions of the porous medium equation, *Comm. P.D.E.*, **9**(1981), 409-437.

[DH] Díaz, J. I. and Herrero, M. A. Estimates on the support of the solutions of some nonlinear elliptic and parabolic problems, *Proc. Royal Soc. Edinburgh*, **89A**(1981), 249-258.

[DV] Díaz, J. I. and Veron, L. Local vanishing properties of solutions of elliptic and parabolic quasilinear equations, *Trans. Amer. Math. Soc.* **290**(1985), 787-814.

[dPV] De Pablo, A. and Vázquez, J. L. Regularity of solutions and interfaces of a generalized porous medium equation in $\mathbb{R}^N$, *Ann. Mat. Pura. Appl.*, to appear.

[dBH] di Benedetto, E. and Herrero, M. A. On the Cauchy problem and initial traces for a degenerate parabolic equation, *Trans. Amer. Math. Soc.*, to appear.

[EV] Esteban, J. R. and Vázquez, J. L. Homogeneous diffusion in $\mathbb{R}$ with power-like nonlinear diffusivity, *Arch. Rat. Mech. Anal.* **103**(1988), 39-80.

[FK] Friedman, A. and Kamin, S. The asymptotic behaviour of gas in an n-dimensional porous medium, *Trans. Amer. Math. Soc.* **262**(1980), 551-563.
[FV] Friedman, A. and Verón, L. Singular solutions of some quasilinear elliptic equations, *Arch. Rat. Mech. Anal.* **96**(1986), 259-287.
[KV] Kichenassamy, S. and Véron, L. Singular solutions of the p-harmonic equation, *Math. Ann.* **275**(1986), 599-615.
[P] Pierre, M. Uniqueness of the solutions of $u_t - \Delta\varphi(u) = 0$ with initial datum a measure, *Nonlinear Anal. T.M.A.* **6**(1982), 175-187.
[V1] Vázquez, J. L. Symétrisation pour $u_t = \Delta\phi(u)$ et applications, *C. R. Acad. Sci. Paris*, I, **295**(1982), 71-74.
[V2] Vázquez, J. L. Symmetrization in nonlinear parabolic problems, *Portugaliae Math.* **41**(1982), 339-345.
[Ve] Verón, L. Coercivité et propriétés régularisantes des semi-groupes non linéaires dans les espaces de Banach, *Ann. Fac. Sci. Toulouse.* **1**(1979), 171-200.

Soshana Kamin*
School of Mathematical Sciences
Raymond and Beverly Sackler
Faculty of Exact Sciences
Tel-Aviv University,
ISRAEL

Juan Luis Vázquez†
Departamento de Matemáticas
Universidad Autónoma de Madrid
28049 Madrid,
SPAIN

* Supported by grant No. 06-0682-00000 from the United States-Israel Binational Science Foundation (BSF).
† Supported by DGICYT (Spain) grant PB86-0112-C02 and EEC Contract SC1-0019-C.

AMS Classification: 35K65, 35B40.

Key Words and Phrases: Nonlinear parabolic equation, p-Laplacian, porous medium equation, fundamental solutions, asymptotic behaviour.

REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.^{os} 3 Y 4, 1988

Estimates of Kernels on Three-Dimensional *CR* Manifolds

C. L. Fefferman and J. J. Kohn

Introduction

(i) Description of results.

Let $\mathfrak{M} = b\Omega$ be the three-dimensional smooth pseudo-convex boundary of a bounded domain $\Omega \subset \mathbb{C}^2$. On $\mathfrak{M}$ there is a natural Cauchy-Riemann operator $\bar{\partial}_b$ mapping functions to sections of a suitable line bundle $\mathcal{B}^{1,0}$ of 1-forms. (See below.) The restriction to $\mathfrak{M}$ of an analytic function on Ω satisfies the tangential Cauchy-Riemann equations $\bar{\partial}_b F = 0$. We denote by $\mathcal{H}_b$ the closed subspace of $L^2(\mathfrak{M})$ annihilated by $\bar{\partial}_b$, and define the Szegö projection S_b as the orthogonal projection from $L^2(\mathfrak{M})$ onto $\mathcal{H}_b$. The Szegö projection is given in terms of an integral kernel as

$$S_b f(x) = \int_{\mathfrak{M}} K(x, y) f(y)\, dy.$$

One of our goals is to describe the singularities of the Szegö kernel $K(x, y)$.

Another goal is to solve the inhomogeneous Cauchy-Riemann equations $\bar{\partial}_b u = f$. To obtain a well-posed problem, we must take into account that both $\bar{\partial}_b$ and its adjoint have large nullspaces. Thus we pick a Hermitean metric on the bundle $\mathcal{B}^{1,0}$ and let $S_b^{1,0}$ be the orthogonal projection from L^2-sections of $\mathcal{B}^{1,0}$ to the subspace $\mathcal{H}_b^{1,0}$ annihilated by $\bar{\partial}_b^*$. The correct statement of the inhomogeneous Cauchy-Riemann equations is

$$\bar{\partial}_b u = f - S_b^{1,0} f \quad \text{with} \quad u \perp \mathcal{H}_b.$$

Again, the solution u is given by an integral kernel,

$$u(x) = \int_{\mathfrak{M}} H(x, y) f(y)\, dy$$

and another of our goals is to describe the singularities of $H(x, y)$. Finally we will also consider the equation

$$\Box_b u = \bar{\partial}_b^* \bar{\partial}_b u = f - S_b f$$

with $u \perp \mathcal{H}_b$. Then, u is given by

$$u(x) = \int_{\mathfrak{M}} R(x, y) g(y)\, dy$$

and we will describe the singularities of $R(x, y)$.

Let r denote a defining function of $b\Omega$, such that r is C^∞ in a neighborhood of $b\Omega$, $dr \neq 0$, and $r(P) > 0$ if $P \notin \bar{\Omega}$, $r(P) = 0$ if $P \in b\Omega$, $r(P) < 0$ if $P \in \Omega$. Let L, $\bar{L}$ and T be the vector fields on $b\Omega$ defined by

$$L = r_{z_2} \frac{\partial}{\partial z_1} - r_{z_1} \frac{\partial}{\partial z_2}, \tag{1}$$

$$\bar{L} = r_{\bar{z}_2} \frac{\partial}{\partial \bar{z}_1} - r_{\bar{z}_1} \frac{\partial}{\partial \bar{z}_2},$$

$$T = \sqrt{-1}\left(r_{\bar{z}_1} \frac{\partial}{\partial z_1} + r_{\bar{z}_2} \frac{\partial}{\partial z_2} - r_{z_1} \frac{\partial}{\partial \bar{z}_1} - r_{z_2} \frac{\partial}{\partial \bar{z}_2} \right).$$

These vector fields are linearly independent at each $P \in b\Omega$ and the vector field $[L, \bar{L}]$ is tangent to $b\Omega$ hence we may write

$$[L, \bar{L}] = \frac{1}{\sqrt{-1}} \theta T + aL + b\bar{L}. \tag{2}$$

The condition that Ω is pseudo-convex is then expressed by $\theta \geqslant 0$. Let $\mathcal{H}_b \subset L_2(b\Omega)$ denote the space of square integrable fuctions which are annihilated by $\bar{L}$ in the distribution sense (*i.e.* $h \in \mathcal{H}_b$ means $h \in L_2(b\Omega)$ and $(h, \bar{L}^* v) = 0$ for all $v \in C^\infty(b\Omega)$). It is proved in [K] and in [BS] that the range of $\bar{L}$ in $L_2(b\Omega)$ is closed. Thus given $f \in L_2(b\Omega)$ with f orthogonal to the null space of $\bar{L}^*$ there exists a unique solution u such that

$$\bar{L} u = f \quad \text{and} \quad u \perp \mathcal{H}_b. \tag{3}$$

Also, given $g \perp \mathcal{H}_b$ there exists a unique u such that

$$\bar{L}^* \bar{L} u = g \quad \text{and} \quad u \perp \mathcal{H}_b. \tag{4}$$

Denote by S_b: $L_2(b\Omega) \to \mathcal{H}_b$ the orthogonal projection. Note that if $f, \bar{L}f \in \in L_2(b\Omega)$ and if u satisfies

$$\bar{L}u = \bar{L}f \quad \text{with} \quad u \perp \mathcal{H}_b \tag{5}$$

then, by linear algebra we have

$$S_b f = f - u. \tag{6}$$

Thus there exist distributions K, H and R on $b\Omega \times b\Omega$ such that

$$S_b f(x) = \int_{b\Omega} K(x,y) f(y)\, dy, \tag{7}$$

$$u(x) = \int_{b\Omega} H(x,y) f(y)\, dy \tag{8}$$

implies that u satisfies (3) whenever f is orthogonal to the null space of $\bar{L}^*$, and

$$u(x) = \int_{b\Omega} R(x,y) g(y)\, dy \tag{9}$$

implies that u satisfies (4) whenever $g \perp \mathcal{H}_b$.

The purpose of this paper is to prove the optimal non-isotropic estimates for K, H and R in a neighborhood of (P,P) where $P \in b\Omega$, is of finite type. Such estimates have been obtained by M. Christ (see [Ch2]) and the estimates of the Szegö kernel are contained in Nagel, Rosay, Stein, and Wainger (see [NRSW]). We will describe our results more precisely below but first we formulate the above in a slightly more general setting.

Let $\mathfrak{X}$ be a two-dimensional complex manifold and let $\Omega \subset \mathfrak{X}$ be an open set such that $\bar{\Omega}$ is compact and its boundary $b\Omega$ is smooth and pseudo-convex. Now the formulas given in (1) make sense only locally and, in fact, there may not be a global non-vanishing vector field tangent to $b\Omega$ which is of type $(1,0)$, *i.e.* a combination of the $\partial/\partial z$. To overcome this difficulty let $T^{1,0}(b\Omega)$ be the subbundle of the complexified tangent bundle $\mathbb{C}T(b\Omega)$ consisting of vectors of type $(1,0)$. Let $T^{0,1}(b\Omega) = \overline{T^{1,0}(b\Omega)}$ and denote by $B^{0,1}(b\Omega)$ the dual space of $T^{0,1}(b\Omega)$. We define the operator $\bar{\partial}_b$: $C^\infty(b\Omega) \to$ sections of $B^{0,1}(b\Omega)$, by

$$\langle (\bar{\partial}_b u)_P, \bar{Z} \rangle = \bar{Z}(u), \tag{10}$$

where $P \in b\Omega$ and $\bar{Z} \in T^{0,1}_P(b\Omega)$.

On each fiber $\mathbb{C}T_P(b\Omega)$ we construct a hermitean inner product, depending smoothly on P, so that $T^{1,0}_P(b\Omega)$ is orthogonal to $T^{0,1}_P(b\Omega)$. We choose a volume element on $b\Omega$ which is compatible with this inner product and we define the spaces of square-integrable functions $L_2(b\Omega)$ and square-integrable $(0,1)$-forms $L^{0,1}_2(b\Omega)$; this is the completion of the space of sections of

$B^{0,1}(b\Omega)$. Now, in case the range of $\bar{\partial}_b$ is closed, there exists a unique u such that

$$\bar{\partial}_b u = f \quad \text{and} \quad u \perp \mathcal{H}_b, \tag{3'}$$

whenever $f \in L_2^{0,1}(b\Omega)$ and $f \perp$ null space of $\bar{\partial}_b^*$, here $\mathcal{H}_b$ denotes the null space of $\bar{\partial}_b$ in $L_2(b\Omega)$. We also have for each $g \in L_2(b\Omega)$, $g \perp \mathcal{H}_b$ a unique u such that

$$\Box_b u = \bar{\partial}_b^* \bar{\partial}_b u = g \quad \text{and} \quad u \perp \mathcal{H}_b. \tag{4'}$$

The condition that the range of $\bar{\partial}_b$ is closed is not always satisfied. In [K] it is proved that the range of $\bar{\partial}_b$ is closed whenever there exists a strictly plurisubharmonic function defined in a neighborhood of $b\Omega$. Thus, in particular, if Ω is a pseudo-convex domain in a Stein manifold then the range of $\bar{\partial}_b$ is closed. The kernels defined by (7), (8), and (9) can thus be interpreted in this more general situation and our results also apply to them. Our results apply to the case when $b\Omega$ is replaced by an abstract three-dimensional, pseudo-convex, compact *CR* manifold $\mathfrak{M}$. The *CR* structure on a three-dimensional manifold $\mathfrak{M}$ is given by a one-dimensional subbundle $T^{1,0}(\mathfrak{M})$ of the complexified tangent bundle $\mathbb{C}T(\mathfrak{M})$ with the property that

$$T^{1,0}(\mathfrak{M}) \cap \overline{T^{1,0}(\mathfrak{M})} = \{0\}.$$

All of the above notions can then be defined on $\mathfrak{M}$ and if the range of $\bar{\partial}_b$ is closed then the kernels given by (7), (8), and (9) make sense and our results apply to them.

Consider $P_0 \in \mathfrak{M}$, with $\mathfrak{M}$ a three-dimensional, pseudo-convex *CR* manifold. Then there exist a neighborhood U of P and linearly independent vector fields L, $\bar{L}$ and T on U such that L has values in $T^{1,0}(\mathfrak{M})$, T is real and (2) is satisfied with $\theta \geqslant 0$. We say that P_0 is of *finite type* if for some m the vector fields L, $\bar{L}$ and their commutators of order less than or equal to m span the tangent space of $b\Omega$ at P. We say that P_0 is of *type* m if m is the least number satisfying this condition. If P_0 is of type m, then there is a small neighborhood U of P_0 in which every point is of type less than or equal to m. For the statements of our results we fix such a neighborhood U.

The results in this paper are stated in terms of a family of non-Euclidean balls on «cylinders». That such cylinders are crucial is clear to anyone who has studied the work of Stein *e.g.* [NSW1], [NSW]. We define these cylinders in Section 2 and recall their basic properties. For each $P \in U$ and each $\delta \in (0, 1]$ the «cylinder» is a neighborhood of P denoted by $B(P, \delta)$. If $\delta_1 < \delta_2$ then $B(P, \delta_1) \subset B(P, \delta_2)$ and $\bigcap_\delta B(P, \delta) = \{P\}$. The $B(P, \delta)$ have height proportional to δ and a base whose radius is proportional to $\gamma(P, \delta)$ with $C\delta^{1/2} \leqslant \gamma(P, \delta) \leqslant C'\delta^{1/m}$. Let $X_1 = \mathrm{Re}\,(L)$ and $X_2 = \mathrm{Im}\,(L)$. Our main theorem is formulated in terms of the following definition.

Definition. *If Q is a distribution on $\mathfrak{M} \times \mathfrak{M}$ and if $P_0 \in \mathfrak{M}$ is of type m with a neighborhood U as above then Q is of non-isotropic order q, with $q = 0,1$, and 2 if whenever p_k is a homogeneous polynomial in four non-commuting variables then there exists C_k such that*

$$|p_k(X_x, X_y)Q(x,y)| \leqslant C_k \frac{\gamma(P,\delta)^{q-k}}{\gamma(P,\delta)^2\delta} \tag{11}$$

whenever $x \in B(P, \delta/3)$ and $y \in B(P,\delta)\backslash B(P,\delta/3)$. Here X_x and X_y denote X_1, X_2 acting on x and y respectively.

The purpose of this paper is to give a self-contained proof of the following theorem:

Main theorem. *Let $\mathfrak{M}$ be a three-dimensional compact pseudo-convex CR manifold. Suppose that the operator $\bar{\partial}_b$ has closed range in $L_2(\mathfrak{M})$ (as noted above this assumption is satisfied whenever $\mathfrak{M}$ is the boundary of a domain in a Stein manifold). Suppose further that $P_0 \in \mathfrak{M}$ is of type m. Then the distributions defined by* (7), (8), *and* (9) *satisfy the following*:

(*A*) *K is of non-isotropic order zero.*
(*B*) *H is of non-isotropic order one.*
(*C*) *R is of non-isotropic order two.*

These results may be found in Christ [Ch. 1, 2] and in Nagel, Rosay, Stein, and Wainger [NRSW]. The main application of this theorem is to the understanding of the mapping properties of the fundamental solutions on function spaces. This is carried out in [NRSW] by proving also certain cancellation properties of K.

(ii) Description of the proof.

In the equation $\bar{\partial}_b u = f$ the right hand side is of course a section of a line bundle (*i.e.* $B^{0,1}(\mathfrak{M})$). However, virtually all our analysis is local, and we use a local trivialization to write the $\bar{\partial}_b$ equation on U in the form $\bar{L}u = f$. Here, by abuse of notation, f is written as a function, not a section. Similarly the operator $\bar{\partial}_b^*$ when applied to sections over U can be identified with $\bar{L}^*$. By $\bar{L}^*$ we mean the formal adjoint of $\bar{L}$ which can be expressed as

$$\bar{L}^* v = -Lv + av \tag{12}$$

where $a \in C^\infty(\mathfrak{M})$.

The condition that the range of $\bar{\partial}_b$ is closed is equivalent to the estimate

$$\|u\| \leqslant C\|\bar{\partial}_b u\|, \tag{13}$$

for all $u \perp \mathcal{H}_b$. It is also equivalent to the condition that the range of $\bar{\partial}_b^*$ is closed which, in turn, is equivalent to the estimate

$$\|v\| \leqslant c\|\bar{\partial}_b^* v\|, \tag{14}$$

for all v orthogonal to the null space of $\bar{\partial}_b^*$. We will make essential use of the fact that if the range of $\bar{\partial}_b$ is closed then $\mathcal{H}_b$ is the orthogonal complement of the range of $\bar{\partial}_b^*$ and hence whenever $u \perp \mathcal{H}_b$ there exists v so that $u = \bar{\partial}_b^* v$.

The main theorem is a consequence of the following three estimates. First, there exists $C > 0$ such that

$$\|u\|_{B(P,\delta)} \leqslant C\gamma(P,\delta)\left(\sum_1^2 \|X_j u\| + \|u\|\right), \tag{15}$$

for $u \in C^\infty(\mathfrak{M})$, $P \in U$, and $\delta \in (0, 1]$, where $\|u\|_{B(P,\delta)}$ denotes the L_2-norm over $B(P, \delta)$.

Second, there exists $\epsilon > 0$ such that if $\zeta, \zeta' \in C_0^\infty(U)$ with $\zeta' = 1$ in a neighborhood of the support of ζ then for every $s \in \mathbb{R}$ there exists $C_s > 0$ such that

$$\|\zeta u\|_{s+\epsilon} \leqslant C(\|\zeta' \bar{L} u\|_s + \|u\| + \|v\|), \tag{16}$$

whenever $u = \bar{L}^* v$, more precisely we should have $u = \bar{\partial}_b^* G$ with $G = v\bar{\omega}$ in U and v should really be replaced by G. Throughout this paper we will continue to treat sections of $B^{0,1}(\mathfrak{M})$ as if they were functions. This will not lead us to any difficulties since any operations we perform on these sections will be confined to U and on U the sections can be identified with functions as noted above.

Third, suppose that $\bar{\partial}_b u = f$. Then, if $u \perp \mathcal{H}_b$ we have

$$\|L(\zeta u)\| \leqslant C\|f\|. \tag{17}$$

These estimates imply the theorem, as in Christ [Ch. 1], [Ch. 2] and in Nagel, Rosay, Stein and Wainger [NRSW]. We are grateful to E.M. Stein for conversations directing our attention to the search for a simple, direct proof of (17). It follows from (16) that the kernels K, H and R are C^∞ off the diagonal. This can be expressed by the following inequalities: given closed sets $B_1, B_2 \subset U$ with $B_1 \cap B_2 = \varnothing$ and $k > 0$ then there exists $C > 0$ such that

$$\sup |D_x^\alpha D_y^\beta K(x, y)| \leqslant C, \tag{18}$$

where the supremum is taken over $x \in B_1$, $y \in B_2$, and $|\alpha| + |\beta| \leqslant k$. The same inequalities hold for H and R. To go from (18) to the desired estimates, given by (11), we «rescale». Roughly speaking this means that for each $P \in U$ and $\delta \in (0, 1]$ we construct a one-to-one map of $B(P, \delta)$ to a fixed open set $\tilde{U} \subset \mathbb{R}^3$.

This map takes the *CR* structure on U to a *CR* structure on $\tilde{U}$ and the kernels K, H and R to kernels $\tilde{K}$, $\tilde{H}$ and $\tilde{R}$ respectively. Estimates (15), (16) and (17) are then used to show that (18) holds for $\tilde{K}$, $\tilde{H}$ and $\tilde{R}$ with C independent of P and δ. The proof of (18) for the operator $\tilde{H}$ depends on a two step rescaling devised by Christ [Ch. 2], we reproduce his proof in Section 7. The estimates (11) then follow from scaling back to the original kernels.

We will briefly discuss the proofs of the estimates (15), (16) and (17). Estimate (15) follows from the results of Rothschild and Stein (see [RS]), we give a self contained proof using microlocal techniques. Estimate (16) is proved in [K1] and a more detailed proof is presented in this paper. Here we indicate the idea of this proof since it is used again in proving (17).

Choose a coordinate system y_0, y_1, y_2 in U with origin at P_0 such that $T = g\partial/\partial y_0$ with $g(0) > 0$ we assume that U is small enough so that $g \geqslant \text{const.} > 0$ on U. Let ξ_0, ξ_1, ξ_2 be the dual coordinates and let $\xi' = (\xi_1, \xi_2)$. We will microlocalize by making use of multipliers of the Fourier transform in the following classes.

$$\text{(19)}\quad \mathcal{C} = \{p \in C^\infty(\mathbb{R}^3): p \geqslant 0,\ p(t\xi) = p(\xi) \text{ whenever } t \geqslant 1 \text{ and } |\xi| \geqslant 1\}$$

$$\mathcal{C}^+ = \{p \in \mathcal{C}: \text{there exists } a > 0 \text{ such that } p(\xi) = 0 \text{ when } \xi_0 < a|\xi'| \text{ and } |\xi| \geqslant 1\}.$$

$$\mathcal{C}^0 = \{p \in \mathcal{C}: \text{there exists } a > 0 \text{ such that } p(\xi) = 0 \text{ when } |\xi_0| > a|\xi'|\}$$

$$\mathcal{C}^- = \{p \in \mathcal{C}: \text{there exists } a > 0 \text{ such that } p(\xi) = 0 \text{ when } \xi_0 > -a|\xi'|\}.$$

For $u \in C_0^\infty(U)$ we define $\mathcal{P}^+u(x)$, $\mathcal{P}^0u(x)$, and $\mathcal{P}^-u(x)$ by

$$\int e^{ix\cdot\xi} p(\xi)\hat{u}(\xi)\, d\xi,$$

where $p \in \mathcal{C}^+$, $p \in \mathcal{C}^0$, and $p \in \mathcal{C}^-$ respectively.

To prove (16) we use the Hörmander estimate, there exist $\epsilon > 0$, $C > 0$ such that

$$\text{(20)}\qquad \|u\|_\epsilon^2 \leqslant C\Big(\sum \|X_j u\|^2 + \|u\|^2\Big)$$

for all $u \in C_0^\infty(U)$, in [H]. Note that

$$\sum \|X_j u\|^2 + \|u\|^2 \sim \|Lu\|^2 + \|\bar{L}u\|^2 + \|u\|^2.$$

Hence to prove (16) for $s = 0$ it is natural to try to estimate $L(\zeta u)$ by $\bar{L}(\zeta u)$. We have

$$\begin{aligned}\|L(\zeta u)\|^2 &= -(\bar{L}L(\zeta u), \zeta u) + \cdots \\ &= -(L\bar{L}(\zeta u), \zeta u) + ([L, \bar{L}](\zeta u), \zeta u) + \cdots \\ &= \|\bar{L}(\zeta u)\|^2 + (\theta T(\zeta u), \zeta u) + \cdots,\end{aligned}$$

where the terms denoted by dots are $O((\|L(\zeta u)\| + \|\bar{L}(\zeta u)\|)\|\zeta u\| + \|\zeta u\|^2)$. Thus we have

$$\|L(\zeta u)\|^2 \leqslant (\theta T(\zeta u), \zeta u) + C(\|\bar{L}(\zeta u)\|^2 + \|\zeta u\|^2). \tag{21}$$

If the first term on the right were negative we could delete it and obtain the estimate for $L(\zeta u)$. In (21) we replace ζu by $\zeta' \mathcal{P}^- \zeta u$, with $\zeta' = 1$ on a neighborhood of supp (ξ). Then the first term on the right can be written as $(R\zeta' \mathcal{P}^- \zeta u, \zeta' \mathcal{P}^- \zeta u)$ where R is a first order pseudo-differential operator with principal symbol equal to $\theta(x)g(x)\xi_0$ when $\xi_0 \leqslant 0$ and 0 if $\xi_0 \geqslant 0$. Hence from Garding's inequality, and (20) we obtain

$$\|\mathcal{P}^- \zeta u\|_\epsilon^2 \leqslant C\|\mathcal{P}^- \zeta \bar{L} u\|^2 + \cdots, \tag{22}$$

where the dots represent terms that get absorbed later. Since $\mathcal{P}^0$ microlocalizes to the elliptic region we also have

$$\|\mathcal{P}^0 \zeta u\|_1^2 \leqslant C\|\mathcal{P}^0 \zeta \bar{L} u\|^2 + \cdots. \tag{23}$$

Let Λ^s be the operator defined by

$$\widehat{\Lambda^s u}(\xi) = (1 + |\xi|^2)^{s/2} \hat{u}(\xi). \tag{24}$$

Replacing ζu by $\Lambda^\epsilon \zeta u$ in (21) and proceeding as in the derivation of (22) we obtain

$$\|\mathcal{P}^- \zeta u\|_{2\epsilon}^2 \leqslant C\|\mathcal{P}^- \zeta \bar{L} u\|_\epsilon^2 + \cdots. \tag{25}$$

Interchanging the role L and $\bar{L}$ we get

$$\|\mathcal{P}^+ \zeta v\|_{2\epsilon} \leqslant C\|\mathcal{P}^+ \zeta \bar{L}^* v\|_\epsilon + \cdots \tag{26}$$

since $\bar{L}^* v = -Lv + av$.

Since $u = \bar{L}^* v$ we have

$$\begin{aligned} \|\mathcal{P}^+ \zeta u\|_\epsilon^2 &= (\mathcal{P}^+ \zeta u, \Lambda^{2\epsilon} \mathcal{P}^+ \zeta \bar{L}^* v) \\ &= (\mathcal{P} u^+ \zeta \bar{L} u, \Lambda^{2\epsilon} \mathcal{P}^+ \zeta v) + \cdots \\ &\leqslant \|\mathcal{P}^+ \bar{L} u\| \, \|\mathcal{P}^+ \zeta v\|_{2\epsilon} + \cdots \\ &\leqslant \|\mathcal{P}^+ \mathcal{P} \zeta \bar{L} u\| \, \|\mathcal{P}^+ \zeta u\|_\epsilon + \cdots. \end{aligned}$$

We thus obtain

$$\|\mathcal{P}^+ \zeta u\|_\epsilon^2 \leqslant C\|\mathcal{P}^+ \zeta \bar{L} u\|^2 + \cdots. \tag{27}$$

Now combining (22), (23) and (27) we obtain (16) with $s = 0$. Then replacing ζu by $\Lambda^s(\zeta u)$ with $s \leqslant k\epsilon$ in (21) and proceeding by induction on k we obtain (16).

Our proof of (17) requires a further microlocalization. Let

$$\psi \in C_0^\infty(\{\xi \in \mathbb{R}^3 : 0 < a < |\xi| < b\}).$$

Let M be a fixed large number and $\delta \in (0, 1]$ then for $u \in C_0^\infty(U)$ we define $\Gamma_\delta u$ by

$$\widehat{\Gamma_\delta u(\xi)} = \psi\left(\frac{\delta}{M}\xi\right)\hat{u}(\xi). \tag{28}$$

In Section 5 we introduce for $s \geqslant 0$, the norms $|||u|||_s$, with $u \in C_0^\infty(U)$, defined by

$$|||u|||_s^2 = \int_0^1 \int_{\mathbb{R}^3} \gamma(P, \delta)^{-2s} |\Gamma_\delta u(P)|^2 \, dP \frac{d\delta}{\delta} + \|u\|^2. \tag{29}$$

We prove that

$$M^\epsilon |||u|||_{s+1} \leqslant C_s \sum_1^r |||X_j u|||_s + C\|u\|, \tag{30}$$

with C_s independent of M. The idea is to prove (17) by using the same type of argument as in the proof of (16) but with the $|||\bullet|||_s$-norms instead of the Sobolev norms and with 1 instead of ϵ. In order to carry this out, and to prove (30), we have to also localize in x-space. To do this we construct functions $\sigma_{P\delta}^0$ which essentially localize to $B(P, \delta)$ but which are supported in larger cylindrical neighborhoods denoted by $\tilde{B}(P, \delta)$. The $\tilde{B}(P, \delta)$ have the same size base as $B(P, \delta)$ but their height is proportional to $\delta\gamma(P, \delta)^{-1}$. The basic fact is that $\gamma(Q, \delta)$ is «almost» constant when $Q \in \tilde{B}(P, \delta)$. We then show that

$$|||u|||_s^2 \sim \int_0^1 \int_{\mathbb{R}^3} \gamma(P, \delta)^{-2s} \|\sigma_{P\delta}^0 \Gamma_\sigma u\|^2 \frac{dP}{\operatorname{vol} B(P, \delta)} \frac{d\delta}{\delta} + \|u\|^2. \tag{31}$$

Now the proof of (30) proceeds by applying (20) to u replaced by $\sigma_{P\delta}^0 \Gamma_\delta u$. We derive an estimate of the form

$$M^{2\epsilon}\gamma(P, \delta)^{-2s-2} \|\sigma_{P\delta}^0 u\|^2 \leqslant C\gamma(P, \delta)^{2s} \sum \|\sigma_{P\delta}^0 \Gamma_s X_j u\|^2 + \cdots, \tag{32}$$

where the dots represent terms whose integral with respect to

$$\frac{dP}{\operatorname{vol} B(P, \delta)} \frac{d\delta}{\delta}$$

is bounded by $C|||u|||_{s+1}$, with C independent of M. The terms represented by the dots arise from various commutators and can be controlled because the dual of the ξ-support of Γ_δ is contained in a ball of radius δ and center P, and

so the distance of this ball to the complement of $\tilde{B}(P, \delta)$ is greater than $c\delta^{1-1/m}$. To prove (17) the main step is to prove an analogue of (26) with $\|\cdot\|_{2\epsilon}$ and $\|\cdot\|_{\epsilon}$ replaced by $M^{\epsilon}|||\cdot|||_2$ and $|||\cdot|||_1$, respectively. This is done replacing ζu in (21) with $\sigma^0_{P\delta}\Gamma_{\sigma}u$ and retracing the steps for (21) to (26) finally obtaining

$$\sum \|\sigma^0_{P\delta}\Gamma_{\delta}X_j\mathcal{P}^+\zeta v\|^2 \leqslant C\|\sigma^0_{P\delta}\Gamma_{\delta}\mathcal{P}^+\zeta\bar{L}^*v\|^2 + \cdots. \tag{33}$$

The dots again represent various commutators and are estimated in such a way that after multiplying by

$$\gamma(P, \delta)^{-2}\frac{dP}{\operatorname{vol} B(P, \delta)}\frac{d\delta}{\delta}$$

and integrating we get

$$\sum |||X_j\mathcal{P}^+\zeta v|||_1^2 \leqslant C|||\mathcal{P}^+\zeta v|||_1^2 + C|||\mathcal{P}^+\zeta v|||_2^2 + \cdots.$$

Choosing M sufficiently large in (30) we obtain an estimate for $|||\mathcal{P}^+\zeta v|||_2$, analogous to (26). We can then estimate $|||\mathcal{P}^+\zeta u|||_2$. Finally we get

$$\|\sigma^0_{P\delta}\Gamma_{\delta}\mathcal{P}^+\zeta u\|^2 \leqslant \|\sigma^0_{P\delta}\Gamma_{\delta}\zeta\bar{L}u\|^2 + \cdots$$

and hence, after integration with respect to

$$\frac{dP}{\operatorname{vol} B(P, \delta)}\frac{d\delta}{\delta},$$

the right hand side is bounded by $C\|\bar{L}u\|$ (or more precisely by $C\|\bar{\partial}_b u\|$ since we use the closed range property). Thus we get the desired bound for $\|L\mathcal{P}^+\zeta u\|$, the bounds for the term $\|L\mathcal{P}^-\zeta u\|$ and $\|L\mathcal{P}^0\zeta u\|$ follow easily from (21) with ζu replaced by $\mathcal{P}^-\zeta u$ and $\mathcal{P}^0\zeta u$, respectively. Combining these we get the desired estimates (17). The principal new point is that (17) has a rather simple proof; all else here will be familiar to experts from [Chr], [M], [NRSW].

(iii) Additional remarks.

Our main objective in writing this paper is to give a self-contained proof of the main theorem. The principal point is to prove (17). Here we wish to describe briefly how (17) also follows from our results in [FK]. Recall that in [FK] we use a pseudo-differential operator A with the following properties

$$A\mathcal{P}^+ \sim -\bar{L}L\mathcal{P}^+ \quad \text{and} \quad A\mathcal{P}^- \sim -L\bar{L}\mathcal{P}^-. \tag{34}$$

First we will show how (17) follows from the inequality

$$\sum \|X_iX_jw\| \leqslant C(\|Aw\| + \|w\|), \tag{35}$$

for all $w \in C_0^\infty(U)$. It suffices to estimate $\|L\mathcal{P}^+ \zeta u\|$, when $u \perp \mathcal{H}_b$, $\bar\partial_b u = f$. We have $u = \bar L^* v \sim -Lv$ so that

$$\bar L \zeta u = -\bar L L \zeta v + \cdots$$

hence

$$-\bar L L \mathcal{P}^+ \zeta v = A\mathcal{P}^+ \zeta v = \zeta \mathcal{P}^+ \zeta' \bar L u + \cdots, \tag{36}$$

the dots, as usual represent terms some of which in the end are absorbed by the left hand side and the rest estimated by the right hand side. From (35) and (36) it then follows that

$$\begin{aligned} \|L\mathcal{P}^+ \zeta u\| &\leqslant C \sum \|L^2 \mathcal{P}^+ \zeta v\| + \cdots \\ &\leqslant C' \|f\|. \end{aligned}$$

The estimate (35) is established starting from the estimate

$$\sum \|X_i v\|^2 \leqslant C|(Av, v)| + C\|v\|^2.$$

Substituting $v = X_j w$ we get

$$\begin{aligned} \sum \|X_i X_j w\|^2 &\leqslant C|(AX_j w, X_j w)| + c\|X_j w\|^2 \\ &\leqslant C\{|(Aw, -X_j^2 w)| + |([A, X_j]w, X_j w)|\} + \cdots \end{aligned}$$

so we have

$$\sum \|X_i X_j w\|^2 \leqslant C' \|Aw\|^2 + C|([A, X_j]w, X_j w)| + \cdots$$

in order to estimate the second term on the right side we have to either microlocalize as above or follow a modification of the calculations in sections 5 and 6 of [FK]. The relevant L^2-estimates for A may also be deduced from the general theory in [F].

We refer the reader to [FK] for a discussion of the many related results in the area and for a more complete reference list. The estimates for the kernels obtained here can be used to prove optimal non-isotropic estimates for the corresponding operators as in Christ [Ch. 2] and Nagel, Rosay, Stein and Wainger [NRSW]. These results are also related to estimates for the $\bar\partial$-equation and the Bergman kernel. In fact, the estimates in the isotropic Hölder norms imply the corresponding estimates for $\bar\partial$ and the Bergman kernel. Recently optimal non-isotropic estimates for the Bergman kernel were obtained by McNeal (see [M]) and by Nagel, Rosay, Stein, and Wainger (see [NRSW]).

A very interesting use of these results has been made by Christ, see [Ch. 3]. He proves that a pseudo-convex, compact, three dimensional *CR* manifold of finite type, on which $\bar\partial_b$ has a closed range, can be imbedded in $\mathbb{C}^n$, for some n. His method shows how these estimates can be used to construct *CR* functions which were not accessible by the use of estimates in Sobolev norms.

We would like to thank M. Christ and E. M. Stein for very helpful conversations, and to M. Kirkham for efforts in typing this paper.

1. The First Microlocalization

Lemma 1.1. *Let $\mathfrak{M}$ be a compact pseudo-convex three-dimensional manifold with $P_0 \in \mathfrak{M}$. Then there exists a neighborhood U of P_0 and a constant $C > 0$ such that*

$$\sum_1^2 \|X_j \mathcal{P}^- w\|^2 \leqslant C\|\mathcal{P}^- \bar{L} w\|^2 + C\|\tilde{\mathcal{P}}^- w\|^2 + C\|\mathcal{R} w\|^2 \tag{1}$$

for all $w \in C_0^\infty(U)$, where $\mathcal{P}^-$ and $\tilde{\mathcal{P}}^-$ are pseudo-differential operators with symbols $p, \tilde{p} \in \mathcal{C}^-$ such that $\tilde{p}(\xi) = 1$ when $\xi \in \operatorname{supp}(p)$, and $\mathcal{R}$ is a pseudo-differential operator of order $-\infty$.

PROOF. First note that

$$\|[\mathcal{P}^-, \bar{L}] w\| \leqslant C\|\tilde{\mathcal{P}}^- w\| + C\|\mathcal{R} w\|, \tag{2}$$

where $\mathcal{R} = [\mathcal{P}^-, \bar{L}](\mathcal{P}^- - 1)$.

Since the X_j are combinations of L and $\bar{L}$ it suffices to show that

$$\|L\mathcal{P}^- w\|^2 \leqslant C\|\bar{L}\mathcal{P}^- w\|^2 + C\|\mathcal{P}^- w\|^2. \tag{3}$$

To prove (3) note that

$$\begin{aligned} \|L\mathcal{P}^- w\|^2 = \|\bar{L}\mathcal{P}^- w\|^2 &+ (\theta T \mathcal{P}^- w, \mathcal{P}^- w) \\ &+ O(\|\mathcal{P}^- w\|(\|L\mathcal{P}^- w\| + \|\bar{L}\mathcal{P}^- w\| + \|\mathcal{P}^- w\|)). \end{aligned} \tag{4}$$

Pseudo-convexity implies that the symbol of $(\mathcal{P}^-)^* \theta T \mathcal{P}^-$ is non-positive on U. Hence by Garding's inequality

$$(\theta T \mathcal{P}^- w, \mathcal{P}^- w) \leqslant C(\|\tilde{\mathcal{P}}^- w\|^2 + \|\mathcal{R} w\|^2). \tag{5}$$

Thus combining (4) and (5) we obtain (3) which completes the proof of the lemma.

Corollary 1.2. *Under the same assumptions, we have*

$$\sum_1^2 \|X_j \mathcal{P}^+ w\|^2 \leqslant C\|\mathcal{P}^+ L w\|^2 + C\|\tilde{\mathcal{P}}^+ w\|^2 + C\|w\|^2. \tag{6}$$

for $w \in C_0^\infty(U)$.

Lemma 1.3. *Under the same assumptions as Lemma* 1.1, *we have*

$$\sum_1^2 \|X_j \mathcal{P}^0 w\|^2 \leqslant C\|\mathcal{P}^0 \bar{L} w\|^2 + C\|\tilde{\mathcal{P}}^0 w\|^2 + C\|\mathcal{R} w\|^2 \tag{7}$$

for $w \in C_0^\infty(U)$.

PROOF. Taking U sufficiently small we see that if $\phi', \phi \in C_0^\infty$, and $\phi = 1$ on U, and $\phi' = 1$ on a neighborhood of $\operatorname{supp}(\phi)$ then

$$\begin{aligned}\|\bar{L}\mathcal{P}^0 w\| &\geqslant \|\phi' L \mathcal{P}^0 w\| - \|\mathcal{R} w\| \\ &\geqslant C\sum_1^2 \left\|\phi' \frac{\partial}{\partial y_j}\mathcal{P}^0 w\right\| - \text{small const.}\left\|\phi' \frac{\partial}{\partial y_0}\mathcal{P}^0 w\right\| - C\|\mathcal{R} w\| \\ &\geqslant C\|\mathcal{P}^0 w\|_1 - C\|\mathcal{R} w\|.\end{aligned}$$

This implies (7) proving the Lemma.

Lemma 1.4. *If in addition to the above assumptions we suppose that* $P_0 \in \mathfrak{M}$ *is of finite type then there exists a neighborhood* U *of* P_0 *such that if* $\zeta, \tilde{\zeta} \in C_0^\infty(U)$, *with* $\tilde{\zeta} = 1$ *on the support of* ζ, *then there exists* $\epsilon > 0$ *and for* $s \in \mathbb{R}$ *a constant* C_s *such that*

$$\|\mathcal{P}^- \zeta u\|_{s+\epsilon} \leqslant C_s(\|\mathcal{P}^- \zeta \bar{L} u\|_s + \|\tilde{\mathcal{P}}^- \tilde{\zeta} \bar{L} u\|_{s-1} + \|\mathcal{P}^- \tilde{\zeta} u\| + \|\mathcal{R}\tilde{\zeta} u\|). \tag{8}$$

PROOF. The fact that P_0 is of finite type implies that there exists a neighborhood U and constants $\epsilon > 0$ and $C > 0$ such that

$$\|w\|_\epsilon^2 \leqslant C\left(\sum_1^2 \|X_j w\|^2 + \|w\|^2\right) \tag{9}$$

for all $w \in C_0^\infty(U)$.

Now we have

$$\begin{aligned}\|\mathcal{P}^- \zeta u\|_{s+\epsilon}^2 &= \|\Lambda^s \mathcal{P}^- \zeta u\|_\epsilon^2 \\ &\leqslant \|\tilde{\zeta}\Lambda^s \mathcal{P}^- \zeta u\|_\epsilon^2 + \|\mathcal{R}\zeta u\|^2 \\ &\leqslant C\left(\sum_1^2 \|X_j \tilde{\zeta}\Lambda^s \mathcal{P}^- \zeta u\|^2 + \|\mathcal{P}^- \tilde{\zeta} u\|_s^2 + \|\mathcal{R}\zeta u\|^2\right) \\ &\leqslant C\left(\sum_1^2 \|X_j \mathcal{P}^- \tilde{\zeta}\Lambda^s \zeta u\|^2 + \|\mathcal{P}^- \tilde{\zeta} u\|_s^2 + \|\mathcal{R}\zeta u\|^2\right) \\ &\leqslant C(\|\mathcal{P}^- \bar{L}\tilde{\zeta}\Lambda^s \zeta u\|^2 + \|\tilde{\mathcal{P}}^- \tilde{\zeta} u\|_s^2 + \|\mathcal{R}\zeta u\|^2) \\ &\leqslant C(\|\mathcal{P}^- \zeta \bar{L} u\|_s^2 + \|\mathcal{P}^- \tilde{\zeta} u\|_s^2 + \|\mathcal{R}\zeta u\|^2).\end{aligned} \tag{10}$$

The desired estimate (8) then follows by induction on k, if we choose $s \leqslant k\epsilon$.

Analogously, using (6) and (7) instead of (3) we obtain

$$\|\mathcal{P}^0 \zeta u\|_{s+\epsilon} \leqslant C_s(\|\mathcal{P}^0 \zeta \bar{L} u\|_s + \|\tilde{\mathcal{P}}^0 \tilde{\zeta} \bar{L} u\|_{s-1} + \|\tilde{\mathcal{P}}^0 \tilde{\zeta} u\| + \|\mathcal{R} \tilde{\zeta} u\|) \tag{11}$$

$$\|\mathcal{P}^+ \zeta v\|_{s+\epsilon} \leqslant C_s(\|\mathcal{P}^+ \zeta \bar{L}^* v\|_s + \|\tilde{\mathcal{P}}^+ \tilde{\zeta} \bar{L}^* v\|_{s-1} + \|\tilde{\mathcal{P}}^+ \tilde{\zeta} v\| + \|\mathcal{R} \tilde{\zeta} v\|). \tag{12}$$

Proposition 1.5. *Under the same assumptions as Lemma* 1.4, *if* $u = \bar{L}^* v$ *then*

$$\|\zeta u\|_{s+\epsilon} \leqslant C_s(\|\tilde{\zeta} \bar{L} u\|_s + \|\tilde{\zeta} u\| + \|\tilde{\zeta} v\|). \tag{13}$$

PROOF. From (8) and (11) we obtain

$$\|\mathcal{P}^- \zeta u\|_{s+\epsilon} + \|\mathcal{P}^0 \zeta u\|_{s+\epsilon} \leqslant C_s(\|\tilde{\zeta} \bar{L} u\|_s + \|\tilde{\zeta} u\|). \tag{14}$$

Hence it will suffice to prove that

$$\|\mathcal{P}^+ \zeta u\|_{s+\epsilon} \leqslant C_s(\|\tilde{\zeta} \bar{L} u\|_s + \|\tilde{\zeta} u\| + \|\tilde{\zeta} v\|). \tag{15}$$

To prove (15) we write

$$\begin{aligned} \|\mathcal{P}^+ \zeta u\|^2_{s+\epsilon} &= (\mathcal{P}^+ \zeta u, \Lambda^{2s+2\epsilon} \mathcal{P}^+ \zeta \bar{L}^* v) \\ &= (\mathcal{P}^+ \zeta \bar{L} u, \Lambda^{2s+2\epsilon} \mathcal{P}^+ \zeta v) \\ &\quad + O(\|\mathcal{P}^+ \zeta u\|_{s+\epsilon}(\|\tilde{\mathcal{P}}^+ \tilde{\zeta} v\|_{s+\epsilon} + \|\tilde{\zeta} v\|)) \\ &\quad + ([\bar{L}, \mathcal{P}^+ \zeta] \tilde{\zeta} u, \Lambda^{2s+2\epsilon} \mathcal{P}^+ \zeta v). \end{aligned} \tag{16}$$

Setting $\mathcal{E} = [\bar{L}, \mathcal{P}^+ \zeta]$ we can write the last term above as

$$\begin{aligned} (\mathcal{E} \tilde{\zeta} u, \Lambda^{2s+2\epsilon} \mathcal{P}^+ \zeta v) &= (\mathcal{P}^+ \zeta u, \Lambda^{2s+2\epsilon} \mathcal{E} \tilde{\zeta} v) \\ &\quad + O(\|\tilde{\mathcal{P}}^+ \tilde{\zeta} u\|^2_{s+\epsilon-1} + \|\tilde{\mathcal{P}}^+ \tilde{\zeta} v\|^2_{s+\epsilon} + \|\tilde{\zeta} u\|^2 + \|\tilde{\zeta} v\|^2). \end{aligned} \tag{17}$$

Combining (16) and (17) we get

$$\begin{aligned} \|\mathcal{P}^+ \zeta u\|^2_{s+\epsilon} &\leqslant \text{large. const. } \|\mathcal{P}^+ \zeta \bar{L} u\|^2_s + \text{small const. } \|\mathcal{P}^+ \zeta v\|^2_{s+2\epsilon} \\ &\quad + \text{small const. } \|\mathcal{P}^+ \zeta u\|^2_{s+\epsilon} + \text{large const. } \|\tilde{\mathcal{P}}^+ \tilde{\zeta} v\|^2_{s+\epsilon} \\ &\quad + C\|\tilde{\mathcal{P}}^+ \tilde{\zeta} u\|^2_{s+\epsilon-1} + C\|\tilde{\zeta} u\|^2 + C\|\tilde{\zeta} v\|^2. \end{aligned} \tag{18}$$

Now by (12) we obtain

$$\|\mathcal{P}^+ \zeta v\|^2_{s+2\epsilon} \leqslant C_s(\|\mathcal{P}^+ \zeta u\|^2_{s+\epsilon} + \|\tilde{\mathcal{P}}^+ \tilde{\zeta} u\|^2_{s+\epsilon-1} + \|\tilde{\zeta} v\|^2) \tag{19}$$

and

$$\|\tilde{\mathcal{P}}^+ \tilde{\zeta} v\|^2_{s+\epsilon} \leqslant C_s(\|\tilde{\tilde{\mathcal{P}}}^+ \tilde{\tilde{\zeta}} u\|^2_s + \|\tilde{\tilde{\zeta}} v\|^2). \tag{20}$$

Combining the above and changing notation we have

$$\|\mathcal{P}^+\zeta u\|^2_{s+\epsilon} \leqslant C\|\mathcal{P}^+\zeta\bar{L}u\|^2_s + C(\|\tilde{\mathcal{P}}^+\tilde{\zeta}u\|^2_s + \|\tilde{\zeta}u\|^2 + \|\tilde{\zeta}v\|^2). \tag{21}$$

Hence, by induction on k with $s \leqslant k\epsilon$, we obtain the desired estimate (15) and conclude the proof.

If $\mathfrak{M}$ has the property that the range of $\bar{\partial}_b$ is closed then $u \perp \mathcal{H}_b$ if and only if there exists v so that $u = \bar{L}^*v$ and furthermore there exists a constant $C > 0$ so that

$$\|u\| \leqslant C\|\bar{L}u\| \quad \text{whenever} \quad u \perp \mathcal{H}_b, \tag{22}$$

and

$$\|v\| \leqslant C\|\bar{L}^*v\| \tag{23}$$

whenever v is orthogonal to the null space of $\bar{L}^*$.

Theorem 1.6. *Suppose that $\mathfrak{M}$ is a compact, three dimensional, pseudo-convex CR manifold on which the range of $\bar{\partial}_b$ is closed. Suppose further that $P_0 \in \mathfrak{M}$ is a point of finite type. Then there exists a neighborhood U of P_0 and an $\epsilon > 0$ such that if u is in the domain of $\bar{L}$, if $u \perp \mathcal{H}_b$, and if $\zeta\bar{L}u \in H^s$ for all $\zeta \in C_0^\infty(U)$ then $\zeta u \in H^{s+\epsilon}$ for all $\zeta \in C_0^\infty(U)$.*

PROOF. Applying (22) and (23) to (13) we obtain

$$\|\zeta u\|_{s+\epsilon} \leqslant C_s(\|\tilde{\zeta}\bar{L}u\|_s + \|\bar{L}u\|). \tag{24}$$

The theorem is then obtained using standard smoothing operators.

Theorem 1.7. *Under the same assumptions as in the theorem above denote by S_b: $L_2(\mathfrak{M}) \to \mathcal{H}_b$ the orthogonal projection. Then there exists a neighborhood U of P_0 such that if f has the property that $\zeta f \in H^s$ for all $\zeta \in C_0^\infty(U)$ then $\zeta S_b(f) \in H^s$ for all $\zeta \in C_0^\infty(U)$.*

PROOF. Let $g = f - S_b(f)$ then $\bar{L}g = \bar{L}f$, with $g \perp \mathcal{H}_b$ and $g = \bar{L}^*v$. Then, applying the previous estimates

$$\begin{aligned}
\|\zeta g\|^2_{s+\epsilon} &= (\zeta g, \Lambda^{2s+2\epsilon}\zeta\bar{L}^*v) \\
&= (\zeta\bar{L}g, \Lambda^{2s+2\epsilon}\zeta v) + O(\|\tilde{\zeta}g\|_{s+\epsilon/2}\|\tilde{\zeta}v\|_{s+3\epsilon/2}) \\
&= (\zeta\bar{L}f, \Lambda^{2s+2\epsilon}\zeta v) + O(\bullet) \\
&= (\zeta f, \Lambda^{2s+2\epsilon}\zeta\bar{L}^*v) + O(\|\tilde{\zeta}f\|_{s+\epsilon}\|\tilde{\zeta}v\|_{s+\epsilon} + \|\tilde{\zeta}g\|_{s+\epsilon/2}\|\tilde{\zeta}v\|_{s+3\epsilon/2}) \\
&= O(\|\tilde{\zeta}f\|_{s+\epsilon}(\|\zeta g\|_{s+\epsilon} + \|\tilde{\zeta}v\|_{s+\epsilon}) + (\|\tilde{\zeta}g\|_{s+\epsilon/2}\|\tilde{\zeta}v\|_{s+3\epsilon/2})) \\
&\leqslant C\|\tilde{\zeta}f\|^2_{s+\epsilon} + \text{small const.}\ \|\zeta g\|^2_{s+\epsilon} + C\|\tilde{\zeta}g\|^2_{s+\epsilon/2}.
\end{aligned}$$

Thus we obtain

$$\|\zeta g\|_{s+\epsilon} \leqslant C(\|\tilde{\zeta} f\|_{s+\epsilon} + \|f\|)$$

hence, replacing $s + \epsilon$ by s, we obtain

$$\begin{aligned}\|\zeta S_b(f)\|_s &\leqslant C\|\zeta f\|_s + C\|\zeta g\|_s \\ &\leqslant C\|\tilde{\zeta} f\|_s + C\|f\|.\end{aligned}$$

Again the proof is concluded by applying smoothing operators.

2. Localization

In this section we recall the localization studied in Sections 2 and 3 of [FK]. Our treatment here is self-contained.

Let X_1 and X_2 be vector fields in a neighborhood U of the origin in $\mathbb{R}^3$. Let T be any vector field on U such that X_1, X_2, and T are linearly independent.

Definition 2.1. *For each k-tuple $(i_1, \ldots, i_k)$, with $i_j = 1$ or 2, we define the functions $\theta_{i_1 \ldots i_k} \in C^\infty(U)$ by*

$$[X_{i_k}[X_{i_{k-1}}, \ldots, [X_{i_2}, X_{i_1}]] \ldots] = \theta_{i_1 \ldots i_k} T \bmod (X_1, X_2).$$

We say that the origin is of type m if there exists an m-tuple $(j_1, \ldots, j_m)$ such that $\theta_{j_1 \ldots j_m}(0) \neq 0$ and if for each k-tuple $(i_1, \ldots, i_k)$ with $k < m$ we have $\theta_{i_1 \ldots i_k}(0) = 0$.

Then we have

$$(1) \quad X_{j_1} \cdots X_{j_s}(\theta_{i_1 \ldots i_k}) = \theta_{i_1 \ldots i_k j_1 \ldots j_s} + \sum_\alpha a_{i_1 \ldots i_k \alpha} \theta_{i_1 \ldots i_k \alpha} + \sum_{|\beta| \leqslant s+1} a_\beta \theta_\beta$$

when $a_{i_1 \ldots i_k \alpha} \in C^\infty(U)$ and the α run over all proper subsets of $\{j_1, \ldots, j_s\}$ and $a_\beta \in C^\infty(U)$.

Note that the $\theta_{i_1 \ldots i_k}$ depend on the choice of T, if $\theta'_{i_1 \ldots i_k}$ corresponds to T' then there is a constant $C > 0$ such that

$$C^{-1}|\theta_{i_1 \ldots i_k}(P)| \leqslant |\theta'_{i_1 \ldots i_k}(P)| \leqslant C|\theta_{i_1 \ldots i_k}(P)| \quad \text{for all} \quad P \in U.$$

Let y_0, y_1, y_2 be coordinates in $\mathbb{R}^3$ chosen so that, for suitably small U, we have X_1, X_2, and $\partial/\partial y_0$ are linearly independent on U. For the remainder of this paper we will set

$$T = \frac{\partial}{\partial y_0}.$$

Definition 2.2. *For each $P \in U$ we define λ_k^P by*

$$\lambda_k^P = \sum_{j \le k} |\theta_{i_1 \dots i_j}(P)|.$$

From now on we will assume that the origin is of type m, then we have $\lambda_m^P \geqslant \text{const.} > 0$ for all $P \in U$.

As in Section 2 of [FK] we define for each $P \in U$ a coordinate system x_0^P, x_1^P, x_2^P by

$$\begin{aligned} x_0^P &= y_0 - y_0(P) - F^P(y_1 - y_1(P), y_2 - y_2(P)) \\ x_0^P &= y_j - y_j(P), \qquad j = 1, 2. \end{aligned} \tag{2}$$

The F^P are polynomials in two variables defined as follows. First we choose vector fields X_1', X_2' such that

$$X_i' = \sum_{j=1}^{2} h_i^j X_j, \tag{3}$$

the h_i^j are smooth with $\det(h_i^j)$ bounded away from zero in a neighborhood U of the origin and

$$X_i' = a_i' \frac{\partial}{\partial y_0} + \frac{\partial}{\partial y_i} \qquad i = 1, 2. \tag{4}$$

Lemma 2.3. *For each $P \in U$ there exists a unique polynomial F^P in two variables, defining x^P by* (2), *such that $F^P(0,0) = 0$, and the functions b_i^P, defined by*

$$X_i' = b_i^P \frac{\partial}{\partial x_0^P} + \frac{\partial}{\partial x_i^P}, \tag{5}$$

satisfy

$$\left[\left(\frac{\partial}{\partial x_1^P} \right)^j b_1^P \right]_P = 0 \quad \textit{for} \quad j = 0, 1, \dots, m-1 \tag{6}$$

and

$$\left[\left(\frac{\partial}{\partial x_1^P} \right)^i \left(\frac{\partial}{\partial x_2^P} \right)^j b_2^P \right]_P = 0 \quad \textit{for} \quad i + j = 0, 1, \dots, m-1. \tag{7}$$

PROOF. From (2), (4), and (5) we obtain

$$\begin{aligned} b_i^P(x^P) = a_i'(x_0^P + y_0(P) + F^P(x_1^P, x_2^P), x_1^P + y_1(P), x_2^P(P) + y_2(P)) \\ - F_{x_i^P}^P(x_1^P, x_2^P). \end{aligned} \tag{8}$$

Then (6) together with (8) determine

$$\left[\left(\frac{\partial}{\partial x_1^P}\right)^i F^P\right]_0 \quad \text{for} \quad i = 1, \ldots, m$$

and together with (7) the

$$\left[\left(\frac{\partial}{\partial x_1^P}\right)^i \left(\frac{\partial}{\partial x_2^P}\right)^j F^P\right]_0$$

are determined for $i + j \leqslant m$. □

In terms of the x^P-coordinates we have

$$X_i = a_i^P \frac{\partial}{\partial x_0^P} + \sum_{j=1}^{2} a_{ij}^P \frac{\partial}{\partial x_j^P}. \tag{9}$$

Lemma 2.4. *There exists $C > 0$ such that for all $P \in U$ we have*

$$\left|\left[\left(\frac{\partial}{\partial x_1^P}\right)^s \left(\frac{\partial}{\partial x_2^P}\right)^t a_i^P\right]_0\right| \leqslant C\lambda_{s+t+1}^P, \tag{10}$$

and

$$\left|\left[\left(\frac{\partial}{\partial x_1^P}\right)^s \left(\frac{\partial}{\partial x_2^P}\right)^t \theta_{i_1 \cdots i_k}\right]_0\right| \leqslant C\lambda_{k+s+t}^P. \tag{11}$$

PROOF. First observe that

$$a_i^P = \sum_{j=1}^{2} g_i^j b_j^P,$$

where (g_i^j) is the inverse of (h_i^j) defined by (3). Hence it suffices to prove (10) with a_i^P replaced by b_i^P, in fact by b_1^P because of (7). Furthermore, defining $\theta'_{i_1 \cdots i_k}$ by

$$[X'_{i_k}, [X'_{i_{k-1}}, \ldots [X'_{i_2}, X'_{i_1}]], \ldots] = \theta'_{i_1 \cdots i_k} T \bmod (X'_1, X'_2), \tag{12}$$

we have

$$\theta_{i_1 \cdots i_k} = \sum_{s \leq k} g^{j_1 \cdots j_s} \theta'_{j_1 \cdots j_s}, \tag{13}$$

where the $g^{i_1 \cdots i_s}$ are polynomials in $X_{r_1} \cdots X_{r_t}(g_i^j)$; thus it suffices to prove (11) with $\theta_{i_1 \cdots i_k}$ replaced by $\theta'_{i_1 \cdots i_k}$. From (12) we see that

$$\theta'_{12} = b_{1x_2}^P - b_{2x_1}^P + b_2^P b_{1x_0}^P - b_1^P b_{2x_0}^P \tag{14}$$

so (10) for $s + t = 1$ and (11) for $s + t = 0$, $k = 2$ follow.

Rewriting (1), we have with $s + t < m$

$$(X_1')^s(X_2')^t(\theta'_{i_1\cdots i_k}) = \theta'_{i_1\cdots i_k \underbrace{1\cdots 1}_{s\text{-times}}\underbrace{2\cdots 2}_{t\text{-times}}} + \sum_{|\alpha|<s+t} a_{i_1\cdots i_k\alpha}\theta_{i_1\cdots i_k\alpha} + \sum_{|\beta|<s+t+1} a_\beta\theta'_\beta.$$

We also have

$$\left(\frac{\partial}{\partial x_1^P}\right)^s\left(\frac{\partial}{\partial x_2^P}\right)^t\theta'_{i_1\cdots i_k} = (X_1')^s(X_2')^t(\theta'_{i_1\cdots i_k})$$

$$+ \sum_{r_1+r_2+q_1+q_2<s+t} a_{r_1r_2q_1q_2}\left[\left(\frac{\partial}{\partial x_1^P}\right)^{r_1}\left(\frac{\partial}{\partial x_2^P}\right)^{r_2}b_j^P\right]\left[\left(\frac{\partial}{\partial x_1^P}\right)^{q_1}\left(\frac{\partial}{\partial x_2^P}\right)^{q_2}\theta'_{i_1\cdots i_k}\right].$$

Differentiating (14) we get

$$\left(\frac{\partial}{\partial x_1^P}\right)^{r_1}\left(\frac{\partial}{\partial x_2^P}\right)^{r_2}b_1 = \left(\frac{\partial}{\partial x_1^P}\right)^{r_1}\left(\frac{\partial}{\partial x_2^P}\right)^{r_2}\theta'_{12}$$

$$+ \sum_{j_1+j_2<r_1+r_2} c_{j_1j_2}\left(\frac{\partial}{\partial x_1^P}\right)^{j_1}\left(\frac{\partial}{\partial x_2^P}\right)^{j_2}b_1$$

$$+ \sum_{j_1+j_2\le r_1+r_2} d_{j_1j_2}\left(\frac{\partial}{\partial x_1^P}\right)^{j_1}\left(\frac{\partial}{\partial x_2^P}\right)^{j_2}b_2.$$

Since b_2 vanishes to order m at 0 we get

$$\left|\left[\left(\frac{\partial}{\partial x_1^P}\right)^{r_1}\left(\frac{\partial}{\partial x_2^P}\right)^{r_2}b_1\right]_0\right| \le C\sum_{j_1+j_2\le r_1+r_2}\left|\left[\left(\frac{\partial}{\partial x_1^P}\right)^{j_1}\left(\frac{\partial}{\partial x_2^P}\right)^{j_2}\theta'_{12}\right]_0\right|.$$

Thus we obtain

$$\left|\left[\left(\frac{\partial}{\partial x_1^P}\right)^s\left(\frac{\partial}{\partial x_2^P}\right)^t\theta'_{i_1\cdots i_k}\right]_0\right| \le C\lambda_{s+t+k}^P$$

$$+ C\sum_{j_1+j_2+q_1+q_2<s+t}\left|\left(\frac{\partial}{\partial x_1^P}\right)^{j_1}\left(\frac{\partial}{\partial x_2^P}\right)^{j_2}\theta'_{12}\right]_0\left|\left|\left[\left(\frac{\partial}{\partial x_1^P}\right)^{q_1}\left(\frac{\partial}{\partial x_2^P}\right)^{q_2}\theta'_{i_1\cdots i_p}\right]\right|\right|.$$

Hence (10) and (11) follow by induction.

Definition 2.5. *For $P \in U$ and $2 \le k \le m$ we define $U^{P,k}$ to be the open set given by*

$$(15)\qquad U^{P,k} = \left\{Q \in U: \sum_{j=1}^{m-k}\lambda_{j+k}^P|(x^P)'(Q)|^j < \frac{\lambda_k^P}{C_0} \text{ and } |x_0^P(Q)| < \frac{\lambda_k^P}{C_0}\right\},$$

where

$$|(x^P)'(Q)| = (x_1^P(Q)^2 + x_2^P(Q)^2)^{1/2}$$

and C_0 is suitably large.

Lemma 2.6. *For all $Q \in U^{P,k}$ we have*

$$\frac{1}{2}\lambda_k^P \leqslant \lambda_k^Q \leqslant 2\lambda_k^P. \tag{16}$$

PROOF. Expanding $\theta_{i_1 \dots i_r}$ in Taylor series in the x^P-coordinates and applying (11) we obtain

$$\begin{aligned}(17)\quad |\theta_{i_1 \dots i_r}(P) - \theta_{i_1 \dots i_r}(Q)| &\leqslant C \sum \left| \left[\left(\frac{\partial}{\partial x_1^P}\right)^s \left(\frac{\partial}{\partial x_2^P}\right)^t \theta_{i_1 \dots i_r} \right]_0 \right| |(x^P)'(Q)|^{s+t} \\ &\quad + C|(x_0^P)(Q)| + C|(x^P)'(Q)|^m \\ &\leqslant C \sum \lambda_{r+j}^P |(x^P)'(Q)|^j + C|(x_0^P)(Q)|.\end{aligned}$$

Taking $r \leqslant k$ we have $\lambda_{r+j}^P \leqslant \lambda_{k+j}^P$ and since $Q \in U^{P,k}$ we have

$$|\theta_{i_1 \dots i_r}(P) - \theta_{i_1 \dots i_r}(Q)| \leqslant \frac{2C}{C_0}\lambda_k^P. \tag{18}$$

Thus (16) is established since

$$\lambda_k = \sum_{r \leqslant k} |\theta_{i_1 \dots i_k}|$$

and C_0 is large.

Definition 2.7. *For each $P \in U$ and $\delta > 0$ we define $\gamma(P, \delta)$ to be the positive number satisfying*

$$\delta = \sum_{j=2}^{m} \lambda_j^P \gamma(P, \delta)^j. \tag{19}$$

We define the δ-order at P, denoted by $k(P, \delta)$, to be the least integer such that

$$\lambda_{k(P,\delta)}^P \gamma(P, \delta)^{k(P,\delta)} = \max_j \lambda_j^P \gamma(P, \delta)^j. \tag{20}$$

Then

$$\lambda_{k(P,\delta)}^P \gamma(P, \delta)^{k(P,\delta)} \leqslant \delta \leqslant (m-1)\lambda_{k(P,\delta)}^P \gamma(P, \delta)^{k(P,\delta)}. \tag{21}$$

Hence

$$\gamma(P,\delta) \sim \left(\frac{\delta}{\lambda^P_{k(P,\delta)}}\right)^{1/k(P,\delta)} = \min_j \left(\frac{\delta}{\lambda^P_j}\right)^{1/j}. \tag{22}$$

Lemma 2.8. *There exists $C > 0$ such that*

$$\gamma(Q,\delta) \leqslant C\gamma(P,\delta) \quad \textit{for all} \quad Q \in U^{P,k(P,\delta)}. \tag{23}$$

PROOF. By (22) and (16), we have

$$\gamma(Q,\delta) \leqslant C\left(\frac{\delta}{\lambda^Q_{k(P,\delta)}}\right)^{1/k(P,\delta)} \leqslant C'\left(\frac{\delta}{\lambda^P_{k(P,\delta)}}\right)^{1/k(P,\delta)} \leqslant C''\gamma(P,\delta).$$

Lemma 2.9. *There exists $\eta > 0$, independent of P and δ such that the set*

$$\tilde{B}(P,\delta) = \left\{Q \in U: |(x^P)'(Q)| \leqslant \eta\gamma(P,\delta),\ |x^P_0(Q)| \leqslant \eta\frac{\delta}{\gamma(P,\delta)}\right\} \tag{24}$$

is contained in $U^{P,k(P,\delta)}$.

PROOF. From (21) we have

$$\begin{aligned} |x^P_0(Q)| &\leqslant \eta\frac{\delta}{\gamma(P,\delta)} \\ &\leqslant \eta(m-1)\lambda^P_{k(P,\delta)}\gamma(P,\delta)^{k(P,\delta)-1} \\ &\leqslant \frac{\lambda^P_{k(P,\delta)}}{C_0} \end{aligned}$$

and

$$\begin{aligned} \sum_{j=1}^{m-k} \lambda^P_{j+k(P,\delta)}|(x^P)'(Q)|^j &\leqslant \eta\sum\lambda^P_{j+k(P,\delta)}\gamma(P,\delta)^j \\ &\leqslant \eta\gamma(P,\delta)^{-k(P,\delta)}\sum\lambda^P_{j+k(P,\delta)}\gamma(P,\delta)^{j+k(P,\delta)} \\ &\leqslant \eta\gamma(P,\delta)^{-k(P,\delta)}\delta \leqslant C\eta\lambda^P_{k(P,\delta)} \leqslant \frac{\lambda^P_{k(P,\delta)}}{C_0}. \end{aligned}$$

Proposition 2.10. *There exists $C > 0$, independent of P and δ, such that*

$$\frac{1}{C}\gamma(P,\delta) \leqslant \gamma(Q,\delta) \leqslant C\gamma(P,\delta) \quad \textit{for all} \quad Q \in \tilde{B}(P,\delta). \tag{25}$$

PROOF. The second inequality follows from Lemmas 1.8 and 1.9. To show that $\gamma(P,\delta) \leqslant C\gamma(Q,\delta)$ first note that (22) can be rewritten as

$$\lambda_j^P \leqslant \delta\gamma(P,\delta)^{-j}.$$

Next, from (17), we have

$$|\theta_{i_1,\ldots,i_r}(P) - \theta_{i_1,\ldots,i_r}(Q)| \leqslant C\sum_{j=1}^{m-r} \lambda_{r+j}^P \gamma(P,\delta)^j + C\delta\gamma(P,\delta)^{-2}$$

$$\leqslant C\delta\gamma(P,\delta)^{-r}.$$

Now

$$\lambda_{k(Q,\delta)}^Q = \sum_{r\leq k(Q,\delta)} |\theta_{i_1,\ldots,i_r}(Q)|$$

$$\leqslant \lambda_{k(Q,\delta)}^P + \sum_{r\leq k(Q,\delta)} |\theta_{i_q,\ldots,i_r}(P) - \theta_{i_1,\ldots,i_r}(Q)|$$

$$\leqslant C'\delta\gamma(P,\delta)^{-k(Q,\delta)}$$

hence

$$\gamma(P,\delta) \leqslant \left(\frac{C'\delta}{\lambda_{k(Q,\delta)}^Q}\right)^{1/k(Q,\delta)} \leqslant \text{const.}\,\gamma(Q,\delta).$$

Lemma 2.11. *There exists $C > 0$ such that*

$$\left|\left[\left(\frac{\partial}{\partial x_1^P}\right)^s\left(\frac{\partial}{\partial x_2^P}\right)^t a_i^P\right]_Q\right| \leqslant C\delta\gamma(P,\delta)^{-s-t-1} \tag{26}$$

and

$$\left|\left[\left(\frac{\partial}{\partial x_1^P}\right)^s\left(\frac{\partial}{\partial x_2^P}\right)^t \theta_{i_1,\ldots,i_k}\right]_Q\right| \leqslant C\delta\gamma(P,\delta)^{-k-s-t} \tag{27}$$

for all $Q \in \tilde{B}(P,\delta)$.

PROOF. Setting

$$D_1 = \frac{\partial}{\partial x_1^P},$$

$$D_2 = \frac{\partial}{\partial x_2^P}$$

and dropping the P's and the δ's we have, using (10)

$$|D_1^t D_2^s a_i(Q)| \leqslant C \sum_{l,r} |D_1^{t+l} D_2^{s+r} a_i(0)| \Big| |x'(Q)|^{l+r} + O(|x_0(Q)| + |x'(Q)|^m)$$
$$\leqslant C\Big(\sum_{l,r} \lambda_{t+l+s+r+1} \gamma^{l+r} + \delta\gamma^{-1}\Big)$$
$$\leqslant C'\delta\gamma^{-s-t-1}$$

thus proving (26). To prove (27) we use (11) and obtain

$$|D_1^t D_2^s \theta_{i_1 \cdots i_k}(Q)| \leqslant C \sum_{l,r} |D_1^{t+l} D_2^{s+r} \theta_{i_1 \cdots i_k}(0)| \Big| |x'(Q)|^{l+r}$$
$$+ O(|x_0(Q)| + |x'(Q)|^m)$$
$$\leqslant C\Big(\sum_{l,r} \lambda_{t+l+s+r+k} \gamma^{l+r} + \delta\gamma^{-1}\Big)$$
$$\leqslant C'\delta\gamma^{-k-s-t}.$$

Definition 2.12. *Let $B(P, \delta)$ denote the subset of $\tilde{B}(P, \delta)$, defined by*

$$B(P, \delta) = \{Q \in \tilde{B}(P, \delta) : |x_0^P(Q)| < \eta\delta\}.$$

Proposition 2.13. *There exists a constant C_1 such that whenever $Q \in B(P, \delta)$ then $B(Q, C_1\delta) \supset B(P, \delta)$ and whenever $Q \in \tilde{B}(P, \delta)$ then $\tilde{B}(Q, C_1\delta) \supset \tilde{B}(P, \delta)$.*

PROOF. Note that, for $C_1 \geqslant 1$ we have

$$C_1^{1/m}\gamma(Q, \delta) \leqslant \text{const.}\ \gamma(Q, C_1\delta) \leqslant C_1^{1/2}\gamma(Q, \delta). \tag{28}$$

Thus it suffices to show that if $Q \in B(P, \delta)$ then there exists a $C_2 > 0$ such that whenever $\bar{Q} \in B(P, \delta)$ we have

$$\begin{cases} |x_0^Q(\bar{Q})| \leqslant C_2\eta\delta \\ |(x^Q)'(\bar{Q})| \leqslant C_2\eta\gamma(Q, \delta) \end{cases} \tag{29}$$

and whenever $Q, \bar{Q} \in \tilde{B}(P, \delta)$ then

$$\begin{cases} |x_0^Q(\bar{Q})| \leqslant C_2\eta\delta\gamma^{-1}(Q, \delta) \\ |(x^Q)'(\bar{Q})| \leqslant C_2\eta\gamma(Q, \delta). \end{cases} \tag{29}'$$

To prove (29) and (29)′ we express the x^Q-coordinates in terms of the x^P-coordinates, we have

$$\begin{cases} x_0^Q = x_0^P - x_0^P(Q) - F^{Q,P}(x_1^P - x_1^P(Q), x_2^P - x_2^P(Q)) \\ (x^Q)' = (x^P)' - (x^P)'(Q) \end{cases} \tag{30}$$

where $F^{Q,P}$ is a polynomial of order m of the form

$$F^{Q,P}(z_1, z_2) = \sum_{i+j=1}^{m} C_{ij}^{Q,P} z_1^i z_2^j. \tag{31}$$

As in (8) the polynomial $F^{Q,P}$ satisfies

$$F_{x_i^Q}^{Q,P}(x_1^Q, x_2^Q) = b_i^P(x_0^Q + x_0^P(Q) + F^{Q,P}(x_1^Q, x_2^Q), x_1^Q + x_1^P(Q), x_2^Q + x_2^P(Q)) - b_i^Q(x_1^Q, x_2^Q). \tag{32}$$

Hence

$$\sum_{i+j=1} |c_{ij}^{QP}| \leqslant \sum_{1}^{2} |b_i^P(Q)| \leqslant C\delta\gamma(P, \delta)^{-1}.$$

Differentiating (32) with respect to x_1^Q and x_2^Q and evaluating at Q we obtain

$$\sum_{i+j=r} |c_{ij}^{QP}| \leqslant C\delta\gamma(P, \delta)^{-r}.$$

Hence from (30) we get

$$|x_0^Q(\bar{Q})| \leqslant \text{const. } \eta\delta$$

and

$$|(x^Q)'(\bar{Q})| \leqslant \text{const. } \eta\gamma(P, \delta),$$

when $Q, \bar{Q} \in B(P, \delta)$, and

$$|x_0^Q(\bar{Q})| \leqslant \text{const. } \eta\delta\gamma^{-1}(P, \delta)$$
$$|(x^Q)'(\bar{Q})| \leqslant \text{const. } \eta\gamma(P, \delta),$$

when $Q, \bar{Q} \in \tilde{B}(P, \delta)$.

The desired inequalities (29) and (29)′ then follow from Proposition 2.10.

To conclude this section we define $\sigma_{P\delta}^0 \in C_0^\infty(B(P, \delta))$ and $\sigma_{P\delta} \in C_0^\infty(\tilde{B}(P, \delta))$. Let $\tau^0, \tau^1 \in C_0^\infty([0, \infty))$ such that $\tau^0 = 1$ in a neighborhood of 0, $\tau^1 = 1$ in a neighborhood of the support of τ^0, and such that the support of τ^1 is sufficiently small so that

$$\text{supp } \{\tau^1(|x_0^P|\gamma(P, \delta)\delta^{-1})\tau^1(|(x^P)'|\gamma(P, \delta)^{-1})\} \subset \tilde{B}(P, \delta).$$

Setting

$$\phi_{P\delta}(x^P) = \tau^1(|x_0^P|\gamma(P, \delta)\delta^{-1})\tau^1(|(x^P)'|\gamma(P, \delta)^{-1}) \tag{33}$$

we define

$$\sigma_{P\delta}^0(x^P) = \tau^0(|x_0^P|\delta^{-1})\tau^0(|(x^P)'|\gamma(P, \delta)^{-1}), \tag{34}$$

and

$$\sigma_{P\delta}(x^P) = \left(1 + \frac{|x_0^P|}{\delta}\right)^{-N} \phi_{P\delta}(x^P), \tag{35}$$

where N is a large number that will be fixed later.

From (26) we easily deduce

$$|X_{i_1} \cdots X_{i_k} \sigma^0_{P\delta}| \leqslant C\gamma(P, \delta)^{-k} \sigma_{P\delta}. \tag{36}$$

3. Second Microlocalization

In this section we study the operators Γ_δ which localize the Fourier transform of a function to a region where $|\xi| \approx M/\delta$. We prove a number of fairly standard properties of these operators, several of these results (or slight variants thereof) were already established in [FK] and we include them here for completeness.

Definition 3.1. *Let* $\psi \in C_0^\infty(\{\xi \in \mathbb{R}^3 : 0 < a < |\xi| < b\})$. *For* $u \in H^{-s_0}$ *and* $\delta > 0$ *we define* $\Gamma_\delta u$ *by*

$$\widehat{\Gamma_\delta u}(\xi) = \psi\left(\frac{\delta}{M}\xi\right)\hat{u}(\xi).$$

The following is part A of Lemma 1 in Section 1 of [FK].

Lemma 3.2. *Let* $R_1, R_2, \ldots, R_n$ *be pseudo-differential operators such that the supports of their symbols have no point in common. If one of the* $R_1, \ldots, R_n$ *has symbol supported in* $|\xi| \approx M/\delta$ *(that is, in the region* $aM/\delta < |\xi| < bM/\delta$*), then*

$$\|R_1 \cdots R_n u\|_{s_0} \leqslant C\left(\frac{\delta}{M}\right)^{\text{power}} \|u\|_{-s_0}. \tag{1}$$

The «power» and s_0 *may be taken arbitrarily large.*

PROOF. Say R_j has symbol supported in $\{|\xi| \approx M/\delta\}$. Let Q be a pseudo-differential operator of order K, with symbol $q(\xi)$ such that

$$q(\xi) = \left(\frac{M}{\delta}\right)^K$$

on the support of R_j and

$$|\partial^\alpha q(\xi)| \leqslant C_\alpha (1 + |\xi|)^{K - |\alpha|}$$

with C_α independent of δ and M. Then, by the standard theory of pseudo-differential operators, we have

$$\|R_1 \cdots R_j Q R_{j+1} \cdots R_n u\|_{s_0} \leqslant \text{const. } \|u\|_{-s_0}$$

where the constant is independent of δ and M. The desired inequality (1) then follows since

$$R_j = \left(\frac{\delta}{M}\right)^K R_j Q.$$

The following result is again a slight variant of a well-known property of standard pseudo-differential operators, it is essential in our analysis and appears as Lemma 2 of Section 1 in [FK], we reproduce it here for completeness.

Lemma 3.3. *Let $T = a(x, D)$ with symbol $a(x, \xi)$ of order zero supported in $|\xi| \approx M/\delta$. Let ϕ be a smooth change of coordinates defined in a fixed neighborhood of 0. If u is supported in a small neighborhood of 0, then*

$$Su(x) = [T(u \circ \phi^{-1})] \circ \phi(x)$$

is well-defined for x in a small neighborhood of 0. There exists a symbol $\tilde{a}(x, \xi)$ of order 0, supported in $|\xi| \approx M/\delta$, so that

$$\|Su - \tilde{a}(x, D)u\|_{H^s(\text{small nbd of } 0)} \leqslant c\delta^{\text{power}} \|u\|_{-s}.$$

Here, «power» and s are as large as we please.

PROOF. For x near 0 and u supported near 0 we have

$$Su(x) = \int e^{i\xi(\phi(x) - \phi(y))} a(\phi(x), \xi) \frac{1}{\det \phi'(y)} u(y)\, dy\, d\xi. \tag{2}$$

Now

$$\xi(\phi(x) - \phi(y)) = \xi\left[\int_0^1 \phi'(tx + (1 - t)y)\, dt\right](x - y) = \eta(x - y)$$

with

$$\eta = G(x, y)\xi, \qquad G(x, y) = \left[\int_0^1 \phi'(tx + (1 - t)y)\, dt\right]^t$$

a smooth matrix-valued function. We have $G(x, x) = (\phi'(x))^t$, so G is invertible for x near y. In our integral x and y are both near 0, so we may introduce

the smooth matrix-valued function $\tilde{G}(x, y) = (G(x, y))^{-1}$. Changing variables from y, ξ to y, η in (2), we have

$$(3) \qquad Su(x) = \int e^{i\eta(x-y)}[\det \tilde{G}(x, y)a(\phi(x), \tilde{G}(x, y)\eta)] \frac{1}{\det \phi'(y)} u(y)\, dy\, d\eta$$

$$= \int e^{i\eta(x-y)} b(x, \eta, y)u(y)\, dy\, d\eta$$

with b satisfying

$$(4) \qquad \begin{cases} |\partial_{x,y}^{\alpha} \partial_{\eta}^{\beta} b| \leqslant C_{\alpha\beta}\left(\dfrac{M}{\delta}\right)^{-|\beta|} \\ \operatorname{supp} b(x, \eta, y) \subset \left\{|\eta| \approx \dfrac{M}{\delta}\right\} \end{cases}$$

Taylor-expanding b in y about $y = x$, we have

$$(5) \qquad Su(x) = \sum_{|\alpha| \leqslant 10l} \frac{1}{\alpha!} \int e^{-\eta(x-y)}(y - x)^{\alpha}[\partial_y^{\alpha} b(x, \eta, y)|_{y=x}]u(y)\, dy\, d\eta$$

$$+ \int e^{i\eta(x-y)} b^{\#}(x, \eta, y)u(y)\, dy\, d\eta$$

where $b^{\#}$ satisfies (4) and also vanishes to order $10l$ at $x = y$.

Integrating by parts in η in (5), we get

$$Su(x) = \sum_{|\alpha| \leqslant 10l} C_{\alpha} \int e^{i\eta(x-y)}[\partial_{\eta}^{\alpha} \partial_y^{\alpha} b(x, \eta, y)|_{y=x}]u(y)\, dy\, d\eta$$

$$+ C \int \{e^{i\eta(x-y)}[|x - y|^{-2l} \Delta_{\eta}^{l} b^{\#}(x, \eta, y)]\}\, dy\, d\eta.$$

On the right side, the first term is of the form $\tilde{a}(x, D)u(x)$ with $\tilde{a}(x, \xi)$ a symbol of order zero supported in $|\xi| \approx M/\delta$.

In the second term, the integrand and its derivatives up to order l in x, y are $O(M/\delta)^{-2l}$ and supported in $|\eta| \approx M/\delta$. Hence, taking l large, we see that

$$Su(x) = \tilde{a}(x, D)u(x) + \int K(x, y)u(y)\, dy,$$

$$|\partial_x^{\alpha} \partial_y^{\beta} K| \leqslant C_l \delta^l, \qquad |\alpha| + |\beta| \leqslant l;$$

l as large as we please. The conclusion of the lemma is now obvious.

Corollary. *Let Φ be a smooth coordinate change defined on $\{|y| < 1\} \subset \mathbb{R}^3$, and let $\theta \in C_0^{\infty}(|y| < 1)$. Thus $\theta[(\Gamma_{\delta} u) \circ \phi]$ is well-defined on all $\mathbb{R}^3$. Let P be*

a pseudo-differential operator of order zero with symbol $\sigma(\xi)$ *supported in* $\{|\xi| < c_1 M/\delta\}$. *If* $c_1 << 1$, *then*

$$\|P\{\theta[(\Gamma_\delta u) \circ \phi]\}\|_s \leqslant C\delta^{\text{power}}\|u\|_{-s}.$$

PROOF. Set $\tilde{\theta} = \theta \circ \Phi^{-1}$, and take $\psi \in C_0^\infty$ equal to 1 in a neighborhood of supp $(\tilde{\theta})$. Lemma 3.3 shows that

$$\theta[(\Gamma_\delta \psi u) \circ \Phi] = \theta\tilde{a}(x, D)[(\psi u) \circ \Phi] + \mathcal{E}[(\psi u) \circ \Phi]$$

with supp $\tilde{a}(x, \xi) \subset \{|\xi| \approx M/\delta\}$ and $\mathcal{E}: H^{-s} \to H^s$ with norm $O(\delta^{\text{power}})$. Hence

$$P\{\theta[(\Gamma_\delta u) \circ \Phi)]\} = \{P\theta\tilde{a}(x, D)\}[(\psi u)\Phi] + P\mathcal{E}[(\psi u)\Phi] + P[\{\tilde{\theta}\Gamma_\delta(1 - \psi)u\} \circ \Phi].$$

Lemma 3.2 applies to $P\theta\tilde{a}(x, D)$ and to $\tilde{\theta}\Gamma_\delta(1 - \psi)$, so the corollary follows at once.

Lemma 3.4. *Let* R *be a pseudo-differential operator of order zero with symbol* $R(x, \xi)$ *which, when expressed in terms of* x^P*-coordinates has support in the support of* Γ_δ. *Then*

$$\|\sigma_\delta^0 Ru\| \leqslant C\|\sigma_\delta \tilde{\Gamma}_\delta u\| + C'\delta^{\text{power}}\|u\|_{-s_0}, \tag{6}$$

here $\sigma_\delta^0 = \sigma_{P\delta}^0$ *and* $\sigma_\delta = \sigma_{P\delta}$. *The constant* C *is independent of* M *and* δ, *and* C' *depends on* M *but not on* δ.

PROOF.

$$\begin{aligned} Ru(x) &= \int R(x, \xi)e^{ix\xi}\hat{u}(\xi)\,d\xi \\ &= \int R(x, \xi)e^{i(x-y)\xi}u(y)\,dy\,d\xi \\ &= \int H(x, y)u(y)\,dy \end{aligned} \tag{7}$$

with

$$H(x, y) = \int R(x, \xi)e^{i(x-y)\xi}d\xi. \tag{8}$$

Using the identity

$$\left(1 + \frac{M}{\delta^2}|x - y|^2\right)^{-K}\left(1 - \frac{M^2}{\delta^2}\Delta_\xi\right)^K e^{i\xi(x-y)} = e^{i\xi(x-y)} \tag{9}$$

and integrating by parts repeatedly, we may rewrite H as

$$H(x, y) = \left(1 + \frac{M}{\delta^2}|x - y|^2\right)^{-K}\int_{\mathbb{R}^3} e^{i\xi(x-y)}\left[\left(I - \frac{M^2}{\delta^2}\Delta_\xi\right)^K R(x, \xi)\right]d\xi. \tag{10}$$

The quantity in brackets is $O(1)$ and supported in $|\xi| \approx M/\delta$ in view of our hypothesis on $R(x, \xi)$. Consequently,

$$|H(x,y)| \leqslant C_K\left(\frac{M}{\delta}\right)^3\left(1+\frac{M^2}{\delta^2}|x-y|^2\right)^{-K}.$$

Substituting this into (7), we get

$$|\sigma_\delta^0(x)Ru(x)| \leqslant C'_K\left(\frac{M}{\delta}\right)^3\int\frac{|\sigma_\delta^0(x)|}{\left(1+\frac{M^2}{\delta^2}|x-y|^2\right)^K}|u(y)|\,dy. \tag{12}$$

Now it follows from the definitions of σ_δ^0 and σ_δ that

$$\frac{|\sigma_\delta^0(x)|}{\left(1+\frac{M}{\delta}|x-y|\right)^{2K}} \leqslant C|\sigma_\delta(y)| + \delta^{\text{power}}, \tag{13}$$

where «power» can be taken as large as we wish if K and N are sufficiently large (recall that N occurs in the definition of σ_δ).

Now from (12) we get

$$\begin{aligned}(14)\quad &|\sigma_\delta^0(x)Ru(x)|^2\\ &\leqslant C_K\left(\frac{M}{\delta}\right)^6\int\frac{dy}{\left(1+\frac{M}{\delta}|x-y|\right)^K}\int\frac{\sigma_\delta(x)^2|u(y)|^2}{\left(1+\frac{M}{\delta}|x-y|\right)^{2K}}dy\\ &\leqslant C\left(\frac{M}{\delta}\right)^3\int\frac{\sigma_\delta(x)^2|u(y)|^2}{\left(1+\frac{M}{\delta}|x-y|\right)^K}dy\\ &\leqslant C\left(\frac{M}{\delta}\right)^3\int\frac{\sigma_\delta(y)^2u(y)^2}{\left(1+\frac{M}{\delta}|x-y|\right)^K}dy + C\delta^{\text{power}}\|u\|^2.\end{aligned}$$

Integrating with respect to x we get

$$\|\sigma_\delta^0 Ru\| \leqslant C\|\sigma_\delta u\| + \delta^{\text{power}}\|u\|. \tag{15}$$

Substitute now $\tilde{\Gamma}_\sigma u$ for u.

We may use Lemma 3.3 to view R as a pseudo-differential operator in terms of the y-coordinates in which $\tilde{\Gamma}_\delta$ was originally defined. Since the symbol of $\tilde{\Gamma}_\delta$ equals 1 on the support of the symbol of R, we obtain from Lemma 3.3 that $\|R\tilde{\Gamma}_\delta u - Ru\|_{L^2(\operatorname{supp}\sigma_\delta^0)} \leqslant \delta^{\text{power}}\|u\|_{-s_0}$, provided that u is supported in a

small neighborhood of 0. We may remove the restriction on supp (u) by putting ϕu in place of u for a suitable cutoff function ϕ and invoking Lemma 3.2 to control the resulting error terms.

Since also

$$\|\tilde{\Gamma}_\delta u\| \leqslant C\left(\frac{M}{\delta}\right)^{s_0}\|u\|_{-s_0},$$

we obtain

$$\|\sigma_\delta^0 Ru\| \leqslant C\|\sigma_\delta\tilde{\Gamma}_\delta u\| + C'\delta^{\text{power}}\|u\|_{-s_0}. \tag{16}$$

Here C is independent of M and δ, while C' depends on M but not on δ. The proof of Lemma 3.4 is now complete.

Lemma 3.5. *Suppose that R is expressed in x^P-coordinates as a pseudo-differential operator of order zero with symbol $R(x, \xi)$ supported in* $\operatorname{supp}\tilde{\Gamma}_\delta$. *Then, again setting $\sigma_\delta^0 = \sigma_{P\delta}^0$ and $\sigma_\delta = \sigma_{P\delta}$, we have*

$$\|[\sigma_\delta^0, R]\Gamma_\delta u\| \leqslant \frac{C}{M}\|\sigma_\delta\Gamma_\delta u\| + C'\delta^{\text{power}}\|u\|_{-s_0}, \tag{17}$$

here C is independent on δ and M; and C' is independent of δ.

PROOF. We have

$$\sigma_\delta^0(x)Ru(x) = \int \sigma_\delta^0(x)R(x, \xi)e^{i\xi(x-y)}u(y)\,dy\,d\xi. \tag{18}$$

So

$$[\sigma_\delta^0, R]u(x) = \int H(x, y)u(y)\,dy \tag{19}$$

with

$$H(x, y) = \int (\sigma_\delta^0(x) - \sigma_\delta^0(y))R(x, \xi)e^{i\xi(x-y)}\,d\xi. \tag{20}$$

Again using (9) and integrating by parts, we obtain

$$H(x, y) = \left(1 + \frac{M}{\delta^2}|x - y|^2\right)^{-K}(\sigma_\delta^0(x) - \sigma_\delta^0(y)) \int e^{i\xi(x-y)}\left\{\left(I - \frac{M^2}{\delta^2}\Delta_\xi\right)^{-K}R(x, \xi)\right\}d\xi \tag{21}$$

and, just like in Lemma 3.4, that the absolute value of the integral is bounded

by $C_K M^3 \delta^{-3}$, hence

$$|H(x,y)| \leqslant \frac{C_K |(\sigma_\delta^0(x) - \sigma_\delta^0(y))| \left(\frac{M}{\delta}\right)^3}{\left(1 + \frac{M^2}{\delta^2}|x-y|^2\right)^K} \tag{22}$$

$$\leqslant C_K \left(\frac{M}{K}\right)^3 \left\{ \frac{|\sigma_\delta^0(x) - \sigma_\delta^0(y)|}{\left(1 + \frac{|x-y|}{\delta}\right)^K} \right\} \frac{1}{\left(1 + \frac{M}{\delta}|x-y|\right)^K}.$$

We will prove that if K is sufficiently large then

$$\frac{|\sigma_\delta^0(x) - \sigma_\delta^0(y)|}{\left(1 + \frac{|x-y|}{\delta}\right)^K} \leqslant C_K \sigma_\delta(y) \frac{|x-y|}{\delta} + C_K \delta^{\text{power}}. \tag{23}$$

To establish (23) we ask when does the following estimate hold

$$(\delta^{\text{power}} + \sigma_\delta(y)) \left(1 + \frac{|x-y|}{\delta}\right)^K \geqslant \text{const.} \tag{24}$$

Clearly (24) holds if $y \in \operatorname{supp}(\sigma_\delta^0)$, for then $\sigma_\delta(y) \sim 1$. Also, (24) holds if $x \in \operatorname{supp}(\sigma_\delta^0)$, by definition of $\sigma_{P,\delta}^0 = \sigma_\delta^0$. If (24) holds, then

$$\frac{|\sigma_\delta^0(x) - \sigma_\delta^0(y)|}{\frac{|x-y|}{\delta}} \leqslant C \left(1 + \frac{|x-y|}{\delta}\right)^K (\delta^{\text{power}} + \sigma_\delta(y)), \tag{25}$$

which amounts to (23).

The only case in which (24) fails is when neither x nor y is in $\operatorname{supp}(\sigma_\delta^0)$, in which case (23) is trivial.

Putting (23) into (22), we get

$$|H(x,y)| \leqslant C_K \left(\frac{M}{\delta}\right)^3 \frac{|x-y|}{\delta} \left(1 + \frac{M}{\delta}|x-y|\right)^{-K} \sigma_\delta(y) \tag{26}$$

$$+ C_K \delta^{\text{power}} \left(\frac{M}{\delta}\right)^3 \left(1 + \frac{M}{\delta}|x-y|\right)^{-K}$$

$$= H_1(x-y)\sigma_\delta(y) + \delta^{\text{power}} H_2(x-y)$$

with

$$\|H_1\|_{L_1} \leqslant \frac{C_K}{M} \quad \text{and} \quad \|H_2\|_{L_1} \leqslant C_K. \tag{27}$$

Substituting this in (19) we have

$$|[\sigma_\delta^0, R]u| \leqslant H_1 * |\sigma_\delta u| + \delta^{\text{power}} H_2 * |u| \tag{28}$$

pointwise, so that

$$\|[\sigma_\delta^0, R]u\| \leqslant \frac{C}{M}\|\sigma_\delta u\| + c\delta^{\text{power}}\|u\|. \tag{29}$$

Taking $\Gamma_\delta u$ in place of u, and noting that

$$\|\Gamma_\delta u\| \leqslant C\left(\frac{M}{\delta}\right)^{s_0}\|u\|_{-s_0}, \tag{30}$$

we obtain the desired estimate (17) thus completing the proof of Lemma 3.5.

4. Rescaled Subelliptic Estimates

The purpose of this section is to prove the following two estimates

$$M^{2\epsilon}\|\sigma_{P\delta}^0 \Gamma_\delta u\|^2 \leqslant C_1 \gamma(P,\delta)^2 \sum_j \|\sigma_{P\delta}^0 \Gamma_\delta X_j u\|^2 + C_1\|\sigma_{P\delta}\tilde{\Gamma}_\delta u\|^2 + C\delta^{\text{power}}\|u\|^2, \tag{1}$$

where C_1 is independent of δ and of M and where C depends on M but not on δ, and

$$\|u\|_{B(P,\delta)} \leqslant C\gamma(P,\delta)\{\sum \|X_j u\| + \|u\|\}, \tag{2}$$

where $\|u\|_{B(P,\delta)}$ denotes the L_2-norm of u over $B(P,\delta)$ and the constant C is independent of P and δ.

We begin with the proof of (1). For each P and δ we introduce the rescaled coordinates (y_0, y_1, y_2) by

$$y_0 = \frac{x_0^P}{\delta}, \quad \text{and} \quad y_j = \frac{x_j^P}{\gamma(P,\delta)} \quad \text{for} \quad j = 1, 2. \tag{3}$$

Let $\mathcal{B} = \{y \in \mathbb{R}^3 : (\delta y_0, \gamma(P,\delta)y') \in B(P,\delta) \text{ for all } P \in U \text{ and } \delta \in (0,1)\}$. Note that $\mathcal{B}$ is of the form $\{y \in \mathbb{R}^3 : |y_0| < C,\ |y'| < C'\}$ where C and C' are independent of P and δ. Let $Y_i = \gamma(P,\delta)X_i$, that is

$$Y_i = \frac{\gamma(P,\delta)}{\delta} a_i^P(\delta y_0, \gamma(P,\delta)y')\frac{\partial}{\partial y_0} + \sum_{j=1}^{2} a_{ij}^P(\delta y_0, \gamma(P,\delta)y')\frac{\partial}{\partial y_j}, \tag{4}$$

where the a_i^P and a_{ij}^P are given in (9) of Section 2. The Y_i are C^∞ independently of P and δ and their commutators through order m span, uniformly in P, δ

(as in [FK]). It then follows that there exist $C > 0$ and $\epsilon > 0$ independent of P and δ such that

$$\|w\|_\epsilon \leqslant C\Big(\sum \|Y_j w\| + \|w\|\Big) \tag{5}$$

for all $w \in C_0^\infty(\mathcal{B})$.

Lemma 4.1. *If S is a pseudo-differential operator of order ϵ, $0 \leqslant \epsilon \leqslant 1$, if $\zeta \in C_0^\infty(\mathcal{B})$, and if $\rho(x) = (1 + |x|^2)^{-N}$ then there exists $C > 0$ such that*

$$\|\zeta Su\| \leqslant C\|\zeta u\|_\epsilon + C\|\rho u\| \tag{6}$$

for all $u \in C_0^\infty(\mathbb{R}^n)$.

PROOF. We have

$$\begin{aligned} \|\zeta Su\| &\leqslant \|S(\zeta u)\| + \|[\zeta, S]u\|, \\ \|S(\zeta u)\| &= \|\Lambda^\epsilon \Lambda^{-\epsilon} S(\zeta u)\| \leqslant C\|\zeta u\|_\epsilon, \end{aligned} \tag{7}$$

and

$$\|[\zeta, S]u\| \leqslant C\|\rho u\|_{\epsilon - 1} \leqslant C\|\rho u\|, \tag{8}$$

Hence combining (7) and (8) we obtain (6).

Setting $w = \zeta u$ and combining (5) and (6) we get

$$\begin{aligned} \|\zeta Su\|^2 &\leqslant C\sum \|Y_j(\zeta u)\|^2 + C\|\rho u\|^2 \\ &\leqslant C\sum \|\zeta Y_j u\|^2 + C'\|\rho u\|^2. \end{aligned} \tag{9}$$

If supp $\hat{u} \subset \{\xi \in \mathbb{R}^3 : |\xi| > M\}$ then, letting S be the operator with symbol

$$S(\xi) = \chi\left(\frac{|\xi|^2}{M^2}\right) M^\epsilon,$$

where

$$\chi(t) = \begin{cases} 0 & \text{if } t < 1/2 \\ 1 & \text{if } t \geqslant 1, \end{cases} \tag{10}$$

we have $Su = M^\epsilon u$ and then

$$M^{2\epsilon}\|\zeta u\|^2 \leqslant C\sum_j \|\zeta Y_j U\|^2 + C\|\rho u\|^2. \tag{11}$$

By rescaling we then obtain

$$M^{2\epsilon}\|\sigma^0_{P\delta} u\|^2 \leqslant C\gamma(P, \delta)^2 \sum_j \|\sigma^0_{P\delta} X_j u\|^2 + C\|\rho_\delta u\|^2 \tag{12}$$

for u with

$$\operatorname{supp}(\hat{u}) \subset \left\{ \xi \in \mathbb{R}^n : |\xi| > \text{const.} \frac{M}{\delta} \right\},$$

where

$$\rho_\delta(x^P) = \left(1 + \frac{|x_0^P|^2}{\delta^2} + \frac{|(x^P)'|^2}{\gamma(P,\delta)^2}\right)^{-N}$$

Now for arbitrary u, $\Gamma_\delta u = u^\# + u_{\text{error}}$ with

$$\operatorname{supp}(\hat{u}^\#) \subset \left\{ |\xi| > C\frac{M}{\delta} \right\} \quad \text{and} \quad \|u_{\text{error}}\|_{C^k} \leqslant C\delta^{\text{power}} \|u\|_{-s_0},$$

where k, «power» and s_0 may be chosen arbitrarily large. Applying (12) to $u^\#$ we get

$$(13)\quad M^{2\epsilon}\|\sigma^0_{P\delta}\Gamma_\delta u\|^2 \leqslant C\gamma(P,\delta)^2 \sum_j \|\sigma^0_{P\delta} X_j \Gamma_\delta u\|^2 + C\|\rho_\delta \Gamma_\delta u\|^2 + \delta^{\text{power}}\|u\|^2.$$

Since

$$\sigma^0_{P\delta} X_j \Gamma_\delta u = \sigma^0_{P\delta}\Gamma_\delta X_j u + \sigma^0_{P\delta}[X_j, \Gamma_\delta]u$$

we obtain, by the results of Section 3

$$\|\sigma^0_{P\delta}[X_j, \Gamma_\delta]u\| \leqslant C\|\sigma_{P\delta}\tilde{\Gamma}_\delta u\| + \delta^{\text{power}}\|u\|.$$

Combining this with (13) we get

$$(14)\quad M^{2\epsilon}\|\sigma^0_{P\delta}\Gamma_\delta u\|^2 \leqslant C\gamma(P,\delta)^2 \sum_j \|\sigma^0_{P\delta}\Gamma_\delta X_j u\|^2 + C\|\rho_\delta \tilde{\Gamma}_\delta u\|^2 + \delta^{\text{power}}\|u\|^2,$$

Recalling the definition of $\phi_{P\delta}$, given by (33) in Section 2, we set $\phi_\delta = \phi_{\rho\delta}$ and replace u in (14) by $\phi_\delta u$ obtaining

$$(15) M^{2\epsilon}\|\sigma^0_\delta \Gamma_\delta \phi_\delta u\|^2 \leqslant C\gamma(\delta)^2 \sum_j \|\sigma^0_\delta \Gamma_\delta X_j \phi_\delta u\|^2 + C\|\rho_\delta \tilde{\Gamma}_\delta \phi_\delta u\|^2 + \delta^{\text{power}}\|u\|^2,$$

here we have dropped P, which remains fixed through the rest of this section.

Now observe that $\gamma(P,\delta) < C\delta^{1/m}$. Thus, introducing the coordinates $x^* = x^P\delta^{1/m-1}$ we note that the ϕ_δ are «smooth» in these coordinates, that is all their derivatives are bounded independently of δ. Similarly we define $\phi^1_\delta, \phi^2_\delta$, ϕ^3_δ etc., so that $\phi^{i+1}_\delta = 1$ on a neighborhood of the support of ϕ^i_δ and we set $\phi^0_\delta = \phi_\delta$. The coordinates ξ^* dual to x^* are given by $\xi^* = \xi\delta^{1-1/m}$ and hence the support of the symbol of Γ_δ lies in $\{|\xi^*| > C\delta^{-1/m}\}$. Since $\sigma^0_\delta = \sigma^0_\delta \phi^1_\delta$ we have

$$\sigma_\delta^0\Gamma_\delta u = \sigma_\delta^0\Gamma_\delta\phi_\delta u + \sigma_\delta^0[\phi_\delta^1\Gamma_\delta((1-\phi_\delta)u)]. \tag{16}$$

The L_2-norm of the second term on the right is $O(\delta^{\text{power}}\|u\|)$ since ϕ_δ^1 and $1-\phi_\delta$ have disjoint supports. Similarly

$$\sigma_\delta^0\Gamma_\delta X_j u = \sigma_\delta^0\Gamma_\delta X_j\phi_\delta u + \sigma_\delta^0[\phi_\delta^1\Gamma_\delta X_j(1-\phi_\delta)u] \tag{17}$$

and again the second term is $O(\delta^{\text{power}}\|u\|)$ since δX_j is a first order operator in the x^*-coordinates.

So now we have

$$M^{2\epsilon}\|\sigma_\delta^0\Gamma_\delta u\|^2 \leqslant C\gamma(\delta)^2\sum_j\|\sigma_\delta^0\Gamma_\delta X_j u\|^2 + C\|\rho_\delta\tilde{\Gamma}_\delta\phi_\delta u\|^2 + \delta^{\text{power}}\|u\|^2. \tag{18}$$

Reasoning as above we have

$$\|(1-\phi_\delta^1)\tilde{\Gamma}_\delta\phi_\delta u\|^2 \leqslant \delta^{\text{power}}\|u\|^2. \tag{19}$$

Next we have

$$\phi_\delta^1\tilde{\Gamma}_\delta\phi_\delta = Q\phi_\delta^2\tilde{\tilde{\Gamma}} + \delta^{\text{power}}Q_{\text{error}}, \tag{20}$$

where Q, Q_{error} are pseudo-differential operators of order zero in the x^*-coordinates and the symbol of Q is supported on $\{|\xi^*| \sim M\delta^{-1/m}\}$. This is due to the fact that $\phi_\delta^2\tilde{\tilde{\Gamma}}_\delta$ is elliptic on places where the symbol of $\phi_\delta^1\tilde{\Gamma}_\delta\phi_\delta$ is essentially supported. Thus

$$\|\rho_\delta\phi_\delta^1\tilde{\Gamma}_\delta\phi_\delta u\| \leqslant \|\rho_\delta Q\phi_\delta^2\tilde{\tilde{\Gamma}}_\delta u\| + \delta^{\text{power}}\|u\|. \tag{21}$$

In the x^*-coordinates, ρ_δ is constant on a scale $\sim\delta^{1/m}$ while Q is order zero with symbol supported in $\{|\xi^*| \sim M\delta^{-1/m}\}$. First we show that

$$\|\rho_\delta Q w\| \leqslant C\|\rho_\delta w\|. \tag{22}$$

The proof of (11) in Section 3 gives

$$Qw(x^*) = \int H(x^*, y^*)w(y^*)\,dy^*$$

with

$$|H(x^*, y^*)| \leqslant C_k(M\delta^{-1/m})^3(1 + |x^* - y^*|\delta^{-1/m}M)^{-K}.$$

From the definitions of ρ_δ and the x^*-coordinate system we read off

$$\frac{\rho_\delta(x^*)}{\rho_\delta(y^*)} \leqslant C(1 + |x^* - y^*|\delta^{-1/m})^N.$$

Hence

$$\rho_\delta Q w(x^*) = \int \frac{\rho_\delta(x^*)}{\rho_\delta(y^*)} H(x^*, y^*) \rho_\delta w(y^*)\, dy^*.$$

Taking K much larger than N we obtain (22) from the above.

Now we apply (22) to $w = \phi_\delta^2 \tilde{\tilde{\Gamma}}_\delta u$ and combining with (21) we obtain

$$\|\rho_\delta \tilde{\Gamma}_\delta \phi_\delta u\| \leqslant C \|\rho_\delta \phi_\delta^2 \tilde{\tilde{\Gamma}}_\delta u\| + \delta^{\text{power}} \|u\|. \tag{23}$$

From the definition of σ_δ we have $|\rho_\delta \phi_\delta^2| \leqslant \text{const.}\, |\sigma_\delta|$. Hence, combining with (18), we obtain the desired estimate (1) with $\tilde{\tilde{\Gamma}}_\delta$ instead of $\tilde{\Gamma}_\delta$ so that, changing notation, (1) is proved.

Now we turn to the proof of (2). The first step is to construct a function $\phi \in C_0^\infty(\{y \in \mathbb{R}^3 : |y| < 1\})$ with the property that $\hat{\phi} = 1$ to high order at 0, *i.e.* $\hat{\phi}(0) = 1$ and $D^\alpha \hat{\phi}(0) = 0$ for $|\alpha| \leqslant N$. Throughout this proof we will keep P fixed. Then for each γ we define $\delta(\gamma)$ by

$$\delta(\gamma) = \sum_{j=2}^{m} \lambda_j^P \gamma^j, \tag{24}$$

which is consistent with Definition 2.7. We define ϕ_γ by

$$\phi_\gamma(x_0, x') = \frac{1}{\delta(\gamma)\gamma^2} \phi\left(\frac{x_0}{\delta(\gamma)}, \frac{x'}{\gamma}\right), \tag{25}$$

here (x_0, x') denotes $(x_0^P, (x^P)')$.

Lemma 4.2. *Setting*

$$B\left(\frac{\gamma}{M}\right) = B\left(P, \delta\left(\frac{\gamma(P)}{M}\right)\right)$$

we have

$$|u^* \phi_{2\gamma/M}(0) - u^* \phi_{\gamma/M}(0)| \operatorname{vol} B\left(\frac{\gamma}{M}\right) \leqslant C\gamma \sum \|X_j u\|_{B(\gamma)} + C_s M^{-s} \|u\|_{B(2\gamma)} \tag{26}$$

and

$$\|u - u^* \phi_{\gamma/M}(0)\|_{B(\gamma)} \leqslant C\gamma \sum \|X_j u\|_{B(\gamma)} + C_s M^{-s} \|u\|_{B(2\gamma)}. \tag{27}$$

PROOF. Let Y_j be smooth vector fields given by $Y_j = \gamma X_j$. Given λ_0, λ', $\tilde{\lambda}_0$, $\tilde{\lambda}'$ we define

$$\psi(y_0, y') = \frac{1}{\lambda_0 |\lambda'|^2} \phi\left(\frac{y_0}{\lambda_0}, \frac{y'}{\lambda'}\right)$$

and

$$\tilde{\psi}(y_0, y') = \frac{1}{\tilde{\lambda}_0 |\tilde{\lambda}'|^2} \phi\left(\frac{y_0}{\tilde{\lambda}_0}, \frac{y'}{\lambda'}\right).$$

If $M >> 1$ and if the λ and $\tilde{\lambda}$ lie in $[M^{-K}, M^{-1}]$ with K large then

$$|\psi^* u(0) - \tilde{\psi}^* u(0)|^2 \leqslant C \sum \| Y_j u \|^2_{\{|y|<1\}} + C_s M^{-s} \| u \|^2_{\{|y|<2\}}. \tag{28}$$

PROOF OF (28). Define $\widehat{\Omega u}$ by $\Omega u(\xi) = w(M^{-1/2}\xi) u(\xi)$ with $w \in C^\infty(\mathbb{R}^3)$, $w = 1$ if $|\xi| > 1$, $w = 0$ if $|\xi| < 1/2$. We begin by applying (11) to Ωv. (Note that $\widehat{\Omega v}$ is supported on $|\xi| > M/2$ so (11) applies with $M^{1/2}$ in place of M.) Thus

$$M^\epsilon \| \zeta \Omega v \|^2 \leqslant C \sum_j \| \zeta Y_j \Omega v \|^2 + C \| v \|^2 \leqslant C \sum_j \| \Omega \zeta Y_j v \|^2 + C \| v \|^2. \tag{28}'$$

Apply this to $v = \tilde{\zeta} \tilde{\Omega} u$, where $\tilde{\zeta}$ and $\tilde{\Omega}$ are analogous to $\zeta \Omega$ but symbol of $\tilde{\zeta}\tilde{\Omega} = 1$ on the support of the symbol of ζQ. Lemma 3.2 with $\delta = 1$ and M replaced by $M^{1/2}$, gives

$$\zeta \Omega v = \zeta \Omega \tilde{\zeta} \tilde{\Omega} u = \zeta \Omega u + \mathcal{E} u \quad \text{with} \quad \| \mathcal{E} u \| \leqslant C_s M^{-s} \| u \|.$$

Similarly,

$$\Omega \zeta Y_j v = \Omega \zeta Y_j \tilde{\zeta} \tilde{\Omega} u = \Omega \zeta Y_j u + \mathcal{E}' u \quad \text{with} \quad \mathcal{E}' u \leqslant C_s M^{-s} \| u \|.$$

Hence (28)′ becomes

$$\begin{aligned} M^\epsilon \| \zeta \Omega u \|^2 &\leqslant C \sum_j \| \Omega \zeta Y_j u \|^2 + C \| \tilde{\zeta} \tilde{\Omega} u \|^2 + C_s M^{-s} \| u \|^2 \\ &\leqslant C \sum_j \| \zeta Y_j u \|^2 + C \| \tilde{\zeta} \tilde{\Omega} u \|^2 + C_s M^{-s} \| u \|^2. \end{aligned}$$

The term $\| \tilde{\zeta} \tilde{\Omega} u \|^2$ is analogous to the left-hand side, but we have gained a factor M^ϵ. Hence an obvious induction gives

$$M^\epsilon \| \zeta \Omega u \|^2 \leqslant C \sum_j \| Y_j u \|^2_{L^2\{|x|<1\}} + C_s M^{-s} \| u \|^2. \tag{29}$$

Since $\psi^* \Omega u(0) - \tilde{\psi}^* \Omega u(0) = (\psi - \tilde{\psi})^* \zeta \Omega u(0)$, (29) implies

$$|\psi^* \Omega u(0) - \tilde{\psi}^* \Omega u(0)|^2 \leqslant C \sum_j \| \zeta Y_j u \|^2 + C_s M^{-s} \| u \|^2. \tag{30}$$

The condition $\hat{\phi} = 1$ to high order at 0 implies that the Fourier transform of $\psi - \tilde{\psi}$ has L^2-norm $O(M^{-s})$ on $G = \operatorname{supp}(1 - w(M^{-1/2}|\xi|))$. Hence

$$|\psi^*(I-\Omega)u(0) - \tilde{\psi}^*(I-\Omega)u(0)|^2 \leqslant \|\hat{\psi} - \hat{\tilde{\psi}}\|^2_{L^2(G)} \|(I-\Omega)u\|^2 \leqslant C_s M^{-s}\|u\|^2. \tag{31}$$

Combining (30) and (31), we get

$$|\psi^*u(0) - \tilde{\psi}^*u(0)|^2 \leqslant C\sum_j \|Y_j u\|^2_{L^2\{|x|<1\}} + C_s M^{-s}\|u\|^2.$$

Replacing u by the product of u and the characteristic function of $\{x: |x| < 2\}$), we obtain (28).

Setting

$$\lambda_0 = \delta\left(\frac{\gamma}{M}\right)\delta(\gamma)^{-1}, \qquad \lambda' = (M^{-1}, M^{-1}), \qquad \tilde{\lambda}_0 = \delta\left(\frac{2\gamma}{M}\right)\delta(2\gamma)^{-1},$$

and $\tilde{\lambda}' = (2M^{-1}, 2M^{-1})$ we have

$$\psi(y_0, y') = \phi_{\gamma/M}(x_0, x')\,\mathrm{vol}\,B\left(\frac{\gamma}{M}\right). \tag{32}$$

$$\tilde{\psi}(y_0, y') = \phi_{2\gamma/M}(x_0, x')\,\mathrm{vol}\,B\left(\frac{2\gamma}{M}\right).$$

Thus we obtain (26) by rescaling (28).

The proof of (27) is analogous to the proof of (26). In place of (28) we have

$$\|u - u^*\psi(0)\|_{\{|y|<1\}} \leqslant C\sum_j \|Y_j u\|_{\{|y|<1\}} + C_s M^{-s}\|u\|. \tag{33}$$

This is proved by again setting $u = \Omega u + (I-\Omega)u$. We then obtain (27) by rescaling (33) which concludes the proof of the lemma.

Let $A = \sup_{0<\gamma<1} \gamma^{-2}\|u\|^2_{B(\gamma)}$. Then from (26) we get, for $2^{i+1}\gamma M^{-1} < 1$, the estimate

$$\begin{aligned}|u^*\phi_{2^{i+1}\gamma/M}(0) - u^*\phi_{2^i\gamma/M}(0)|^2 &\leqslant C\delta(2^i\gamma)^{-1}\sum \|X_j u\|^2 \\ &\quad + C_s M^{-s}\delta(2^i\gamma)^{-1}(2^i\gamma)^{-2}\|u\|^2_{B(2^i\gamma)} \\ &\leqslant C\delta(2^i\gamma)^{-1}\sum \|X_j u\|^2 + C_s M^{-s}\delta(2^i\gamma)^{-1}A.\end{aligned} \tag{34}$$

Since $\delta(2^i\gamma)^{-1} \leqslant 4^{-i}\delta(\gamma)^{-1}$, we obtain, by summing over i,

$$|u^*\phi_{\gamma/M}(0)|^2 = \delta(\gamma)^{-1}\left(C\sum \|X_j u\|^2 + C_s M^{-s}A\right) + C_M\|u\|^2. \tag{35}$$

Then

$$\begin{aligned}|u^*\phi_{\gamma/M}(0)|^2\,\mathrm{vol}\,B\left(\frac{\gamma}{M}\right) &\leqslant \gamma^2\left(C\sum \|X_j u\|^2 + C_s M^{-s}A\right) \\ &\quad + C_M\delta(\gamma)\gamma^2\|u\|^2.\end{aligned} \tag{36}$$

Combining this with (27) we obtain

$$A \leqslant C\Big(\sum_j \|\ddot{X}_j u\|^2 + \|u\|^2\Big) + C_s M^{-s} A$$

so that (2) follows by taking M sufficiently large.

5. Scaled Sobolev Norm

Definition 5.1. *For each $s \geqslant 0$ and $u \in C_0^\infty(U)$ we define the scaled Sobolev norm of u, denoted by $|||u|||_s$, by*

$$|||u|||_s^2 = \int_0^1 \int_{\mathbb{R}^3} \gamma(P,\delta)^{-2s} |\Gamma_\delta u(P)|^2 \, dP \frac{d\delta}{\delta} + \|u\|^2, \tag{1}$$

where Γ_δ is defined in terms of ψ in Definition 3.1 *and $\psi(\xi) = 1$ when $a < a' < |\xi| < b' < b$.*

The purpose of this section is to prove that for $M >> 1$

$$M^\epsilon |||u|||_{s+1} \leqslant C_1 \sum |||X_j u|||_s + C\|u\|, \tag{2}$$

where C_1 is independent of M.

Note that $|||u|||_0^{\cdot} \sim \|u\|$ since

$$c' \ln(b'/a) \leqslant \int_0^1 \left|\psi\Big(\frac{\delta}{M}\xi\Big)\right|^2 \frac{d\delta}{\delta} \leqslant c \ln(b/a).$$

Also, if $\gamma(P,\delta)$ is replaced by δ, the resulting norm is equivalent to $\|u\|_s$.

Lemma 5.2. *The following are equivalent to $|||u|||_s^2$*

$$|||u|||_s^2 \sim \|u\|^2 + \int_0^1 \int_{\mathbb{R}^3} \gamma(P,\delta)^{-2s} \|\sigma^0_{P\delta} \Gamma_\delta u\|^2 \frac{dP}{\mathrm{vol}(P,\delta)} \frac{d\delta}{\delta} \tag{3}$$

and

$$|||u|||_s^2 \sim \|u\|^2 + \int_0^1 \int_{\mathbb{R}^3} \gamma(P,\delta)^{-2s} \|\sigma_{P\delta} \Gamma_\delta u\|^2 \frac{dP}{\mathrm{vol}(P,\delta)} \frac{d\delta}{\delta}, \tag{4}$$

where $\mathrm{vol}(P,\delta) = c\gamma(P,\delta)^2\delta$, which is the volume of $B(P,\delta)$.

PROOF. Since $\gamma(Q,\delta) \sim \gamma(P,\delta)$ for $Q \in \tilde{B}(P,\delta)$ it suffices to show that

$$\int |\sigma^0_{P\delta}(Q)|^2 \frac{dP}{\mathrm{vol}(P,\delta)} \sim \int |\sigma_{P\delta}(Q)|^2 \frac{dP}{\mathrm{vol}(P,\delta)} \sim 1. \tag{5}$$

The fact that the first term is equivalent to 1 is an immediate consequence of Proposition 2.13. To see that the second term is equivalent to 1 we have, from (30) in Section 2, that

$$x_0^P(Q) = x_0^Q(P) - F^{Q,P}(-x_1^P(Q), -x_2^P(Q)),$$

furthermore

$$|F^{P,Q}(-x_1^P(Q), -x_2^P(Q))| \leqslant \text{const.}\,\delta,$$

for $Q \in \tilde{B}(P, \delta)$. Hence

$$\left(1 + \frac{|x_0^P(Q)|^2}{\delta^2}\right)^{-N} \sim \left(1 + \frac{|x_0^Q(P)|^2}{\delta^2}\right)^{-N}$$

for $Q \in \tilde{B}(P, d)$. So that (4) follows from this and the definition of $\sigma_{P\delta}$, which completes the proof.

The following proposition shows that the norms $|||\cdot|||_s$ are independent of the choice of ψ in the definition of Γ_δ. If Γ_δ is defined by the function ψ as in Definition 3.1 and if further there exist a_1, b_1 so that $\psi(\xi) = 1$ whenever $a_1 < |\xi| < b_1$ then we will denote the norm $|||u|||_s$, given by (1), by $|||u|||_s^\psi$. Similarly, if $\bar{\psi} \in C_0^\infty(\{\xi: 0 < \bar{a} < |\xi| < \bar{b}\})$ and if $\bar{\psi}(\xi) = 1$ when $\bar{a}_1 < |\xi| < \bar{b}_1$, we have the following result.

Proposition 5.3. *The norms* $|||u|||_s^\psi$ *and* $|||u|||_s^{\bar{\psi}}$ *are equivalent.*

PROOF. First assume that $\bar{\psi}$ is elliptic on the support of ψ, i.e. $|\bar{\psi}(\xi)| > \text{const.} > 0$ if $\xi \in \operatorname{supp} \psi$. Then $\Gamma_\delta = Q\bar{\Gamma}_\delta$, where

$$q(\xi) = \frac{\psi\left(\frac{\delta}{M}\xi\right)}{\bar{\psi}\left(\frac{\delta}{M}\xi\right)}$$

is the symbol of Q then, from (3), (4) and Lemma 3.4, we conclude that $|||u|||_s^\psi \leqslant C|||u|||_s^{\bar{\psi}}$. Now we drop the assumption concerning the support of ψ and we take up the general case. Let

$$r = \frac{\bar{a}_1 + \bar{b}_1}{2\bar{a}_1},$$

let k_0 be an integer so that $\bar{a}_1 r^{-k_0} < a$ and $\bar{b}_1 r^{k_0} \geqslant b$, $\bar{\psi}_j(\xi) = \bar{\psi}(r^j\xi)$, and let

$$\phi = \sum_{-k_0}^{k} \psi_j.$$

The function ϕ is bounded away from zero on the support of ψ and hence we have

$$|||u|||_s^{\psi} \leqslant C|||u|||_s^{\phi} \leqslant C \sum_{-k_0}^{k} |||u|||_s^{\bar{\psi}_j}.$$

Now, changing variables in (1), we get

$$(|||u|||_s^{\bar{\psi}_j})^2 = \int_0^{r^j} \int_{\mathbb{R}^3} \gamma(P, r^{-j}\delta)^{-2s} |\Gamma_\delta^{\bar{\psi}} u|^2 \, dP \frac{d\delta}{\delta} + \|u\|^2.$$

Since $\gamma(P, \delta) \leqslant \text{const.}\, \gamma(P, r^{-j}\delta)$ and since

$$\left| \int_1^{r^j} \int \gamma(P, r^{-j}\delta)^{-2s} |\Gamma_\sigma^{\psi} u|^2 \, dP \frac{d\delta}{\delta} \right| \leqslant C\|u\|^2,$$

we obtain $|||u|||_s^{\psi} \leqslant \text{const.}\, |||u|||_s^{\bar{\psi}}$. Reversing the roles of ψ and $\bar{\psi}$ we conclude that $|||u|||_s^{\psi} \sim |||u|||_s^{\bar{\psi}}$, thus completing the proof.

To prove the desired estimate (2) we proceed as follows. Multiplying (1) in Section 4 by $\gamma(P, \delta)^{-2s-2}$, integrating with respect to $dP/\text{vol}\,(P, \delta) \cdot d\delta/\delta$, and applying the above lemmas we obtain

$$M^{2\epsilon}|||u|||_{s+1}^2 \leqslant C \sum |||X_j u|||_s^2 + C|||u|||_s^2 + C\|u\|^2.$$

Then (2) follows by choosing M large enough.

6. The Basic Estimate

The purpose of this section is to establish the following result.

Theorem 6.1. *Let $\mathfrak{M}$ be a pseudo-convex compact CR manifold of dimension 3 on which $\bar{\partial}_b$ has closed range in L_2 and suppose that $P_0 \in \mathfrak{M}$ is of finite type. Then there exists a neighborhood U of P_0 such that if $\zeta \in C_0^\infty(U)$ then*

$$\sum_1^2 \|X_j(\zeta u)\| \leqslant C\|\bar{L}u\| \tag{1}$$

for all $u \perp \mathcal{H}(\mathfrak{M})$.

PROOF. The estimate (1) is equivalent to

$$\|L(\zeta u)\| \leqslant C\|\bar{L}u\|. \tag{2}$$

The following estimates follow from (1) and (7) in Section 1 and the fact that the range of $\bar{\partial}_b$ is closed.

$$\|L\mathcal{P}^- \zeta u\| + \|L\mathcal{P}^0 \zeta u\| \leqslant C\|\bar{L}u\|. \tag{3}$$

Thus it will suffice to prove

$$\|L\mathcal{P}^+ \zeta u\| \leqslant C\|\bar{L}u\|. \tag{4}$$

We have $u = \bar{L}^* v = -Lv + gv$. The estimates (4) will be proved in three steps. First, we will show that

$$M^\epsilon |||\mathcal{P}^+ \zeta v|||_1 \leqslant C(\|u\| + \|v\|) \tag{5}$$

and

$$M^\epsilon |||\mathcal{P}^+ \zeta v|||_2 \leqslant C(|||\mathcal{P}^+ \zeta u|||_1 + \|u\| + \|v\|). \tag{5}'$$

Second, we will prove that (5) implies that

$$|||\mathcal{P}^+ \zeta u|||_1 \leqslant C\|\bar{L}u\|. \tag{6}$$

Finally, we will prove that (6) implies (4).

We start with (6) in Section 1, namely

$$\sum \|X_j \mathcal{P}^+ w\|^2 \leqslant C\|\mathcal{P}^+ Lw\|^2 + C\|\mathcal{P}^+ w\|^2 + C\|\mathcal{R}w\|^2. \tag{7}$$

We wish to use this to estimate $\|\sigma^0_{P\delta} \Gamma_\delta X_j \mathcal{P}^+ \zeta v\|$, from Lemmas 3.1 and 3.4 we have

$$\|\sigma^0_{P\delta} \Gamma_\delta X_j \mathcal{P}^+ \zeta v\| \leqslant \|\sigma^0_{P\delta} \Gamma_\delta \mathcal{P}^+ X_j \zeta v\| + \|\sigma_{P\delta} \Gamma_\delta \tilde{\mathcal{P}}^+ \zeta v\| + C\delta^{\text{power}} \|v\|. \tag{8}$$

Letting $\tilde{\mathcal{P}}^+_\delta$ be an operator whose symbol is supported in a neighborhood of the symbol of $\Gamma_\delta \mathcal{P}^+$ and which is 1 on the support of $\Gamma_\delta \mathcal{P}^+$, we choose $\tilde{\mathcal{P}}^+_\delta$, Γ_δ and $\mathcal{P}^+$ so that their symbols depend only on ξ hence we have $\Gamma_\delta \mathcal{P}^+ = \tilde{\mathcal{P}}^+_\delta \Gamma_\delta \mathcal{P}^+$ and thus by Lemma 3.5

$$\|\sigma^0_{P\delta} \Gamma_\delta \mathcal{P}^+ X_j \zeta v\| \leqslant \|\tilde{\mathcal{P}}^+_\delta \sigma^0_{P\delta} \Gamma_\delta \mathcal{P}^+ X_j \zeta v\| + \frac{C}{M} \|\sigma_{P\delta} \Gamma_\delta X_j \mathcal{P}^+ \zeta v\| + C'\delta^{\text{power}} \|v\|. \tag{9}$$

The first term on the right above can be estimated by

$$\|\tilde{\mathcal{P}}^+_\delta \sigma^0_{P\delta} \Gamma_\delta \mathcal{P}^+ X_j \zeta v\| \leqslant \|X_j \tilde{\mathcal{P}}^+_\delta \sigma^0_{P\delta} \Gamma_\delta \mathcal{P}^+ \zeta v\| + C\|\sigma_{P\delta} \tilde{\Gamma}_\delta \tilde{\mathcal{P}}^+ \zeta v\| + \frac{C}{\gamma(P,\delta)} \|\sigma_{P\delta} \Gamma_\delta \mathcal{P}^+ \zeta v\| + C\delta^{\text{power}} \|v\|. \tag{10}$$

Now, applying (7), with $\mathcal{P}^+$ replaced by $\tilde{\mathcal{P}}_\delta^+$, to the first term on the right we obtain

$$(11)\quad \sum \|X_j \tilde{\mathcal{P}}_\delta^+ \sigma_{P\delta}^0 \Gamma_\delta \mathcal{P}^+ \zeta v\| \leqslant C\|\tilde{\mathcal{P}}_\delta^+ L \sigma_{P\delta}^0 \Gamma_\delta \mathcal{P}^+ \zeta v\| + C\|\sigma_{P\delta}^0 \Gamma_\delta \mathcal{P}^+ \zeta v\|.$$

The first term on the right in (11) is estimated by

$$(12)\quad \|\mathcal{P}_\delta^+ L \sigma_{P\delta}^0 \Gamma_\delta \mathcal{P}^+ \zeta v\| \leqslant C\|\sigma_{P\delta}^0 \Gamma_\delta \mathcal{P}^+ \zeta \bar{L}^* v\| + \frac{C}{\gamma(P,\delta)} \|\sigma_{P\delta} \Gamma_\delta \mathcal{P}^+ \zeta v\| + C\|\sigma_{P\delta} \tilde{\Gamma}_\delta \tilde{\mathcal{P}}^+ \tilde{\zeta} v\| + C\delta^{\text{power}} \|v\|.$$

Combining all these we obtain

$$(13)\quad \sum \|\sigma_{P\delta}^0 \Gamma_\delta X_j \mathcal{P}^+ \zeta v\|^2 \leqslant C\|\sigma_{P\delta}^0 \Gamma_\delta \mathcal{P}^+ \zeta \bar{L}^* v\|^2 + \frac{C}{M^2} \sum \|\sigma_{P\delta} \tilde{\Gamma}_\delta X_j \mathcal{P}^+ \zeta v\|^2 + \frac{C}{\gamma(P,\delta)^2} \|\sigma_{P\delta} \Gamma_\delta \mathcal{P}^+ \zeta v\|^2 + C\|\sigma_{P\delta} \tilde{\Gamma}_\delta \tilde{\mathcal{P}}^+ \tilde{\zeta} v\|^2 + C\delta^{\text{power}} \|v\|^2.$$

Multiplying by $\dfrac{dP}{\text{vol}\,(P,\delta)} \dfrac{d\delta}{\delta}$ and integrating we obtain

$$\sum \|X_j \mathcal{P}^+ \zeta v\|^2 \leqslant C\|\mathcal{P}^+ \zeta \bar{L}^* v\|^2 + \frac{C}{M^2} \sum \|X_j \mathcal{P}^+ \zeta v\|^2 + C|||\mathcal{P}^+ \zeta v|||_1^2 + C\|v\|^2$$

hence

$$\sum \|X_j \mathcal{P}^+ \zeta v\| < C(\|u\| + \|v\|) + C|||\mathcal{P}^+ \zeta v|||_1.$$

Applying (2) in Section 5 with $s = 0$ and M sufficiently large we obtain (5). Similarly, multiplying (13) by

$$\gamma(P,\delta)^{-2} \frac{dP}{\text{vol}\,(P,\delta)} \frac{d\delta}{\delta}$$

and integrating we obtain

$$\sum |||X_j \mathcal{P}^+ \zeta v|||_1 \leqslant C|||\mathcal{P}^+ \zeta u|||_1 + C|||\mathcal{P}^+ \zeta v|||_2 + C|||\tilde{\mathcal{P}}^+ \tilde{\zeta} v|||_1 + C\|v\|$$

now applying (2) in Section 5 with $s = 1$ and (5) with $\tilde{\mathcal{P}}^+ \tilde{\zeta}$ in place of $\mathcal{P}^+ \zeta$ we get (5)′.

Next, we take up the proof of (6), we have

$$\begin{aligned}\|\sigma^0_{P\delta}\Gamma_\delta \mathcal{P}^+ \zeta u\|^2 &= (\sigma^0_{P\delta}\Gamma_\delta \mathcal{P}^+ \zeta u, \sigma^0_{P\delta}\Gamma_\delta \mathcal{P}^+ \zeta \bar{L}^* v)\\ &\leqslant |(\sigma^0_{P\delta}\Gamma_\delta \mathcal{P}^+ \zeta \bar{L} u, \sigma^0_{P\delta}\Gamma_\delta \mathcal{P}^+ \zeta v)|\\ &\quad + \frac{C}{\gamma(P,\delta)} \|\sigma_{P\delta}\Gamma_\delta \mathcal{P}^+ \zeta u\| \, \|\sigma_{P\delta}\Gamma_\delta \mathcal{P}^+ \zeta v\|\\ &\quad + C\|\sigma^0_{P\delta}\Gamma_\delta \mathcal{P}^+ \zeta u\| \, \|\sigma^0_{P\delta}\tilde{\Gamma}_\delta \tilde{\mathcal{P}}^+ \tilde{\zeta} v\|\\ &\quad + C\|\sigma_{P\delta}\tilde{\Gamma}_\delta \tilde{\mathcal{P}}^+ \tilde{\zeta} u\| \, \|\sigma^0_{P\delta}\Gamma_\delta \mathcal{P}^+ \zeta v\|\\ &\quad + C\delta^{\text{power}}(\|u\|^2 + \|v\|^2)\end{aligned}$$

hence

$$\begin{aligned}\|\sigma^0_{P\delta}\Gamma_\delta \mathcal{P}^+ \zeta u\|^2 &\leqslant C\gamma(P,\delta)^2 \|\sigma^0_{P\delta}\Gamma_\delta \mathcal{P}^+ \zeta \bar{L} u\|^2\\ &\quad + \frac{C}{\gamma(P,\delta)^2} \|\sigma_{P\delta}\Gamma_\delta \mathcal{P}^+ \zeta v\|^2\\ &\quad + \text{small const.} \, \|\sigma_{P\delta}\Gamma_\delta \mathcal{P}^+ \zeta u\|^2\\ &\quad + C\|\sigma^0_{P\delta}\tilde{\Gamma}_\delta \tilde{\mathcal{P}}^+ \zeta v\|^2\\ &\quad + C\gamma(P,\delta)^2 \|\sigma^0_{P\delta}\tilde{\Gamma}_\delta \tilde{\mathcal{P}}^+ \tilde{\zeta} u\|^2\\ &\quad + C\delta^{\text{power}}(\|u\|^2 + \|v\|^2).\end{aligned}$$

Again, multiplying by $\gamma(P,\delta)^{-2} \dfrac{dP}{\text{vol}\,(P,\delta)} \dfrac{d\delta}{\delta}$ and integrating, we obtain

$$\begin{aligned}|||\mathcal{P}^+ \zeta u|||_1^2 &\leqslant C\|\mathcal{P}^+ \zeta \bar{L} u\|^2 + C|||\mathcal{P}^+ \zeta v|||_2^2 + \text{small const.} \, |||\mathcal{P}^+ \zeta u|||_1^2\\ &\quad + C|||\tilde{\mathcal{P}}^+ \tilde{\zeta} v|||_1^2 + C(\|u\|^2 + \|v\|^2).\end{aligned}$$

Then from (5) and (5)′ we get, after choosing M sufficiently large

$$|||\mathcal{P}^+ \zeta u|||_1^2 \leqslant C\|\bar{L} u\|^2 + C(\|u\|^2 + \|v\|^2).$$

The desired inequality (6) then follows from the closed range property.

We are now in a position to finish the proof of the theorem by proving (4). We have

$$\begin{aligned}\|\sigma^0_{P\delta}\Gamma_\delta L \mathcal{P}^+ \zeta u\|^2 &\leqslant \|\sigma_{P\delta}\tilde{\Gamma}_\delta \tilde{\mathcal{P}}^+ \tilde{\zeta} \bar{L} u\|^2\\ &\quad + \frac{C}{\gamma(P,\delta)^2} \|\sigma_{P\delta}\tilde{\Gamma}_\delta \tilde{\mathcal{P}}^+ \tilde{\zeta} u\|^2\\ &\quad + C\gamma(P,\delta)^2 \|\sigma^0_{P\delta}\Gamma_\delta \theta T \mathcal{P}^+ \zeta u\|^2\\ &\quad + C\delta^{\text{power}} \|u\|^2.\end{aligned} \tag{11}$$

This estimate is obtained by again integrating by parts as follows

$$\begin{aligned}\|ALBu\|^2 &= (L^*ALBu, ABu) + (ALBu, [A, L]Bu)\\ &= (ALBL^*u, ABu) + ([L^*, ALB]u, ABu) + (ALBu, [A, L]Bu)\\ &= \|ABL^*u\|^2 + (A[L^*, L]Bu, ABu) + \cdots\end{aligned}$$

and using the estimate

$$|(A[L^*, L]Bu, ABu)| \leqslant \frac{1}{\gamma^2}\|ABu\|^2 + \gamma^2\|A\theta TBu\|^2 + \cdots$$

Since $\delta\Gamma_\delta T$ is of order zero and since

$$|\theta| \leqslant \frac{C\delta}{\gamma(P, \delta)^2}$$

we get

$$\|\sigma^0_{P\delta}\Gamma_\delta\theta T\mathcal{P}^+\zeta u\| \leqslant \frac{C}{\gamma(P, \delta)^2}\|\sigma_{P\delta}\tilde{\Gamma}_\delta\mathcal{P}^+\zeta u\| + C\delta^{\text{power}}\|u\|.$$

Substituting this into (11), multiplying by $\dfrac{dP}{\text{vol}\,(P, \delta)}\dfrac{d\delta}{\delta}$ and integrating we obtain

$$\|L\mathcal{P}^+\zeta u\|^2 \leqslant C\|\bar{L}u\|^2 + C|||\tilde{\mathcal{P}}^+\tilde{\zeta}u|||_1^2 + C\|u\|^2.$$

Hence, applying (6) we obtain (4) and conclude the proof of the theorem.

7. Estimates of Kernels

In this section we prove the main theorem stated in the introduction.

PROOF OF (A). Suppose that f is supported in $B(P, C\delta)$, and let $u = f - S_b f$. We know that $\bar{L}u = \bar{L}f$ and that there exist v and w so that $u = \bar{L}^*v$ and $v = \bar{L}w$. Then we have $\|u\| \leqslant \|f\|$ and

$$\|v\|_{B(P, C\delta)} \leqslant C'\gamma(P, \delta)\|u\| \leqslant C'\gamma(P, \delta)\|f\|. \tag{1}$$

Now we introduce the rescaled functions $\tilde{u}$, $\tilde{v}$, $\tilde{w}$, $\tilde{f}$ defined on $\{\tilde{x}: |\tilde{x}| < 10\}$, here we denote by $\tilde{x}$ the coordinates defined by (3) in Section 4 and we are setting $\tilde{x} = y$. We define

$$\Phi(\tilde{x}_0, \tilde{x}_1, \tilde{x}_2) = \left(\frac{\tilde{x}_0}{\delta}, \frac{\tilde{x}_1}{\gamma(P, \delta)}, \frac{\tilde{x}_2}{\gamma(P, \delta)}\right),$$

$$\tilde{f} = f \circ \Phi, \qquad \tilde{u} = u \circ \Phi, \qquad \tilde{v} = \gamma^{-1} v \circ \Phi, \qquad \tilde{w} = \gamma^{-1} w \circ \Phi.$$

Here we are dealing with fixed P and δ and set $\gamma = \gamma(P, \delta)$. Then we have $\tilde{\bar{L}}\tilde{u} = \tilde{\bar{L}}\tilde{f}$, $\tilde{u} = \tilde{\bar{L}}^* \tilde{v}$, $\tilde{v} = \tilde{\bar{L}}\tilde{w}$ with $\|\tilde{u}\|_{\{|\tilde{x}|<10\}} + \|\tilde{v}\|_{\{|\tilde{x}|<10\}} \leqslant C\|\tilde{f}\|$, where $\tilde{L} = \gamma(P, \delta)L$.

Recall the subelliptic estimates from Section 1, namely

$$\|\phi\tilde{u}\|_{s+\epsilon} \leqslant C\|\phi'\tilde{\bar{L}}\tilde{f}\|_s + C\|\phi'\tilde{u}\| + C\|\phi'\tilde{v}\|. \tag{2}$$

If we pick f so that $\operatorname{supp}\tilde{f} \cap \operatorname{supp}\phi' = \varnothing$, then the first term on the right is zero so we get

$$\|\phi\tilde{u}\|_{s+\epsilon} < C\|\tilde{f}\|. \tag{3}$$

Now $\tilde{u}(\tilde{x}) = \int \tilde{K}(\tilde{x}, \tilde{y})\tilde{f}(\tilde{y})\, d\tilde{y}$ where $\tilde{K}$ is the rescaled Szegö kernel $\tilde{K}(\tilde{x}, \tilde{y}) = \gamma^2 \delta K(\Phi(\tilde{x}), \Phi(\tilde{y}))$. We write, $\tilde{u} = \tilde{S}_b\tilde{f}$. We have shown that for U, U' contained in $\{\tilde{x}: |\tilde{x}| < 10\}$ with $\bar{U} \cap \bar{U}' = \varnothing$, we have

$$\|\tilde{S}f\|_{H^s(U)} \leqslant C\|\tilde{f}\|$$

for $\tilde{f}$ supported in U'. Now $\tilde{S}_b$ is self-adjoint, so we obtain

$$\|\tilde{S}\tilde{f}\|_U \leqslant C\|\tilde{f}\|_{H^{-s}(U')}$$

for $\tilde{f}$ supported in U', by taking duals and interchanging U with U'. Interpolating between the last two estimates, we obtain

$$\|\tilde{S}\tilde{f}\|_{H^{s/2}(U)} \leqslant C\|\tilde{f}\|_{H^{-s/2}(U')}.$$

That means that $\tilde{K} \in C^\infty(U \times U')$ uniformly in δ and P. Hence $|p(\tilde{L}, \tilde{\bar{L}})\tilde{K}| \leqslant C$ for any polynomial p and L, $\bar{L}$ can act either in $\tilde{x}$ or $\tilde{y}$. The desired inequality (2) then follows by reinterpreting this in terms of L, $\bar{L}$ and K.

PROOF OF (B). First we have

$$\|u\|_{B(P, C\delta)} \leqslant C'\gamma(P, \delta)\|f\|. \tag{4}$$

Let $\theta \in C_0^\infty(B(P, C\delta))$ be a function such that $\tilde{\theta} = \theta \circ \Phi \in C_0^\infty(\{x: |x| < 10\})$ uniformly in δ and P. Let $u^\# = \theta u - S(\theta u)$. Now

$$\bar{L}u^\# = (\bar{L}\theta)u + \theta(\bar{L}u),$$

hence if $U \subset B(P, C\delta)$ is a neighborhood on which $\theta = 1$ then $\bar{L}^*\bar{L}u^\# = \bar{L}^*\bar{L}u = \bar{L}^* f$ in U. Moreover $\|u^\#\| \leqslant \|\theta u\| \leqslant C'\gamma(P, \delta)\|f\|$. Since $u^\# \perp \mathcal{H}_b$ we have $u^\# = \bar{L}^* v^\#$. Choosing $v^\#$ orthogonal to $\mathfrak{N}(\bar{L}^*)$ we have

$$\|v^\#\|_U \leqslant C\gamma(P, \delta)\|u^\#\| \leqslant C\gamma(P, \delta)^2\|f\|. \tag{5}$$

Now we rescale f, u, $u^{\#}$, $v^{\#}$ by setting $\tilde{f} = f \circ \Phi$, $\tilde{u} = \gamma(P,\delta)^{-1} u \circ \Phi$, $\tilde{u}^{\#} = \gamma(P,\delta)^{-1} u^{\#} \circ \Phi$, and $\tilde{v}^{\#} = \gamma(P,\delta)^{-2} v^{\#} \circ \Phi$. Then we have $\|\tilde{u}^{\#}\| \leqslant C\|\tilde{f}\|$, $\|\tilde{v}^{\#}\| \leqslant C\|\tilde{f}\|$, $\tilde{\tilde{L}}^* \tilde{\tilde{L}} \tilde{u}^{\#} = \tilde{\tilde{L}}^* \tilde{f}$, and $\tilde{L}^* \tilde{v}^{\#} = \tilde{u}^{\#}$ in $\tilde{U} = \Phi^{-1}(U)$, which we may take to be $\{x: |\tilde{x}| < 10\}$. Returning to the subelliptic estimate, we have from Proposition 1.5

$$\|\phi \tilde{u}^{\#}\|_{s+\epsilon} \leqslant C\|\phi' \tilde{f}\|_s + C\|\phi' \tilde{u}^{\#}\| + C\|\phi' \tilde{v}^{\#}\|.$$

Assuming that $\operatorname{supp} \tilde{f} \cap \operatorname{supp} \phi' = \varnothing$, the first term on the right vanishes, while the last terms are dominated by $C\|\tilde{f}\|$. Thus we obtain $\|\phi \tilde{u}^{\#}\|_{s+\epsilon} \leqslant C\|\tilde{f}\|$. What we really want is $\|\phi \tilde{u}\|_{s+\epsilon} \leqslant C\|\tilde{f}\|$, so we look at the difference $\tilde{u} - \tilde{u}^{\#}$ in $\operatorname{supp} \phi$.

Recall that $u^{\#} = \theta u - S_b(\theta u)$ and that $S_b u = 0$. Hence in $\operatorname{supp} \theta$ we have $u^{\#} = u + S_b((1-\theta)u)$, that is for $x \in \operatorname{supp} \theta$

$$(u^{\#} - u)(x) = \int K(x,y)(1 - \theta(y))u(y)\, dy. \tag{6}$$

If $p_k(L, \bar{L})$ is a monomial of order k in L, $\bar{L}$, then for $x \in \operatorname{supp} \theta$

$$P_k(L,L)(u^{\#} - u)(x) = \int \{p_k(L,\bar{L})K(x,y)\}(1 - \theta(y))u(y)\, dy,$$

where the $p_k(L,\bar{L})$ acts on the x-variable. Since $1 - \theta(y) = 0$ in $B(P, C\delta)$ we may break up the above integral as

$$p_k(L,\bar{L})(u^{\#} - u)(x) = \sum_{i=-10}^{\infty} \int_{y \in B(P,2^i\delta)\setminus B(P,2^{i-1}\delta)} |p_k(L,\bar{L})K(x,y)|\, |u(y)|\, dy. \tag{7}$$

In the region of integration for fixed i we have

$$|p_k(L,\bar{L})K(x,y)| \leqslant \frac{C\gamma(P,2^i\delta)^{-k}}{2^i \delta \gamma(P,2^i\delta)^2}. \tag{8}$$

Also we have

$$\|u\|_{B(P,2^i\delta)} \leqslant C\gamma(P,2^i\delta)\|f\|$$

so by Cauchy-Schwarz

$$\int_{B(P,2^i\delta)} |u(y)|\, dy \leqslant C\gamma(P,2^i\delta)[\gamma(P,2^i\delta)^2 2^i\delta]^{1/2}\|f\|. \tag{9}$$

Inserting (9) and (8) in (7) and recalling that $\gamma(P,2^i\delta) > 2^{i/2}\gamma(P,\delta)$ we obtain

$$|p_k(L,\bar{L})(u^{\#} - u)(x)| \leqslant C\gamma(P,\delta)^{-k}\delta^{-1/2}\|f\|$$

for $x \in \operatorname{supp} \theta$.

Rewriting this estimate in terms of the rescaled quantities we get

$$|p_k(\tilde{L}, \tilde{\bar{L}})\tilde{u}^{\#} - \tilde{u})| \leqslant C' \|f\|$$

for $|\tilde{x}| < 10$.

Since $\partial/\partial\tilde{x}_j$ is a bounded linear combination of the $p_k(\tilde{L}, \tilde{\bar{L}})$ with $k \leqslant m$ in $\{\tilde{x}: |\tilde{x}| < 10\}$ we conclude that for any s

$$\|\phi(\tilde{u} - \tilde{u}^{\#})\|_{H^s} \leqslant C'' \|\tilde{f}\|.$$

Combining this with our previous estimate $\|\phi\tilde{u}^{\#}\|_{s+\epsilon} \leqslant c\|\tilde{f}\|$, we find that

$$\|\tilde{u}\|_{H^s(V)} \leqslant C'' \|\tilde{f}\| \tag{10}$$

if $V, V' \subset \{\tilde{x}: |\tilde{x}| < 10\}$ are neighborhoods with disjoint closures with $\operatorname{supp}\tilde{f} \subset V'$.

Denoting the operator that takes f to u by Nf, that is $u = Nf$, we observe that there is a completely analogous estimate for N^* by interchanging the roles of $\bar{L}$ and $\bar{L}^*$. Then by duality and interchanging of V and V', we get

$$\|\tilde{u}\|_V \leqslant C'' \|\tilde{f}\|_{H^{-s}(V')} \quad \text{where} \quad \operatorname{supp}\tilde{f} \subset V'. \tag{11}$$

Then from (10) and (11) we conclude that $\tilde{H} \in C^{\infty}(V \times V')$ where

$$\tilde{u}(\tilde{x}) = \int \tilde{H}(\tilde{x}, \tilde{y})\tilde{f}(\tilde{y})\, d\tilde{y}.$$

In particular, for a monomial $p_k(\tilde{L}, \tilde{\bar{L}})$ of degree k with each L acting either on $\tilde{x}$ or $\tilde{y}$, we get

$$|p_k(\tilde{L}, \tilde{\bar{L}})\tilde{H}(\tilde{x}, \tilde{y})| \leqslant C_k$$

for $\tilde{x} \in V$, $\tilde{y} \in V'$ with constants C_k independent of γ and P. Returning to the original coordinates we obtain the desired estimate (11) in the Introduction thus completing the proof.

PROOF OF (C). Define kernels $K(x, y)$, $\tilde{K}(x, y)$ to provide the relative fundamental solutions of $\bar{L}$, $\bar{L}^*$ respectively. Thus

$$u(x) = \int K(x, y) f(y)\, dy$$

solves $\bar{L}u = f$ modulo the nullspace of $\bar{L}^*$, and u is orthogonal to the nullspace of $\bar{L}$; while

$$\tilde{u}(x) = \int \tilde{K}(x, y) f(y)\, dy$$

solves $\bar{L}^*\tilde{u} = f$ modulo the nullspace of $\bar{L}$, and $\tilde{u}$ is orthogonal to the nullspace of $\bar{L}^*$. From K, $\tilde{K}$ we define composed kernels

$$K_2(x,z) = \int \tilde{K}(x,y)K(y,z)\,dy$$
$$K_3(x,z) = \int K(x,y)K_2(y,z)\,dy.$$

In particular, K_2 is the kernel for the relative fundamental solution of $\bar{L}^*\bar{L}$, so (C) in the Main Theorem amounts to proving estimates on K_2. First we estimate the size of K_2, K_3. Recall from (B) the estimates

$$|K(x,y)|, |\tilde{K}(x,y)| \leqslant C/\delta(x,y)\gamma(x,y).$$

Here $\gamma(x,y)$ is the least γ for which $y \in B(x,\gamma)$, while $\delta = \delta(x,\gamma)$. From these estimates follows easily

$$\int_{B(x,\gamma)} |\tilde{K}(x,y)|\,dy \leqslant C\gamma, \qquad \int_{B(z,\gamma)} |K(y,z)|\,dy \leqslant C\gamma.$$

To estimate K_2, we write $\gamma = \gamma(x,z)$ and $\delta(\gamma) = \delta(z,\gamma)$, then break up the region of integration in the definition of K_2. We get

$$\begin{aligned}
|K_2(x,z)| &\leqslant \int_{y\in B(x,\gamma/10)} |\tilde{K}(x,y)|\,\frac{C}{\gamma\delta(\gamma)}\,dy \\
&\quad + \int_{y\in B(z,2\gamma)\setminus B(x,\gamma/10)} \frac{C}{\gamma\delta(\gamma)}\,|K(y,z)|\,dy \\
&\quad + \sum_{i\geqslant 1} \int_{y\in B(z,2^{i+1}\gamma)\setminus B(z,2^i\gamma)} \left[\frac{C}{(2^i\gamma)\delta(2^i\gamma)}\right]^2 dy \\
&\leqslant \frac{C'}{\delta(\gamma)} + \sum_{i\geqslant 1} \frac{C}{\delta(2^i\gamma)} \leqslant \frac{C''}{\delta(\gamma)}.
\end{aligned}$$

Similarly, using our estimate for the size of K_2 in the definition of K_3,

$$\begin{aligned}
|K_3(x,z)| &\leqslant \int_{y\in B(x,\gamma/10)} |K(x,y)|\,\frac{C}{\delta(\gamma)}\,dy \\
&\quad + \int_{y\in B(z,2\gamma)\setminus B(x,\gamma/10)} \frac{C}{\gamma\delta(\gamma)}\,|K_2(y,z)|\,dy \\
&\quad + \sum_{i>1} \int_{y\in B(z,2^{i+1}\gamma)\setminus B(z,2^i\gamma)} \frac{C}{(2^i\gamma)\delta(2^i\gamma)}\,\frac{C}{\delta(2^i\gamma)}\,dy \\
&\leqslant \frac{C\gamma}{\delta(\gamma)} + \sum_{i>1} C'\frac{2^i\gamma}{\delta(2^i\gamma)} \leqslant \frac{C''\gamma}{\delta(\gamma)},
\end{aligned}$$

since (19) in Section 2 implies that $\delta(2^i\gamma) \geqslant 4^i\delta(\gamma)$.

Now suppose f is a function supported in $B(0, \gamma/10)$ and define functions

$$u_1(x) = \int K(x, y) f(y)\, dy,$$
$$u_2(x) = \int K_2(x, y) f(y)\, dy,$$
$$u_3(x) = \int K_3(x, y) f(y)\, dy.$$

On $B(0, \gamma)$ we rescale f, u_1, u_2, u_3 by setting $\tilde{f} = f \circ \Phi$, $\tilde{u}_1 = \gamma^{-1} u_1 \circ \Phi$, $\tilde{u}_2 = \gamma^{-2} u_2 \circ \Phi$, $\tilde{u}_3 = \gamma^{-3} u_3 \circ \Phi$. These functions are defined in $\{\tilde{x}: |\tilde{x}| < 1\}$. The definitions of our kernels show that $\bar{L}^* u_2 = u_1$, $\bar{L} u_3 = u_2$, hence $\tilde{\bar{L}}^* \tilde{u}_2 = \tilde{u}_1$, $\tilde{\bar{L}} \tilde{u}_3 = \tilde{u}_2$ in $\{\tilde{x}: |\tilde{x}| < 1\}$. Therefore, the subelliptic estimates of Section 1 give

$$\|\tilde{\phi} \tilde{u}_2\|_{s+\epsilon} \leqslant C \|\tilde{\phi} \tilde{u}_1\|_s + C \|\tilde{u}_2\|_{L^2\{|\tilde{x}| < 1\}} + C \|\tilde{u}_3\|_{L^2\{|\tilde{x}| < 1\}}$$

for $\tilde{\phi} \in C_0^\infty(\{|\tilde{x}| < 1\})$. Our estimates for the size of the kernels K_2, K_3 show easily that $\|\tilde{u}_2\|_{L^2\{|\tilde{x}| < 1\}}$, $\|\tilde{u}_3\|_{L^2\{|\tilde{x}| < 1\}} \leqslant C \|\tilde{f}\|$. Thus,

$$\|\tilde{\phi} \tilde{u}_2\|_{s+\epsilon} \leqslant C \|\tilde{\phi} \tilde{u}_1\|_s + C \|\tilde{f}\|.$$

In the proof of (B), we showed that $\|\tilde{\phi} \tilde{u}_1\|_s \leqslant C \|\tilde{f}\|$, provided $\tilde{\phi}$ is supported away from $\{\tilde{x}: |\tilde{x}| < 1/10\}$, where $\tilde{f}$ is supported.

So we get

$$(*) \qquad \|\tilde{\phi} \tilde{u}_2\|_{s+\epsilon} \leqslant C' \|\tilde{f}\|$$

for $\operatorname{supp}(f) \subset \{x: |x| < 1/10\}$, $\operatorname{supp}(\tilde{\phi})$ disjoint from $\{\tilde{x}: |\tilde{x}| < 1/10\}$. To interpret this, we introduce the rescaled operator $\tilde{G}_2$ and kernel $\tilde{K}_2(x, y)$, defined by

$$\tilde{G}_2 \tilde{f}(\tilde{x}) = \tilde{u}_2(\tilde{x}) = \int \tilde{K}_2(\tilde{x}, \tilde{y}) f(\tilde{y})\, d\tilde{y}.$$

For $x, y \in B(0, \gamma)$, $K_2(x, y)$ is related to $\tilde{K}_2(\tilde{x}, \tilde{y})$ by an obvious scaling. Estimate $(*)$ says that

$$(**) \qquad \tilde{G}_2 : L^2(\tilde{U}) \to H^{s+\epsilon}(\tilde{V})$$

for U, V disjoint neighborhoods in $\{\tilde{x}: |\tilde{x}| < 1\}$.

Note also that $\tilde{G}_2$ is self-adjoint, since $K_2(x, y)$ provides the relative fundamental solution of the self-adjoint operator $\bar{L}^* \bar{L}$. Hence by duality we deduce from $(**)$: $\tilde{G}_2 : H^{-s-\epsilon}(V) \to L^2(U)$. Reversing the roles of $\tilde{U}$ and $\tilde{V}$ and interpolating with $(**)$, we find that $\tilde{G}_2 : H^{-(s-\epsilon)/2}(\tilde{U}) \to H^{(s+\epsilon)/2}(\tilde{V})$. This means that the kernel $\tilde{K}_2(\tilde{x}, \tilde{y})$ is C^∞ away from the diagonal. After rescaling, we obtain the estimate (11) of the Introduction with $q = 2$ thus proving (C).

References

[BS] Boas, H. P. and Shaw, M. C. Sobolev estimates for the Lewy operator on weakly pseudo-convex boundaries, *Math. Ann.* **274**(1986), 221-231.

[C] Catlin, D. Estimates of invariant metrics on pseudo-convex domains of dimension two, *Math. Zeitschrift* (to appear).

[Ch. 1] Christ, M. Regularity properties of the $\bar{\partial}_b$ equation on weakly pseudo-convex *CR* manifolds of dimension 3, *Journal Amer. Math. Soc.* **1**(1988), 587-646.

[Ch. 2] Christ, M. Pointwise estimates for the relative fundamental solution of $\bar{\partial}_b$, *Proc. Amer. Math. Soc.*, **104**(1988), 787-792.

[Ch. 3] Christ, M. Embedding compact three-dimensional *CR* manifolds of finite type in $\mathbb{C}^n$, *Annals of Math.*, **129**(1989), 195-213.

[F] Fefferman, C. L. The Uncertainty Principle. *Bull. Amer. Math. Soc.*. **9**(1983), 129-206.

[FK] Fefferman, C. L. and Kohn, J. J. Hölder estimates on domain of complex dimension two and on three-dimensional *CR* manifolds. *Advances in Math.* **69**(1988), 223-303.

[FS] Fefferman, C. L. and Sánchez-Calle, A. Fundamental solutions for second order subelliptic operators, *Annals of Math.* **124**(1986), 247-272.

[H] Hormander, L. Hypoelliptic second order differential equations, *Acta Math.* **119**(1967), 147-171.

[Kl] Kohn, J. J. Estimates for $\bar{\partial}_b$ on pseudo-convex *CR* manifolds, *Proc. Symp. Pure Math.* **43**(1985), 207-217.

[K2] Kohn, J. J. The range of the tangential Cauchy-Riemann operator, *Duke Math. J.* **53**(1986), 525-545.

[M] McNeal, J. Boundary behavior of the Bergman kernel, *Duke Math. J.*, **58**(1989), 499-512.

[NRSW] Nagel, A., Rosay, J. P., Stein, E. M. and Wainger, S. Estimates for the Bergman and Szegö kernels in $\mathbb{C}^2$, *Annals of Math.* (to appear).

[NSW1] Nagel, A, Rosay, J. P., Stein, E. M. and Wainger, S. Boundary Behaviour of Functions Holomorphic in Domains of Finite Type. *Proc. Nat. Acad. Sci. U.S.A.* **78**(1981), 6596-99.

[NSW2] Nagel, A., Stein, E. M. and Wainger, S. Balls and metrics defined by vector fields I: Basic properties, *Acta Math.* **155**(1985),

[RS] Rothschild, L. P. and Stein, E. M. Hypoelliptic differential operators and nilpotent groups, *Acta Math.* **137**(1976), 247-320.

C. L. Fefferman* and J. J. Kohn*
Department of Mathematics,
Princeton University
Princeton N.J. 08544
U.S.A.

* Both authors were partially supported by NSF grants at Princeton University.

REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.^{os} 3 Y 4, 1988

Infinite Group Actions on Spheres

Gaven J. Martin

Introduction

This paper is mainly intended as a survey of the recent work of a number of authors concerning certain infinite group actions on spheres and to raise some as yet unanswered questions. The main thrust of the current research in this area has been to decide what topological and geometrical properties characterise the infinite conformal or Möbius groups. One should then obtain reasonable topological or geometric restrictions on a subgroup G of the homeomorphism group of a sphere which will imply that it can be made conformal after a change of coordinates. That is G is topologically conjugate to a Möbius group. Many aspects of the theory of Kleinian and Möbius groups can be found in the books [Ah 3], [Bea], [Mas 1], [MB] and [Th].

Another focus of an investigation into the topological nature of Möbius groups is that it helps us to better understand these groups and enables us to decide what features are essential to the development of certain aspects of the general theory. As a relatively simple question one may ask whether the classification into elliptic, parabolic and loxodromic elements for a conformal group is analytic or topological in nature? What properties determine the structure of the limit set? To gain deeper insight one is led to ask are the combination and decomposition theorems for Kleinian groups (a lá Maskit), either topological, geometric or analytic in nature? How does Ahlfors' finiteness theorem fit into the picture? Is it really topological in nature or does it definitely depend on the analytic structure of Kleinian groups (as does the proof). Perhaps it is essentially an algebraic statement about finitely generated Kleinian groups! Also what of Bers' Area Theorem and Sullivan's Finiteness of Cusps. Are these too more general phenomena? Is Selberg's theorem on the

existence of torsion free finite index subgroups of finitely generated linear groups a phenomena exhibited by any reasonable topological action on a sphere?

In this introduction we will outline the main ideas and some of their consequences, specific details of statements and some proofs appear later. I hope the careful reader (and especially the relevant authors) will forgive the simplifications I have had to make in the proofs that do appear.

A first natural generalisation of a group of conformal homeomorphisms of a sphere is that of a *quasiconformal group* of homeomorphisms. Such groups can be thought of as acting with a uniformly bounded amount of distortion, whereas conformal groups have no distortion (at least infinitesimally). We use J. Väisälä's book [Vä] as our standard reference for the theory of quasiconformal mappings. The following question was first asked by F. W. Gehring and B. Palka [GP] in connection with their work on quasiconformally homogeneous domains:

Is every uniformly quasiconformal group the quasiconformal conjugate of a conformal group?

The answer to this question is unknown in dimension one (where quasiconformal maps are more commonly called quasisymmetric maps), yes in dimension two and generally false in higher dimensions. These, together with some positive results, are mainly due to D. Sullivan [Su 1] and P. Tukia [Tu 1, 2, 3].

We regard quasiconformal groups as a geometric generalisation of conformal groups. A topological generalisation was proposed by Gehring and the author [GM] in their study of quasiconformal groups. The idea is to retain the normal families properties of conformal and quasiconformal mappings, as exhibited by the fact that any sequence of conformal homeomorphisms contains a subsequence whose limit is either a conformal homeomorphism or a constant mapping. This compactness property is clearly a necessary condition for the existence of a topological conjugacy and we call group of homeomorphisms of the n-sphere with this property a convergence group. We then ask.

When is a convergence group topologically conjugate to a conformal group?

Again the answer to this question is largely unknown although there are good positive results. In dimension one it is thought to be always the case while in higher dimensions some additional hypotheses are clearly necessary. We ask the question with regard to conformal groups since we are primarily interested in lower dimensional phenomena or cases where the natural geometric can-

didate for a possible conjugacy is a conformal group, for instance as in the case of topological analogues of Schottky groups. As a simple example, the group of complex biholomorphisms of the unit ball in $\mathbb{C}^n$, $n \geqslant 2$, acts as a convergence group on $\mathbb{S}^{2n-1}$ but by Margulis super-rigidity this group or any of its discrete cocompact subgroups cannot even be isomorphic to conformal groups as the enveloping Lie algebras are distinct. Actually these groups cannot even be topologically conjugate to quasiconformal groups. There are however quite reasonable hypotheses to make in a general setting which yield good answers to the above question.

There are many closely related questions to those raised above. For instance:

When does the group action on the n-sphere extend to the $(n+1)$*-ball*?

This question is motivated by the classical Poincaré extension which extends conformal actions on the n-sphere to the $(n+1)$-ball.

In dimension one this extension problem is equivalent to the general conjugacy problem, [MT]. And due to results of G. Mess [Me] and P. Scott [Sc], the general conjugacy problem is equivalent to the Seifert conjecture for three manifolds:

Is a compact three manifold with a normal infinite cyclic subgroup in its fundamental group is a Seifert fibered space?

Roughly a Seifert fibered space is a three manifold foliated by circles in a nice way. There are other interesting consequences related to Teichmüller theory. Indeed the one dimension question on the conjugacy of convergence groups to Fuchsian groups is essentially a universal version of the Nielsen Realisation Problem. The usual Nielsen Realisation Problem, solved by S. Kerchoff [Ker], is implied by the statement (see [Tu 3]):

A group of homeomorphisms acting on the circle with a Fuchsian subgroup of finite index is topologically conjugate to a Fuchsian group.

Necessarily a group of homeomorphisms with a Fuchsian subgroup of finite index is a convergence group. P. Tukia has shown that convergence groups of the circle are conjugate to Fuchsian groups unless they contain a semitriangle group of finite index [Tu 3]. Such groups are so special as to be presumed not to exist. In addition he and the author have shown that convergence groups of the circle which can be extended to the entire disk are conjugate to Fuchsian groups [MT].

In higher dimensions the problem is essentially more complicated. It is known that the fundamental group of a strictly negatively curved manifold acts as a convergence group on the sphere at infinity of the universal covering space [MS]. Thus reasonable affirmative solutions to the problem imply results concerning the homotopy type of negatively curved compact manifolds. This is especially interesting in the three dimensional case where it is thought that every negatively curved compact three manifold is hyperbolic. There is enough machinery in place, mainly due to Thurston, so that the answer to this question is essentially implied by the conjecture:

A convergence group acting on the two sphere and whose limit set is the whole two sphere is topologically conjugate to a Kleinian group?

For partial results along this line see [MS] and [MT]. There is a relationship here between these questions and those of Gromov concerning his so called hyperbolic groups [Gr. 2]. Also the fact that the fundamental group of a strictly negatively curved manifold acts as a convergence group on the sphere at infinity explains many of the geometric features of the isometry groups of visibility manifolds discovered by Eberlin and O'Neil in their work [EO].

M. Freedman has shown that in higher dimensions, the question of the existence of a solution to the extension problem for even a rather restrictive class of convergence groups is equivalent to the surgery problem [Fr]. In particular the four dimensional surgery problem, at present unsolved, is equivalent to the following question concerning a topological characterisation of certain Schottky groups acting on the three sphere.

Given a convergence group G acting on the three sphere, isomorphic to the free group on a finite number of generators, with limit set a Cantor set and which is of compact type, does G extend to a convergence group of the four ball?

These questions are all very interesting, however the simplicity of their statements often belies the fact that they are mostly quite difficult. But as we will see even partial solutions are often revealing and unify many approaches to varied problems.

Finally we remark that the theory of finite group actions on spheres is subsumed in trying to understand general convergence groups. Fortunately this theory is well developed but there are still many unanswered questions which are rather difficult and deep, for a good survey (particularly the three dimensional aspects) see Morgan and Bass [MB] and the references therein. We will try to concentrate on the aspects of the theory which specifically relate to the group being infinite.

1. Notation and Definitions

We denote by Hom $(\mathbb{S}^n)$ the homeomorphism group of the unit sphere $\mathbb{S}^n$ of euclidean $(n+1)$-space, $\mathbb{R}^{n+1}$. The topology of Hom $(\mathbb{S}^n)$ is the usual compact open topology, which is the same as that of the topology of uniform convergence in the *chordal distance* $\rho(x, y)$ that $\mathbb{S}^n$ inherits as a metric subspace of $\mathbb{R}^{n+1}$. A self homeomorphism f of $\mathbb{S}^n$ is said to be *K-quasiconformal*, $1 \leqslant K < \infty$, if for each $x \in \mathbb{S}^n$

$$\limsup_{r \to 0} \frac{\mathfrak{L}(x, r)}{\ell(x, r)} \leqslant K$$

where

$$\mathfrak{L} = \max \{\rho(f(x), f(y)) : \rho(x, y) = r\}$$

and

$$\ell = \min \{\rho(f(x), f(y)) : \rho(x, y) = r\}.$$

By a *conformal* homeomorphism of $\mathbb{S}^n$ we mean a Möbius transformation, that is the restriction to $\mathbb{S}^n$ of a finite composition of reflections in spheres or hyperplanes orthogonal to $\mathbb{S}^n$ in $\mathbb{R}^{n+1}$, [Ah 3]. In all dimensions, a self homeomorphism of the sphere is conformal if and only if it is 1-quasiconformal [Ge]. Conformal mappings are of course the boundary values of isometries of the usual Riemannian hyperbolic metric $ds^2 = (1 - |x|^2)^{-2}|dx|^2$ of constant negative curvature on the unit ball B^{n+1}.

A subfamily $\mathfrak{F}$ of Hom $(\mathbb{S}^n)$ is said to have the *convergence property* if each infinite subfamily of $\mathfrak{F}$ contains a subsequence $\{f_j\}$ for which one of the following two properties holds

(*a*) There is an $f \in \text{Hom}(\mathbb{S}^n)$ such that $f_j \to f$ and $f_j^{-1} \to f^{-1}$ uniformly.

(*b*) There are points x and y (possibly $x = y$) such that $f_j \to x$ locally uniformly in $\mathbb{S}^n - \{y\}$ and $f_j^{-1} \to y$ locally uniformly in $\mathbb{S}^n - \{x\}$.

One should note that there is a slight redundancy in the definition with regard to the convergence of the inverse mappings.

Definition 1.1. *A subgroup G of* Hom $(\mathbb{S}^n)$ *is called a* ***convergence group*** *if G has the convergence property. A subgroup G of* Hom $(\mathbb{S}^n)$ *is called a* ***quasiconformal group*** *if there is a finite K such that every $g \in G$ is K-quasiconformal and G is called a Möbius group (sometimes* ***conformal group****) if every $g \in G$ is a Möbius transformation.*

Our interest in the convergence property is twofold. Firstly it is the characteristic compactness property of conformal and quasiconformal mappings, see [GM Thm. 3.2].

Theorem 1.2. *The family $\mathfrak{F}_K$ of all K-quasiconformal homeomorphisms of $\mathbb{S}^n$ has the convergence property.*

Corollary 1.3. *A quasiconformal group G is a convergence group.*

Secondly the convergence property is topologically invariant.

Theorem 1.4. *If G is a convergence group and h is a self homeomorphism of $\mathbb{S}^n$, then the group $H = hGh^{-1}$ is also a convergence group.*

Thus the topological conjugates of Möbius and quasiconformal groups are convergence groups. Clearly then the convergence property is a necessary condition for a group G to be the topological conjugate of a quasiconformal group.

Definition 1.5. *Let Ω be a region in the sphere and $x \in \Omega$. A subfamily $\mathfrak{F}$ of* Hom $(\mathbb{S}^n)$ *is said to act* ***discontinuously at*** *x if there is a neighbourhood U of x in Ω such that for all but finitely many $f \in \mathfrak{F}$,*

$$f(U) \cap U = \varnothing.$$

We say that $\mathfrak{F}$ acts ***discontinuously in*** *Ω if it acts discontinuously at each point of Ω. We say that $\mathfrak{F}$ acts* ***properly discontinuously*** *in Ω if for each compact subset E of Ω*

$$f(E) \cap E = \varnothing$$

for all but finitely many $f \in \mathfrak{F}$.

Definition 1.6. *Let G be a convergence group acting on $\mathbb{S}^n$. The* ***ordinary set*** *for G is the set*

$$O(G) = \{x \in \mathbb{S}^n : G \text{ acts discontinuously at } x\}.$$

The ***limit set*** *of G is the set*

$$L(G) = \mathbb{S}^n - O(G).$$

We say that is ***discrete*** *if the identity is isolated in G.*

Notice that by definition $O(G)$ is open and $L(G)$ is closed. Recall that even in the conformal case it is possible that G is discrete while $O(G)$ is empty. Whenever G is discrete the convergence property (*a*) can never occur for if $g_j \to g$, then the sequence $g_j g_{j+1}^{-1}$ converges uniformly to the identity in G. The convergence property (*b*) guarantees $L(G)$ is not empty as soon as G is infinite and discrete. In fact the condition (*b*) easily implies.

Theorem 1.7. *A discrete convergence group G acts properly discontinuously in its ordinary set O(G).*

We recall that a *Schottky group* is a Möbius group of the n-sphere whose limit set is a *Cantor set*, that is a totally disconnected perfect set. The classical example of a Schottky group is the group generated by reflections in a finite disjoint collection of codimension one round spheres with the property that no sphere separates the collection. Schottky groups are a good example from which to build a general theory, they being the simplest infinite Möbius groups that exhibit interesting geometric and dynamic behaviour. The following result [GM Thm. 7.8] shows that any reasonable topological generalisation of a Schottky group will necessarily be a convergence group.

Theorem 1.8. *Let G be a subgroup of* Hom $(\mathbb{S}^n)$ *which acts properly discontinuously in* $\mathbb{S}^n - E$, *where E is a closed totally disconnected set. Then G is a discrete convergence group and L(G) lies in E.*

SKETCH OF PROOF. Let $\{g_i\}$ be an infinite sequence of elements of G and $x_0 \in \mathbb{S}^n - E$. Without loss of generality we may assume that $g_i(x_0) \to a$ and $g_i^{-1}(x_0) \to b$. Necessarily $a, b \in E$. We claim $g_i \to a$ locally uniformly in $\mathbb{S}^n - \{b\}$. Let $E_\delta = \{x \in \mathbb{S}^n : \rho(x, E) < \delta\}$. Then $E = \cap E_\delta$ and the maximum diameter of a component of E_δ tends to zero with δ. Assume that E_δ has connected complement. For sufficiently large i, $g_i(\mathbb{S}^n - E_\delta) \cap (\mathbb{S}^n - E_\delta) = \varnothing$ as G acts properly discontinuously in the complement of E. Since this set is connected its image must lie in a component of E_δ and as x_0 eventually lies in E_δ and $g_i(x_0) \to a$ we see that eventually $g_i(\mathbb{S}^n - E_\delta)$ lies in the component U_δ of E_δ containing a. The boundary of all the components of E_δ also must lie in this set and so the image of all but one of these components must lie in U_δ. It is not difficult to see that this one component must be that component V_δ of E_δ containing b. Since $U_\delta \to a$ and $V_\delta \to b$ the desired convergence follows. The details are in showing the assumption that the complement of E_δ is connected is not necessary.

Here is another useful characterisation of convergence groups. Let T_n denote the *triple space*

$$T_n = \mathbb{S}^n \times \mathbb{S}^n \times \mathbb{S}^n - \Delta$$

where Δ is the big diagonal $\Delta = \{(x, y, z) : x = y \text{ or } y = z \text{ or } x = z\}$. Thus T_n is the space of distinct triples of points on the sphere.

The space T_1 is homeomorphic to $B^2 \times \mathbb{S}^1$. Let $n \geqslant 2$ and $\zeta = (x, y, z) \in T_n$. We think of x, y and z as being three distinct points on the boundary $\mathbb{S}^n$ of hyperbolic $(n + 1)$-space B^{n+1}. Let $\mathcal{L}$ be the hyperbolic line joining x to y and

$p(\zeta)$ as the projection of z onto $\mathfrak{L}$. That is $p(\zeta)$ is the point of $\mathfrak{L}$ such that the hyperbolic ray joining $p(\zeta)$ to z is orthogonal to $\mathfrak{L}$. For each $w \in B^{n+1}$ the fiber $p^{-1}(w)$ is homeomorphic to the space of two frames of the tangent space $T_w B^{n+1}$ the fiber $p^{-1}(w)$ is homeomorphic to the space of two frames of the tangent space $T_w B^{n+1}$ (the inner product comes from the hyperbolic metric). The homeomorphism is given by the directions of x and z at $w = p(\zeta)$. Since $T_w B^{n+1} = B^{n+1}$ we find

$$T_n = B^{n+1} \times V_{n+1,2},$$

where $V_{n+1,2} = \{(u, v) \in \mathbb{S}^n \times \mathbb{S}^n : u \perp v\}$ is the Steiffel manifold of two frames on hyperbolic space. Thus we think of T_n as being hyperbolic space up to a compact factor.

Given a subgroup G of $\mathrm{Hom}(\mathbb{S}^n)$ there is a natural homomorphism ϕ embedding G in $\mathrm{Hom}(T_n)$ given by $\phi(g)(\zeta) = (g(x), g(y), g(z))$ where $\zeta = (x, y, z)$.

The point of this exercise is the following (see [GM 2]).

Theorem 1.9. *The following assertions are equivalent*:

(1) *G is a discrete convergence group acting on $\mathbb{S}^n$.*
(2) *$\phi(G)$ acts properly discontinuously on T_n.*

Furthermore, the action of $\phi(G)$ on T_n is effective if G is torsion free.

Usually we will forget the homomorphism ϕ and say that G acts on the triple space.

P. Scott has conjectured that a finitely generated group G of homeomorphisms of $\mathbb{S}^1$ is isomorphic to a Fuchsian group if and only if $\phi(G)$ acts properly discontinuously on T_1 [Sc. pp. 175]. This is conjecture is evidently implied by the existence of a topological conjugacy of any one dimensional convergence group to a Fuschsian group. It is here that we find the relationship with the Seifert conjecture. If G is convergence group acting on $\mathbb{S}^1$, then $M^3 = T_1/\phi(G)$ is a three manifold for which we have the short exact sequence

$$1 \to \mathbb{Z} \to \pi_1(M^3) \to G \to 1.$$

If the Seifert conjecture is true, then M^3 is Seifert fibered and G is isomorphic to a Fuchsian group. Given this isomorphism one can construct a topological conjugacy [MT]. Conversely, given an orientable three manifold M with $\mathbb{Z}$ as a normal subgroup and if M is not a circle bundle over the sphere then the universal cover of M is $B^2 \times \mathbb{R}^1$. Passing to the appropriate subcover given by the normal $\mathbb{Z}$ one can construct an action of the quotient group on the circle (which is thought of as ∂B^2) which acts properly discontinuously on the

triple space $T_1 \approx B^2 \times \mathbb{S}^1$. If this action is topologically conjugate to a Fuchsian group, then the quotient group is isomorphic to a Fuchsian group and consequently M is a Seifert fibered space, [Sc] [Me].

From [St. p. 132] we find a calculation of the homotopy groups of the Steiffel manifolds. In particular $\pi_i(V_{n+1,2}) = 0$, $i < n-1$ and $\pi_n(V_{n+1,2})$ is $\mathbb{Z}$ or $\mathbb{Z}_2$ depending on whether n is odd or even. Hence T_n is simply connected for $n > 2$ and thus a torsion free discrete convergence group of $\mathbb{S}^n$, $n > 2$, is isomorphic to the fundamental group of a $3n$-manifold.

This also suggests that an approach to the extension problem (from $\mathbb{S}^n$ to B^{n+1}) might be to try and split off the compact manifold factor from the action of the group on the triple space.

Theorem 1.9 has other important consequences. If M is a negatively curved $(n+1)$ manifold, then it is well known that the universal cover of M is diffeomorphic to the open $(n+1)$ ball. One can define a compactification of this universal cover: A point at infinity is an equivalence class of asymptotic geodesics (that is geodesics which remain a bounded distance apart). If the sectional curvature $K(M) \leqslant K < 0$, then the points at infinity form a topological n-sphere S_∞ which is the boundary of the universal covering space [EO]. It is not difficult to see that the elements of $\pi_1(M)$ act naturally on S_∞ via their action on the geodesics and that this action is as a group of homeomorphisms. The curvature assumption also implies that there is a unique geodesic line in the universal cover connecting any two points of the sphere at infinity. Using this fact, together with the fact that $\pi_1(M)$ is a discrete group of isometries and is therefore properly discontinuous on the universal cover, it is not too difficult to establish the following consequence of Theorem 1.9, see [MS Thm. 5.6].

Theorem 1.10. *Let M be a strictly negatively curved $(n+1)$-manifold. Then $\pi_1(M)$ acts naturally as a convergence group on the sphere at infinity of the universal cover.*

This provides yet another reason for studing convergence groups. Results of this type for compact M are implicit in earlier work of Gromov [Gr 3].

2. The Limit Set and the Elementary Convergence Groups

The following lemma is a useful tool in the study of the limit set of a discrete convergence group.

Lemma 2.1. *Let G be a discrete convergence group. Then for each point $x_0 \in L(G)$ there is a point $y_0 \in L(G)$ and a sequence $\{g_i\}$ of elements of G such*

that

$$g_i \to x_0 \text{ locally uniformly in } \mathbb{S}^n - \{y_0\}$$

and

$$g_i^{-1} \to y_0 \text{ locally uniformly in } \mathbb{S}^n - \{x_0\}$$

It is essentially this property which implies the following classical result on the structure of the limit set.

Theorem 2.2. *Let G be a discrete convergence group. Then*

(1) *$L(G)$ consists of either 0, 1, 2 points or else $L(G)$ is a perfect set.*
(2) *$L(G)$ is either nowhere dense or coincides with $\mathbb{S}^n$.*
(3) *If $\operatorname{card} L(G) > 2$, then $L(G)$ is the smallest closed G invariant set.*

If $L(G) = \mathbb{S}^n$, then we say G is of the *first kind.* Otherwise G is of the *second kind.*

Theorem 2.2 holds in much more generality of course. A more general construction of a limit set was given by R. Kulkarni [Ku]. He points out that there is a close connection between the number of components of the limit set and the number of ends of the group (a purely algebraic concept). Extending ideas of H. Hopf and H. Freudenthal he went on to develop the ends inequality [Ku, Thm. 4] which for (finitely generated) convergence groups is

The number of components of $L(G)$ is no more than the number of ends of G.

This result has a nice consequence (realised by Kulkarni in the conformal case) on the algebraic structure of certain convergence groups which we will see later.

In a discrete Möbius group there are three types of elements; elliptic, parabolic and loxodromic. Here is a topological generalisation of these notions.

Definition 2.3. *Let G be a discrete convergence group. We define*

$$\operatorname{ord}(g) = \inf\{m: m \geqslant 1 \text{ and } g^m = \text{identity}\}$$

and

$$\operatorname{fix}(g) = \{x \in \mathbb{S}^n: g(x) = x\}.$$

*We say g is **elliptic** if $\operatorname{ord}(g) < \infty$, g is **parabolic** if $\operatorname{ord}(g) = \infty$ and $\operatorname{card}\{\operatorname{fix}(g)\} = 1$ and finally g is **loxodromic** if $\operatorname{ord}(g) = \infty$ and $\operatorname{card}\{\operatorname{fix}(g)\} = 2$.*

Definition 2.4. *Let G be a discrete convergence group. We say that G is elementary if card $L(G) \leqslant 2$.*

Then, analogously to the classical case, it is quite natural to classify the elementary discrete convergence groups [GM 1].

Theorem 2.5. *$L(G) = \varnothing$ if and only if G is a finite group of elliptic elements.*

$L(G)$ consists of a single point if and only if G is an infinite group consisting only of elliptic and parabolic elements all of which must fix this point.

$L(G)$ consists of a pair of points if and only if G is an infinite group consisting only of elliptic and loxodromic elements each of which fixes or interchanges these points.

In the last case it is not difficult to show that G does indeed contain a loxodromic element. It is not known whether there must be a parabolic element in the case that $L(G)$ is a point. This is equivalent to the following question. (Recall we do not require the action to be effective!)

Is there an infinite purely torsion convergence group acting properly discontinuously on euclidean n-space? What if we suppose in addition that the group is finitely generated?

The answer is no in both cases when $n = 1$ or 2. This question should be regarded as a first step towards a more general Selberg Lemma [Se] as the existence of a torsion free subgroup of finite index (in the finitely generated case) implies the group is finite. In the conformal case, the answer is always no, [Wat].

The following result must be regarded as the first real interplay between topology and algebra in this theory. We repeat the sketch of proof from [GM 3].

Theorem 2.6. *Let G be a discrete convergence group which is abelian. Then G is elementary.*

SKETCH OF PROOF. Suppose that x_1, x_2, x_3 are three distinct points of $L(G)$. From Lemma 2.1 we obtain sequences and points such that for $k = 1, 2, 3$

$$g_{j,k} \to x_k \quad \text{as} \quad j \to \infty \quad \text{locally uniformly in } \mathbb{S}^n - \{y_k\}.$$

By symmetry we need only consider the two cases

(i) $x_1 \neq y_2$ and $x_2 \neq y_1$.

(ii) $x_1 \neq y_3$, $x_2 \neq y_1$ and $x_3 \neq y_2$.

Suppose (i) holds. Choose disjoint neighbourhoods U_1, U_2 of x_1, x_2 such that $y_1 \notin cl(U_1)$ and $y_2 \notin cl(U_2)$. Fix $y \notin \{y_1, y_2\}$. Then by the convergence property for all sufficiently large j

$$U_1 \supseteq g_{j,1}(\{y\} \cup U_2) \quad \text{and} \quad U_2 \supseteq g_{j,2}(\{y\} \cup U_1).$$

Thus

$$g_{j,2} \circ g_{j,1}(x) \in U_2 \quad \text{and} \quad g_{j,1} \circ g_{j,2}(x) \in U_1$$

and this is impossible as G is abelian. The argument is similar but a little more complicated when (ii) holds.

If we now consider the cyclic group generated by an element of a discrete convergence group we find that our classification of elements is exhaustive.

Corollary 2.7. *Let G be a discrete convergence group. Then every element of G is elliptic, parabolic, or loxodromic.*

A classical result of Newman [Ne] which says that periodic homeomorphisms of the sphere which fix an open set are the identity together with the above classification now implies a uniqueness result reminiscent to analyticity.

Proposition 2.8. *Let G a discrete convergence group and $f, g \in G$. If $f = g$ on an open set, then $f \equiv g$.*

The convergence property also implies the following dynamical behaviour of the elements of infinite order which is identical to that of the elements of infinite order in a discrete Möbius group.

Theorem 2.9. *If g is parabolic with fixed point x_0, then*

$$g^{\pm j} \to x_0 \quad \text{as} \quad j \to \infty \quad \text{locally uniformly in} \quad \mathbb{S}^n - \{x_0\}$$

If g is loxodromic with fixed points x_0 and y_0, then these points can be labeled so that

$$g^j \to x_0 \quad \text{as} \quad j \to \infty \quad \text{locally uniformly in} \quad \mathbb{S}^n - \{y_0\}$$

and

$$g^{-j} \to y_0 \quad \text{as} \quad j \to \infty \quad \text{locally uniformly in} \quad \mathbb{S}^n - \{x_0\}.$$

In light of this result and in view of the general question we have raised concerning topological conjugacy, one is naturally led to ask.

Are the elements in a discrete convergence group topologically conjugate to Möbius transformations?

We now turn to answer this fundamental question. Notice before we proceed that the classification necessarily applies to the elements of a quasiconformal group and so one can raise the above question in the quasiconformal context. We say g is a K-quasiconformal parabolic (respectively loxodromic) if the cyclic group $\langle g \rangle$ generated by g is a K-quasiconformal group and g is a parabolic (respectively loxodromic) element. We consider the three cases separately.

(1) Suppose g is parabolic.

Then in dimension one the answer is easily seen to be yes. In higher dimensions the convergence property for parabolics is known as Sperner's condition and homeomorphisms of $\mathbb{S}^n$ with this property are called *quasitranslations.* Results due to Sperner and Kérékjárto [Ke 1] show that the answer is again yes in dimension two (basically the action of $\langle g \rangle = \{g^m : m \in \mathbb{Z}\}$ on $\mathbb{R}^2$ is effective and so must cover an annulus). However it is not true in general that quasitranslations are topologically conjugate to translations, for examples see Kinoshita [Ki] in dimension 3 and Husch [Hu 1] for all higher dimensions. Here is an outline of a construction of a counterexample in dimension four. Let W^3 be a Whitehead three manifold. That is W^3 is a contractible three manifold not homeomorphic to $\mathbb{R}^3$. It is known that $W^3 \times \mathbb{R}$ is homeomorphic to $\mathbb{R}^4$. Then $W^3 \times \mathbb{S}^1$ and $\mathbb{R}^3 \times \mathbb{S}^1$ have the same universal covering space $\mathbb{R}^4$. The covering action extends to a quasitranslation of the sphere and these two actions are topologically distinct as the quotients are not homeomorphic.

If in addition g is a K-quasiconformal parabolic, the existence of a K'-quasiconformal conjugacy to a translation, where K' depends only on K and not g, is due to Hinkkanen [Hi] in dimension one (where it is surprisingly difficult) and the aforementioned results of Sullivan and Tukia in dimension two (which we outline later). The general question in the quasiconformal case is unknown in higher dimensions and we ask

Is every quasiconformal parabolic transformation topologically (quasiconformally) conjugate to a Möbius translation?

A generalization of Sullivan and Tukia's method yields an affirmtive answer for parabolic quasiconformal diffeomorphisms of the sphere, see [Ma 2]. Here the conjugacy is quasiconformal with dilatation depending only on that of the group and the dimension. There is an example however of a purely parabolic abelian quasiconformal group of $\mathbb{S}^n$, $n \geqslant 3$, which is not quasiconformally conjugate to a group of translations (it is topologically conjugate, see Section 5).

(2) Suppose g is loxodromic.

Then in all dimensions loxodromic transformations are topologically conjugate to Möbius transformations. Homeomorphisms satisfying the conclusion of Theorem 2.9 for loxodromic mappings are called *topological dilations.* Results due to Kérékjárto [Ke 2], Homma and Kinoshita [HK] and Husch [Hu 2] imply that orientation preserving topological dilations or their inverses are always conjugate to the dilation $x \to 2x$. Here is an outline of the proof of this fact:

We may assume that the fixed points of g are 0 and ∞ and that $g^m \to \infty$ as $m \to \infty$. Assume that there is a topological sphere S of codimension one separating 0 and ∞ and such that $g(S) \cap S = \emptyset$. Set $S_m = g^m(S)$. The convergence properties imply that S_m is a disjoint collection of spheres with 0 and ∞ as their only accumulation point. The Annulus Theorem implies that the region between S_m and S_{m+1} is homeomorphic to an annulus. Now it is easy to see how to use this annular structure to construct the desired conjugacy. We point out however that the use of the Annulus Theorem here is an essential use of this deep theorem which has only recently been established in all dimensions. We are left with the problem of constructing the sphere S. The convergence properties imply that eventually every sphere S separating 0 from ∞ moves of itself as it is compactly supported away from the limit set. In low dimensions one can use a finite cut and paste process to modify such an S to find one that moves off itself. In higher dimensions one must use more sophisticated methods. When $n \geqslant 3$, $\mathbb{S}^n - \{0, \infty\}$ is simply connected. The action of $\langle g \rangle$ on this space is effective and must cover a compact manifold with fundamental group $\mathbb{Z}$. This manifold must fiber over the circle and using some homotopy theory one can show that there is a sphere transverse to the fibration which lifts. This is the sphere we seek.

The existence of a topological conjugacy implies of course that

$$(\mathbb{S}^n - \{0, \infty\})/\langle g \rangle = \mathbb{S}^{n-1} \times \mathbb{S}^1.$$

In the quasiconformal case the existence of a topological conjugacy together with the fact that the quotient is compact implies (via the quasiconformal Hauptvermutung) that there is a quasiconformal conjugacy. It is not known if the dilatation of this conjugating map can be made to depend only on the dilatation of the group and the dimension. However this is the case if g is a diffeomorphism (where a more direct proof can be found, see Section 5).

(3) Suppose g is elliptic.

Then by a Theorem in part due to Brouwer, Kérékjárto and Eilenberg [Ei] every periodic homeomorphism of the two sphere is topologically conjugate to a rotation. By suspension this also implies the result in dimension one.

However this result is not true in higher dimensions. In dimension three Montgomery and Zippin [MZ] constructed an involution of $\mathbb{S}^3$ whose fixed point set is a wild knot. Such a map cannot of course be conjugate to an orthogonal rotation. However the affirmative solution to the Smith conjecture [MB] implies that periodic diffeomorphisms of $\mathbb{S}^3$ with a nonempty fixed point set are conjugate. The result is unknown if the fixed point set is empty. Examples due to Giffen and others [Gi] of smooth diffeomorphisms of all orders whose fixed points sets are smoothly knotted codimension two spheres show that there is no hope of topological conjugacy. Note that a finite group of diffeomorphisms is a quasiconformal group. For the general quasiconformal case in dimension one the existence of a quasiconformal conjugacy whose dilatation depends only on that of the group and not the order of the group is established in [Hi], while in dimension two the result follows form [Tu].

This summarises what is known about the elements in a discrete convergence group. There is the remaining question about the elements in a convergence group which is not discrete. It is clear we need only consider the cyclic case and we ask what can be said about a self homeomorphism of $\mathbb{S}^n$ with the property that for some sequence of integers $m(j)$ we have

$$g^{m(j)} \to \text{Identity uniformly in } \mathbb{S}^n.$$

In dimension one it is simple consequence of Denjoy's Theorem that such maps are conjugate to irrational rotations. The quasiconformal case can be found in [Hi]. In dimension two again it is always true that such maps are conjugate to irrational rotations of the two sphere [HM]. I do not know of any results in higher dimensions. Presumably it is not the case that such topological irrational rotations are conjugate to orthogonal irrational rotations in any higher dimension, even in the quasiconformal case. I do not know the answer to the analogous identity theorem for such maps;

Suppose that g is a topological irrational rotation and that g is the identity on an open set. Is g necessarily the identity?

It is worthwhile to recall a result of A. Hinkkanen, [Hi 2], who shows that every non-discrete convergence group of $\mathbb{S}^1$ is topologically conjugate to a Möbius group.

Returning to the structure of the limit set we observe from the classification of the elementary groups that every nonelementary discrete convergence group contains at least one loxodromic element. The following theorem shows that actually the loxodromic fixed points are pairwise dense in $L(G) \times L(G)$, see [GM, Thm. 6.7].

Theorem 2.10. *Let G be a nonelementary discrete convergence group and suppose that U and V are open sets which both meet the limit set. Then there is a loxodromic element in G with one fixed point in U and the other in V.*

Another result in this direction analogous to a classical result is

Theorem 2.11. *Let G be a discrete convergence group and $f, g \in G$. If f and g have a common fixed point and if g is loxodromic, then f and g have two fixed points in common and there is a integer k such that*

$$f \circ g^k = g^k \circ f.$$

A simple consequence of the plethora of loxodromic elements in a nonelementary discrete convergence group is

Theorem 2.12. *Let G be a discrete nonelementary convergence group. Then for every m, G contains a subgroup isomorphic to the free group of rank m.*

Finally here is another nice interplay between algebra and topology for convergence groups (for a conformal version see [Ku]). Stalling's theorem [Sta] states that a finitely generated group G with infinitely many ends splits as a free product with amalgamation or as an HNN extension. If G is a convergence group whose limit set is a Cantor set, then the ends inequality (see the note following Theorem 2.2) and Stalling's theorem imply that G splits. Each of these spliting subgroups has fewer than three ends or has infinitely many and the limit set correspondingly must consist of fewer than three points or is again a Cantor set. One can then inductively proceed to split those splitting subgroups with infinitely many ends. In the end (if this splitting process terminates) we are left with a complete splitting of G into subgroups all of whose limit set consists of fewer than three points. These are the elementary groups.

Theorem 2.13. *Let G be a finitely generated discrete convergence group whose limit set is a Cantor set. Then G splits as* HNN *extensions and free products with amalgamation (possibly infinite) and each factor is an elementary convergence group.*

In the special case that G is torsion free the splitting process must stop by Grushko's theorem. As we have seen the elementary convergence groups with a single limit point are hard to classify (they could be isomorphic to Fuchsian groups for instance). However Theorem 2.11 easily implies those torsion free elementary convergence groups with two limit points are just the infinite cyclic

groups generated by a loxodromic. There is an easy way to assert that in a convergence group with Cantor limit set there are no elementary subgroups with a single limit point and that is to assume that every element is loxodromic. Thus we have the following nice corollary to the above [GM 2].

Corollary 2.14. *Let G be a finitely generated convergence group of $\mathbb{S}^n$ whose limit set is a Cantor set and for which every element is loxodromic. Then G is isomorphic to a free group.*

B. Maskit has a stronger version of this result, including a partial converse, in the Kleinian case [Mas 2].

3. Groups Acting on the Circle

Here is the best known result concerning discrete convergence groups acting on the circle. It is due to P. Tukia [Tu, Thm. 6 B]. We will sketch a proof from notes of Tukia's talk at the Nevanlinna Colloquium. Tukia's proof is motivated by the partial solutions to the Nielsen realisation problem obtained by Nielsen and Zieschang. Central to the proof is the existence of a simple axis.

Theorem 3.1. *Let G be a discrete convergence group of the circle. Then either G is topologically conjugate to a Fuchsian group or G has a semitriangle group of finite index.*

A *semitriangle group* is a discrete nonelementary convergence group generated by two elements f and g such that for some $p, q, r > 1$

$$f^p = g^q = (f \circ g)^{-r}.$$

Such groups are factors of triangle groups and are of such a special nature it might be supposed that they do not exist (as subgroups of $\mathrm{Hom}(\mathbb{S}^1)$). Here are some conditions which imply that G does not contain a semitriangle group of finite index and so G is topologically conjugate to a Fuchsian group.

(1) G is torsion free.
(2) G is isomorphic to a Fuchsian group.
(3) $L(G) \neq \mathbb{S}^1$.
(4) G contains a parabolic element.
(5) G is infinitely generated.
(6) $T_1/\phi(G_0)$ is noncompact or virtually Haken (G_0 is the orientation preserving subgroup).

It should be noted that if there are semitriangle convergence groups. Then it is not true (as is conjectured) that every 3-manifold is virtually Haken (that is finitely covered by a Haken manifold). The 3-manifold $T_1/\phi(G)$ is the counterexample.

SKETCH OF PROOF. We assume that G is orientation preserving. From the above every element of G is topologically conjugate to a Möbius transformation. We want to extend the action of G to the unit disk and apply the following result of Martin and Tukia [MT, Thm. 4.4] whose proof we also outline.

Theorem 3.2. *Let G be a discrete convergence group acting on the closed unit disk $\mathfrak{D}$. Then G is topologically conjugate to a Fuchsian group.*

PROOF. $(\mathfrak{D} - L(G))/G$ is a (possibly bordered) surface with isolated branch points. Using this surface construct a G-invariant triangulation of $\mathfrak{D} - L(G)$. Redefine the conformal structure of $\mathfrak{D} - L(G)$ so that G now acts conformally, this can be done as in Ahlfors and Sario [AS, p. 127]. This conformal structure on int $(\mathfrak{D})$ is conformally that of the standard unit disk or the entire complex plane. This last case is impossible as it implies G is elementary. If $L(G)$ is totally disconnected the conjugacy is easily extended to the whole disk $\mathfrak{D}$ while if $L(G) = \mathbb{S}^1$ one can extend the conjugacy using the fact that every limit point has a neighbourhood basis bounded by loxodromic axes by Theorem 2.10.

Returning to our sketch of Tukia's proof, we define an *axis* to be a pair (a, b) of distinct points of $\mathbb{S}^1$. We identify and axis with the hyperbolic line joining the points and we say two axes intersect if the corresponding hyperbolic lines meet. If $A = (a, b)$ is an axis and $g \in G$ we set $g(A) = (g(a), g(b))$. We say an axis is *simple* with respect to G if

(i) A and $g(A)$ do not intersect for any $g \in G - \{\text{identity}\}$, and
(ii) for any $\epsilon > 0$, $|g(a) - g(b)| < \epsilon$ for all but finitely many $g \in G$.

Suppose that there is a simple axis A. We form the one complex

$$X_A = \mathbb{S}^1 \cup \{g(A): g \in G\}.$$

The simplicity of A easily enables us to extend the action of G to this complex retaining the convergence property. A component C of $\mathfrak{D} - X_A$ has boundary homeomorphic to a circle and the stabilizer G_C of this boundary acts as a convergence group of this circle. The limit set of G_C lies in $\mathbb{S}^1$ and so the group is of the second kind. It is not difficult to show that there is a simple axis for groups of the second kind (the endpoints of an interval of discontinuity). We repeat the construction of the one complex in this component C and then

transfer it to all components conjugate to C in G. In this way we obtain a convergence group of an extended complex $Y_A \supseteq X_A$.

We repeat this process, perhaps infnitely often, to find a one complex X on which G acts as a convergence group and such that if C is a component of $\mathfrak{D} - X$, then the stabilizer of C is an elementary convergence group. It is easy to conjugate the elementary groups to Fuchsian groups and so we are able to extend the action of G on X to the whole disk.

This then leaves us with the problem of constructing a simple axis or showing that there is a semitriangle group of finite index. It is important to note that not all Fuchsian groups have simple axes, for instance cocompact triangle groups do not. The basic starting point is the existence of a *regular axis.* This is an axis A satisfying (ii) above and such that $\{g \in G - G_A : g(A) \cap A = \varnothing\}$ is a finite set, here $G_A = \{g \in G : g(A) = A\}$. The cardinality of this set is the intersection number of A. It can be shown that the fixed points of a loxodromic element of G form a regular axis, this follows more or less from the convergence property. Thus we start with a loxodromic axis A which is regular but not simple and try to produce another axis with smaller intersection number. This is done by carefully studying the intersection patterns of those axes which are conjugate in G to A and meet A. One can show that there is an axis with a smaller intersection number unless a specific pattern occurs which then leads to the construction of a semitriangle group. An argument based on the action of the group on the triple space shows this semitriangle group to be of finite index.

The remaining problem of dealing with the semitriangle groups seems to be quite difficult. Tukia has suggested a way of conjugating every discrete convergence group of the circle to a quasisymmetric group (the one dimensional analogue of quasiconformal groups). The hope is that the problem will be easier given a more geometric constraint on the group. Here is what is known at present about quasisymmetric groups (that does not follows from Tukia's result above), the results are due to Hinkkanen [Hi].

Theorem 3.3. *Let G be a K-quasisymmetric group of the circle such that either G is elementary or G is not discrete. Then G is conjugate to a Fuchsian group by a K'-quasisymmetric mapping, where K' depends only on K and not on G.*

P. Tukia has kindly pointed out the following

Theorem 3.4. *Let G be a discrete convergence group acting on $\mathbb{S}'$. Then G is topologically conjugate to a uniformly quasisymmetric group.*

PROOF. If G does not contain a semitriangle group of finite index, then G is topologically conjugate to a Möbius group, Theorem 3.1. Otherwise the

action of G on the triple space is cocompact and by a result of G. Mess [Me] the group G is coarse quasi-isometric to the hyperbolic plane. That is, after a change of coordinates, G acts as a quasisymmetric group of $\mathbb{S}^1$.

Presumably some version of the following question has an affirmative answer and so one can apply Tukia's results to the quasisymmetric case as well.

Suppose that G is a discrete K-quasisymmetric group which is topologically conjugate to a Fuchsian group. Is G conjugate to a Fuchsian group by a K'-quasisymmetric mapping, where K' depends only on K and not on G?

It is not difficult to see a possible argument in the case that G is of the first kind.

We point out here that if one could extend a quasisymmetric group of the circle to a quasiconformal group of the unit disk, then this extension is necessarily quasiconformally conjugate to a Fuchsian group and so too therefore is the boundary group. There are many ways to extend a quasisymmetric map of the circle to the disk and some of these are canonical in many respects, see especially the extension found by Douady and Earle [DE]. However no extension can respect the composition of general quasisymmetric maps and so the extension of a group to the disk may no longer be a group acting on the disk. It might be however, that there is a canonical extension respecting composition for quasisymmetric groups.

We present here an interesting connection between the problem of conjugacy of quasisymmetric groups to Fuchsian groups and Teichmüller theory. The reader who is familiar with Teichmüller theory will recall that the Nielsen realisation problem (the problem of realising a finite group of homotopies of a given surface F^2 as a group of isometries of F^2 in some conformal structure) is implied by showing that a finite group of isometries of the Teichmüller space of F^2 has a common fixed point.

The *Schwarzian derivative* $S(\varphi)$ of a holomorphic mapping φ of the unit disk $\mathfrak{D}$ is defined as

$$S(\varphi) = \left(\frac{\varphi''}{\varphi'}\right)' - \frac{1}{2}\left(\frac{\varphi''}{\varphi'}\right)^2.$$

The Schwarzian derivative measures the deviation of a holomorphic map from a Möbius transformation as seen by the fact that it precisely annihilates the latter. We define

$$\mathfrak{U} = \{S(\varphi): \varphi \text{ is univalent in } \mathfrak{D}\}.$$

Universal Teichmüller space is $\mathfrak{I} = \text{int}(\mathfrak{U})$. A quasidisk is the image of the unit disk under a quasiconformal self homeomorphism of $\mathbb{R}^2$. Here is the connection between quasisymmetric mappings and Schwarzian derivatives, relevant details and proofs can be found in [Le].

Let g be a quasisymmetric homeomorphism of the circle. Then there is a quasidisk Ω and univalent mappings $\varphi: \mathfrak{D} \to \Omega$ and $\psi: \mathbb{C} - \mathfrak{D} \to \mathbb{C} - \Omega$, such that

$$g = \psi^{-1} \circ \varphi \mid \mathbb{S}^1.$$

This construction is essentially unique so that the map $g \to S(\varphi) = S_g$ is well defined and, given some normalisation of g, is bijective (it is really the Bers' embedding).

Universal Teichmüller space has a few natural metrics. The Teichmüller metric is defined by

$$d(S_g, S_f) = \inf \{\log K: g \circ f^{-1} \text{ has a } K\text{-quasiconformal extension to } \mathfrak{D}\},$$

other natural metrics and discussions of their various properties can be found in the recent books [Ga], [Le] and [Na]. We point out that this metric is sufficiently regular to enable one to construct geodesics between any two points. Our interest in all of this lies in the following representation of a quasisymmetric group G as a subgroup of a compact subgroup of the isometry group of Universal Teichmüller space.

Let g be a quasisymmetric map of the circle. Define the map $g^*: \mathfrak{I} \to \mathfrak{I}$ by the rule

$$g^*(S_f) = S_h,$$

where $h = f \circ g^{-1}$.

It is not difficult to see that g^* is an isometry of $\mathfrak{I}$ with the Teichmüller metric and that the set $G^* = \{g^*: g \in G\}$ is a group of isometries. Since G is a quasisymmetric group, the set

$$S(G) = \{S_g: g \in G\}$$

is a bounded subset of $\mathfrak{I}$. It is also G^* invariant. This enables us to prove that the closure of G^* is a compact group of isometries. If the elements of G^* have a common fixed point S_h, then

$$g^*(S_h) = S_h \Leftrightarrow h \circ g^{-1} = \gamma \circ h$$

for some Möbius transformation γ. Thus h conjugates G to a Möbius group. There are many natural ways to try to find a fixed point for a compact group action on a contractible space (as $\mathfrak{I}$ is) using the metric structures available.

For instance if the set $S(G)$ has a well defined barycenter. Or if there is a metric with enough «convexity» in which G acts as isometries. The fact that geodesics are not always unique for the Teichmüller metric causes problems for constructions based on this metric. In the finite dimensional case, where this construction works perfectly well, G^* acts as isometries in the Weil-Petersson metric. S. Wolpert [Wo] has shown that this metric is negatively curved and so by an old result of Cartan G^* has a common fixed point. This is another proof of the Nielsen problem. We therefore finish this section by posing the following problem:

Find a metric ρ on universal Teichmüller space in which the action g^* of a quasisymmetric map g is as an isometry and such that a compact group of isometries has a common fixed point.

It is quite possible that the Teichmüller metric is such a metric.

4. Groups Acting on the Two Sphere

As we have mentioned, every quasiconformal group acting on $\mathbb{S}^2$ (or more generally any subdomain) is quasiconformally conjugate to a Möbius group. This result was first proved by Sullivan [Su 1] and his proof was formalised and generalised by Tukia [Tu 2]. The methods developed for the two dimensional quasiconformal case have useful applications in higher dimensions. Thus we will give here a brief outline of the idea of the proof and then present most of the details in a more general setting in Section 5.

Theorem 4.1. *Let G be a K-quasiconformal group acting on $\mathbb{S}^2$. Then there is a K'-quasiconformal self homeomorphism f of the sphere such that the group fGf^{-1} is a Möbius group. Here $K' \leqslant \min\{K^{1/\sqrt{2}}, \sqrt{2K}\}$.*

IDEA OF PROOF. Suppose that G is a group of diffeomorphisms and $x \in \mathbb{S}^2$. What it means for G to have bounded distortion is that for each $g \in G$ the linear mappings $g'(x)^t g'(x)$: $T_x\mathbb{S}^2 \to T_{g(x)}\mathbb{S}^2$ send circles to ellipses of uniformly bounded eccentricity. Since we are only interested in the associated conformal structures we normalise these ellipses so that the product of the length of their axes is one (that is $\det g'(x) = 1$). For each $g \in G$ we construct the ellipse coming from $g^{-1}(x)$ under the map induced by g. This family of ellipses based at x is G-invariant. That is, the family of ellipses based at $g^{-1}(x)$ is sent to the family based at x under the map induced by g. So too then is the average ellipse which we can construct because of the uniform bound on eccentricity and diameter. This constructs a G-invariant ellipse field; a field of ellipses in the tangent bundle

which is invariant under the map induced by any element of G and has uniformly bounded eccentricity. The smoothness assumption on the elements of G can be relaxed (quasiconformal mappings are differentiable almost everywhere). In this way we construct a measurable G-invariant ellipse field. It is this much of the construction that can be done in all dimensions. Next in dimension two the existence theorem for quasiconformal mappings (or as it has lately become known, the measurable Riemann mapping theorem) implies that any measurable ellipse field is almost everywhere the ellipse field induced by some quasiconformal mapping. Conjugating by the quasiconformal homeomorphism h corresponding to the G-invariant ellipse field we see that the group $H = hGh^{-1}$ leaves the standard field of round circles invariant. Thus every element of H is conformal and so H is a Möbius group.

As a first application of this result we recall from Theorem 1.10 that the action of the fundamental group Γ of a negatively curved three manifold on the universal cover is a convergence group of the two sphere at infinity S_∞. Evidently this sphere carries a quasiconformal structure in which Γ acts quasiconformally if and only if M has the homotopy type of a hyperbolic space form.

If M has quarter pinched sectional curvatures $-4 < K(M) \leqslant -1$, then one knows that the sphere at infinity has a compatible C^1-structure in which the fundamental group acts as diffeomorphisms, [Gre] and [HP]. D. Sullivan has made the following conjecture in this case [Su 2].

If M is a quarter pinched three manifold, then $\pi_1(M)$ acts as a quasiconformal group on the two sphere at infinity.

As a consequence of course we find that quarter pinched three manifolds have the homotopy type of hyperbolic manifolds. Sullivan has related this conjecture to other questions concerning the expanding properties of the geodesic flow on the tangent spaces to the horospheres. Examples of Gromov and Thurston [GT] of n-manifolds, $n \geqslant 4$, with sectional curvatures close to minus one that do not have the homotopy type of hyperbolic manifolds show that this quasiconformality must be a low dimensional phenomena.

We return to the question of whether or not convergence groups are also conjugate to Möbius groups in dimension two. Here is a simple example that is easily generalised to all dimensions to show that this is not the case (although recall that in dimension two every element is individually conjugate).

Example 4.2. Let G be a nonelementary Fuchsian group acting on $\mathbb{S}^2$. Then G leaves the disk $\mathbb{D}^2$ invariant. Identify $\mathbb{D}^2$ to a point x_0 and extend the action of G over this point by agreeing that every element of G fixes x_0. This pro-

duces a group of homeomorphisms H of $\mathbb{S}^2/\mathfrak{D} \approx \mathbb{S}^2$ acting properly discontinuously in the complement of the point x_0. Thus H is a convergence group and is not conjugate to any Möbius group since the stabilizer of a point in a Möbius group is virtually abelian and the stabilizer of x_0 is $H \cong G$.

One can see how to generalise this process of taking an equivariant decomposition of the sphere to produce more complicated examples of convergence groups. The point to this example is that in dimension two it seems quite likely that this is the only way to construct convergence groups which are not the topological conjugates of Möbius groups. We now outline the results of R. Skora and the author [MS]. To begin with we need the following terminology

Definition 4.3. *A subset X of $\mathbb{S}^n$ is* ***cellular*** *if $X = \cap B_j$, where $\{B_j\}$ is a collection of closed topological n-balls (cells) such that B_{i+1} lies inside* Int (B_j). *A map $\eta: \mathbb{S}^n \to \mathbb{Y}$ is* ***cellular*** **if** *$\eta^{-1}(y)$ is cellular for every $y \in \mathbb{Y}$. A* ***decomposition*** *$\mathfrak{R}$ of $\mathbb{S}^n$ is a disjoint collection of subsets of $\mathbb{S}^n$ such that every point of $\mathbb{S}^n$ lies in some element of $\mathfrak{R}$. We say a decomposition $\mathfrak{R}$ is* ***cellular*** *if the quotient map $\mathbb{S}^n \to \mathbb{S}^n/\mathfrak{R}$ is cellular.*

A subset X of the two sphere is cellular if and only if X is a closed connected nonseparating set (an easy application of the Riemann mapping Theorem).

A decomposition $\mathfrak{R}$ of $\mathbb{S}^n$ is *upper semi-continuous* if the quotient map is a closed map and it is *G-invariant* or simply *equivariant* if its elements are permuted by G. The applications we have in mind rest on the following important result of R. L. Moore [Mo].

Theorem 4.4. *If $\mathfrak{R}$ is a cellular, upper semi-continuous decomposition of $\mathbb{S}^2$, then $\mathbb{S}^2/\mathfrak{R}$ is homeomorphic to $\mathbb{S}^2$.*

An immediate consequence is

Corollary 4.5. *Let G be a discrete convergence group of $\mathbb{S}^2$ and $\mathfrak{R}$ an equivariant upper semi continuous cellular decomposition. Then the induced action of G on $\mathbb{S}^2/\mathfrak{R}$ ($\approx \mathbb{S}^2$) is a convergence group of $\mathbb{S}^2$.*

We need to make the following technical definition.

Definition 4.6. *A finitely generated group Γ is* ***accessible*** *if every sequence of nontrivial algebraic splittings of Γ as free products with amalgamation along finite subgroups or as* HNN *extensions along finite subgroups is finite.*

It is conjectured by C.T.C. Wall that every finitely generated group is accessible [Wa]. This conjecture has been established in the case that Γ has

uniformly bounded torsion or is finitely presented by P. Linnell [Li] and M. Dunwoody [Du] respectively. It is a simple consequence of Grushko's theorem that a finitely generated torsion free group is accessible and since Möbius groups are virtually torsion free (Selberg's Lemma) they too are always accessible.

The main result of [MS] is really the Decomposition Theorem. It implies that an accessible convergence group G with $O(G) \neq \emptyset$ is the fundamental group of a finite graph of groups. The edge groups are finite cyclic and the vertex groups are convergence groups for which every component of the ordinary set is simply connected. A consequence of this result will be that under certain circumstances $O(G)/G$ is of finite topological type. This is a version of Ahlfors' Finiteness Theorem. The proof of the Decomposition Theorem is analogous to Abikoff and Maskit's decomposition of finitely generated Kleinian groups [AM]. If some component of the ordinary set is not simply connected we use a topological version of Maskit's Planarity Theorem [Mas 1] to find a «nice» invariant union $\mathcal{C}$ of simple loops in $O(G)$. From this we construct a tree on which G acts without inversions as follows. Define an equivalence relation on $O(G) - \mathcal{C}$ as by saying that two points are equivalent in no element of $\mathcal{C}$ separates them. If $[x]$ and $[y]$ are equivalence classes then at most finitely many elements of $\mathcal{C}$ separate them. Thus $\mathcal{C}$ determines a tree where every vertex is an equivalence class of points of $O(G) - \mathcal{C}$ and two vertices are joined by an edge if exactly one element of $\mathcal{C}$ separates them. G acts on this tree without inversions if it is orientation preserving. For a vertex v we set $G_v = \{g \in G: g(v) = v\}$ and similarly define the stabilizer of an edge. Then by the Bass-Serre theory of groups acting on graphs G admits a spliting as $G \cong G_1 *_{\mathbb{Z}/n\mathbb{Z}} G_2$ or $G \cong G_1 *_{\mathbb{Z}/n\mathbb{Z}}$ and each factor G_i is a vertex group and the amalgamating groups are the edge groups. One then proceeds inductively to find splitings of G_1 and G_2. If this process terminates we have a complete understanding of G up to these factors, this is why we need the hypothesis of accessibility. In the case that G is a Kleinian group this splitting process terminates by Ahlfors' Finiteness Theorem which says that $O(G)/G$ is of finite type [Ah 2].

Decomposition Theorem 4.7. *Let G be an orientation preserving convergence group which is accessible. Then there exists an invariant union $\mathcal{C}$ of disjoint simple loops in $O(G)$ such that $\mathcal{C}/G$ is a disjoint collection of simple loops and each component of $O(G_v)$ is simply connected for every vertex v in the tree determined by $\mathcal{C}$.*

Actually one gets somewhat more from this result. Namely a picture of how the limit set is built up from the limit sets of the vertex groups [MS, pp 4-5].

The decomposition theorem reduces the problem of classifying those accessible convergence groups with $O(G) \neq \emptyset$ to classifying those for which every component of the ordinary set is simply connected. Unfortunately even in the case of Kleinian groups this classification is far from complete, containing Web groups and all groups of the first kind! In any case we have achieved one of our objectives. It is a weak form of Ahlfors' Finiteness Theorem.

Theorem 4.8. *Let G be an accessible convergence group such that every component of $L(G)$ separates the two sphere into finitely many (and therefore at most two* [GM]*) pieces. Then $O(G)/G$ has finite type. That is it has finitely many components, each component has finite genus, a finite number of ends and a finite number of orbifold branch points.*

There have been other results along this line. R. Kulkarni and P. Shalen [KS] showed that if G is a finitely generated torsion free discrete convergence group acting on the closed unit 3-ball, then $(O(G) \cap \mathbb{S}^2)/G$ is of finite type, except for the possibility that there are infinitely many disks or annuli. It is unknown whether this last possibility can occur. Recently Kulkarni has found a version of Theorem 4.8 in the case that the limit set is a Cantor set. And M. Feighn and D. McCullogh refined the techiques of Kulkarni and Shalen [FM] to obtain topological analogues of Bers' Area Theorem [Be] and Sullivan's Finiteness of Cusps [Su 3] and [Kr]. Our methods too provide weak (but topological) forms of these estimates. One can directly check that under the hypotheses of the Decomposition Theorem and Theorem 4.8, that $\chi(O(G)/G) \geqslant 2(1 - N)$, where χ is the orbifold characteristic and N is the fewest number of generators of G. This is Bers' Area Theorem. Similarly it implies that there are a finite number of conjugacy classes of parabolic fixed points of G and under the additional hypothesis of G being torsion free it implies this number is no more than $3N/2$. This is Sullivan's finiteness of cusps.

We now proceed by making an assumption on the structure of the limit set of the convergence group G. We want to identify those convergence groups G_j in the splitting of G that comes from the Decomposition Theorem as topological conjugates of Fuchsian groups. We wish to eliminate from our considerations, those convergence groups G_j whose ordinary set is simply connected and whose limit set is a nonseparating continuum and those convergence groups G_j whose limit set is connected and has infinitely many components in its complement (principally because we cannot show that these groups are conjugates of Kleinian groups in any reasonable sense). In the first case we would like to say that every such convergence group is obtained from a Fuchsian group by some identification on the limit set $\mathbb{S}^1$. Unfortunately it is not known whether this is even true for Kleinian groups although it is con-

jectured to be the case. This is the question of the local connectedness of the limit set of a finitely generated function group. The second case concerns Web groups and seems even more difficult. This leaves us with essentially two types of convergence groups to deal with. Firstly the elementary groups. We have already classified these: The finite groups are topologically conjugate to finite Möbius groups, the groups with a single limit point are either conjugate to a euclidean group or are constructed exactly as in Example 4.2. The former are identified by the fact that the stabiliser of the limit point is virtually abelian. Secondly there are the convergence groups whose limit set is a topological circle. These groups are topologically conjugate to Fuchsian groups, see [MT]. This follows more or less from Theorem 3.2.

Given a convergence group G such that every component of the limit set is a point or a circle we decompose G into the groups G_j as given by the decomposition Theorem. We conjugate the G_j to Fuchsian groups by homeomorphisms f_j where this is possible. For the groups arising as in Example 4.2 we need a semi-conjugacy by a cellular mapping η_j. One then recombines these Kleinian conjugate groups using the Klein-Maskit Combination Theorem to obtain a Kleinian group Γ. The map which will «conjugate» G to Γ is locally defined by the f_j and the η_j and will correspond to some equivariant (with respect to Γ) cellular decomposition of the sphere. The elements of the decomposition which are not points will be disks whose boundary lies in the limit set of Γ. Thus the convergence group G is obtained from Γ via this cellular decomposition. We capture this notion in the following definition

Definition 4.9. *Let G and H be convergence groups of $\mathbb{S}^n$. We say that G is covered by H if there is an isomorphism $\varphi: H \to G$ and a cellular map $\eta: \mathbb{S}^n \to \mathbb{S}^n$ such that $\eta/\eta^{-1}(O(G))$ is a homeomorphism and the following diagram commutes*

$$\begin{array}{ccc} H \times \mathbb{S}^n & \longrightarrow & \mathbb{S}^n \\ \downarrow \varphi \quad \downarrow \eta & & \downarrow \eta \\ G \times \mathbb{S}^n & \longrightarrow & \mathbb{S}^n. \end{array}$$

Our discussion has established the first part of the next theorem. The second part follows since if the fixed point data of G is correct, then all the conjugacies are topological.

Theorem 4.10. *Let G be a finitely generated accessible convergence group such that every component of the limit set $L(G)$ is either a point or a topological circle. Then G is covered by a Kleinian group. If in addition G has the correct fixed point data, namely the stabilizer of any point in $L(G)$ is virtually abelian, then G is topologically conjugate to a Kleinian group.*

As a corollary to this and Theorem 1.8 we obtain the following topological characterisation of Schottky groups in the plane, once we observe that as soon as there is an invariant component of the ordinary set of a finitely generated convergence group, then it is accessible [MS, Prop. 5.1 and Cor. 5.5].

Corollary 4.11. *Let G be a finitely generated group of homeomorphisms acting properly discontinuously in the complement of a totally disconnected set. If the stabilizer G_x of any point $x \in \mathbb{S}^2$ is virtually abelian, then G is topologically conjugate to a conformal Schottky group.*

Of course in Corollary 4.11, the assumption on the fixed point data is necessary.

We conclude by conjecturing that Theorem 4.10 is true in much more generality and is really the only way of constructing convergence groups on $\mathbb{S}^2$.

Conjecture. *Every convergence group G acting on $\mathbb{S}^2$ is covered by a Kleinian group.*

This conjecture implies that every negatively curved three manifold has the homotopy type of a hyperbolic space form. It implies the following (weaker?) conjecture which is also true if G is as in Theorem 4.10.

Conjecture. *Every convergence group G acting on $\mathbb{S}^2$ extends to a convergence group G' of B^3.*

This would be a topological Poincaré extension theorem.

5. Higher Dimensions

The examples of C. Giffen [Gi] of smooth involutions of the n-sphere, $n \geqslant 4$, which we mentioned above provide simple examples of smooth uniformly quasiconformal groups which are not the topological conjugates of Möbius groups. Presumably it is possible to construct a quasiconformal version of the example of Montgomery and Zippin [MZ] to provide a three dimensional example (the affirmative solution to the Smith conjecture implies that there is no smooth version).

To obtain more complicated examples, one can use the quasiconformal analogue of the Klein-Maskit combination theorems [Ma, Thm. 4.3] to combine one of the examples above with a various Möbius groups. Indeed such a construction is possible for any Möbius group Γ with nonempty ordinary set. We can easily arrange that the smooth involution acts conformally on some open set and (after an appropriate conjugacy of Γ by a suitable Möbius

transformation) that this set contains the complement of a fundamental domain for the action of Γ on $O(\Gamma)$. Using the combination theorem we obtain a uniformly quasiconformal group G which is isomorphic as group to $\Gamma * \mathbb{Z}_p$ and which is not topologically conjugate to any Möbius group.

There are many natural questions that these examples leave unanswered. For instance, is the existence of a fake involution the only way to construct such quasiconformal examples? If we were to assume that the group was torsion free or that every element of the group has the correct fixed point data could there still be such examples? Example 4.2 generalises to give examples in all dimensions of convergence groups that are not topologically conjugate to Möbius or even quasiconformal groups. Is this essentially the only way to construct such examples as it seems in dimension two? We will discuss these questions and then turn to consider some affirmative results regarding the existence of a conjugacy.

The first relevant example is due to P. Tukia [Tu 4].

Theorem 5.1. *For every $n \geqslant 3$ there is a uniformly quasiconformal action of a solvable Lie group on $\mathbb{S}^n$ which is not isomorphic as a topological group to any Möbius group.*

This example was an extension of another example, due also to him, of a uniformly Lipschitz (in the euclidean topology of $\mathbb{R}^n$) action which was topologically conjugate but not quasiconformally conjugate to a group of translations. This latter example was constructed by arranging that the orbit of a point was a hyperplane that was not locally quasiconformally flat (in fact this hyperplane is the product of the Von Koch snowflake and $\mathbb{R}^{n-2}$). This group was not discrete. Tukia's methods were subsequently modified and extended by the author in [Ma 1] to prove the following theorem.

Theorem 5.2. *For each $n \geqslant 3$ there is a discrete uniformly quasiconformal group of parabolic transformations acting on $\mathbb{S}^n$, isomorphic as a group to $\mathbb{Z}^{n-1}$ and topologically but not quasiconformally conjugate to a Möbius group.*

This theorem is especially interesting since as we shall see all smooth abelian examples are quasiconformally conjugate to Möbius groups. But this does raise the question.

Suppose G is a uniformly quasiconformal group of $\mathbb{R}^n$ containing only parabolic transformations (fixing ∞). Is G topologically conjugate to a group of translations?

The next interesting examples are due to M. Freedman and R. Skora [F.S.].

Firstly a convergence group which has all the correct fixed point data and a simple algebraic structure but cannot be made quasiconformal by a change of coordinates. We recall that a Cantor set is *tame* in $\mathbb{S}^n$ if there is a self homeomorphism of $\mathbb{S}^n$ taking the Cantor set to a subset of a smoothly embedded arc in $\mathbb{S}^n$. Otherwise the Cantor set is *wild*.

Theorem 5.3. *For every $n \geqslant 3$ there is a convergence group G with the following properties.*

(i) *G is algebraically isomorphic to the free group on two generators and contains only loxodromic transformations,*
(ii) *$O(G)/G$ is compact and $L(G)$ is a wild Cantor set,*
(iii) *G is not topologically conjugate to any uniformly quasiconformal group.*

The example of Theorem 5.3 provides a nice counterpoint to M. Freedman's topological characterisation of certain Schottky groups [Fr].

Theorem 5.4. *Let G be a discrete convergence group of $\mathbb{S}^n$, $n \neq 4$, with the following properties.*

(i) *G is algebraically isomorphic to a free group and contains only loxodromic elements.*
(ii) *$O(G)/G$ is compact.*
(iii) *$L(G)$ is a tame Cantor set.*

Then G is topologically conjugate to a Schottky group.

Freedman goes on in that paper to equate the four dimensional surgery problem and the five dimensional s-cobordism «theorem» to the problem of extending to the four ball every convergence group G of $\mathbb{S}^3$ satisfying (i) and (ii) of Theorem 5.4 and whose limit set is a Cantor set. This reduction is quite complicated and related to the reduction of the surgery problem to the so-called atomic surgery problems [CF] (whether it is possible to solve the surgery problem in four manifolds with fundamental group isomorphic to a free group).

The hypothesis that $O(G)/G$ is compact effectively forbids the existence of parabolics and hence the sort of elementary convergence groups that arose as in Example 4.2. We say that G is of *compact type* if $O(G)/G$ is compact. A relevant question is

Suppose that G is a (finitely generated?) discrete convergence group with $O(G) \neq \emptyset$. Show G is of compact type if and only if G contains no parabolics.

The second result of Freedman and Skora is a quasiconformal example similar to Theorem 5.3.

Theorem 5.5. *For every $n \geqslant 3$ there is an $r > 1$ and a smooth uniformly quasiconformal group G of compact type, acting on $\mathbb{S}^n$ and having the following properties*:

(i) *G is algebraically isomorphic to the semidirect product of a free group of rank r and a cyclic group of order $2r$.*
(ii) *$L(G)$ is a wild Cantor set.*
(iii) *G contains only loxodromic and elliptic elements and every element of G is quasiconformally conjugate to a Möbius transformation.*
(iv) *G is not the topological conjugate of any Möbius group.*

This example is good because it does not rely on the existence of a fake elliptic and it has simple algebraic structure. Unfortunately there is still the problem of an elliptic. Also we know that r is necessarily large in this example (is this really necessary in general?). The fact that in both of the examples of Theorems 5.3 and 5.5 the limit set is a wild Cantor set (as evidenced by the fact that its complement is not simply connected) led Freedman and Skora to raise the hope that this was enough to guarantee that a convergence (or quasiconformal) group was not conjugate to a Möbius group. Unfortunately we now know this is not the case as D. Cooper and M. Bestvina [BC] have constructed a Schottky group satisfying (i) and (ii) of Theorem 5.3. However Freedman and Skora's results do raise the problem

Find invariants of Cantor sets which imply they cannot be the limit of Schottky groups. Then if possible construct convergence (or quasiconformal) groups with these sets as limit sets.

As Freedman and Skora point out, the example of [BC] shows such invariants must be subtle. In particular it is not known if the existence of the elliptic in the example of Theorem 5.5 is necessary. That is whether the free part of the action alone is not conjugate. Hopefully subtle invariants may be used to decide that in a very real sense the limit set of the example is wilder than the limit sets of the so far constructed Schottky groups. In any case this still leaves the problem

Construct a discrete uniformly quasiconformal group of $\mathbb{S}^n$, $n \geqslant 3$, which is torsion free and not topologically conjugate to a Möbius group.

It was pointed out to me by R. Skora that there is still the possibility that every convergence group of $\mathbb{S}^n$ whose limit set is a Cantor set arises from some Schottky group via an equivariant cellular decomposition (not necessarily supported on the limit set). Is this possible?

To give a rough idea of what is going on in these examples we outline the constructions. Recall first that the natural way to construct a classical Schottky group is to take a collection of pairs of round hyperspheres, none of which separates the collection. For each pair choose the (essentially) unique orientation preserving Möbius transformations that takes one of the paired spheres to the other. The group generated by these pairings is a Schottky group with limit set a Cantor set. This construction works more or less because the inside and outside of a sphere are homeomorphic. It was Freedman and Skora's observation that one could play the same game with tori (the inside and outside of a torus in $\mathbb{S}^3$ are homeomorphic). One could take a collection of paired tori (no one separating the collection) and look at the group generated by a suitable choice of homeomorphisms of $\mathbb{S}^3$ identifying elements of the pairing. This is not so easy to do conformally as attached to a torus is the conformal-invariant of its modulus. This construction of identifying tori provides nothing new unless one chooses the tori so that some linking occurs. Then one finds that the homeomorphism which maps the inside of one of the paired tori A to the outside of the other B causes the remaining tori to form a nontrivial link inside B. This process repeats itself as we look at other elements of the group. We find a construction going on like that of Antoine's necklace [Ru]. In general one must be a little careful to arrange that the limit set is indeed a Cantor set. If the linking is very tight (few tori all linked), then there is so much stretching (to ensure that the image of the link wraps around the core of the torus under the identifying homeomorphism) and squeezing (to ensure that everyting fits in) that the group cannot be quasiconformal. It cannot be made quasiconformal because this stretching and squeezing pattern is topological in nature (the linking pattern of tori is preserved by any homeomorphism). This is the key to the first example. If one makes the linking less severe by using a large number of tori one can arrange that the amount of distortion stays bounded. The analogy is with a coarse necklace with few pieces, each piece must bend a lot for the whole thing to bend a lot, and a fine necklace, each pieces suffers little even when one bends the whole. At this stage it would be nice to conclude that the limit set built out of this linking pattern (in the quasiconformal case) is not the limit set of a conformal group. Unfortunately this has not been done. However if one is careful in the construction and symmetrically places the linked tori one can extend the group by a finite rotation group preserving this symmetry. We then use the linking pattern to show the fixed point circles of certain conjugates of the elements of the group preserving the symmetry also form a link and that this link cannot be realised by the linking of round circles (which must be the fixed points of Möbius elliptics). This shows the group is not topologically conjugate to a Möbius group and explains the presence of the elliptic in Freedman and Skora's second example.

Having provided a number of counterexamples to the general question we shall now go on to look at what can be said in specific circumstances for a uniformly quasiconformal group G. The results are actually quite interesting and useful. The most powerful tool is the construction of a G-invariant ellipse field (as it is in dimension two). One obtains conjugacy results by showing this invariant field is the pullback of the standard field (of round spheres in the tangent space) under a quasiconformal mapping (an existence theorem) or by somehow linearising this field to find a new invariant field which is the pullback of the standard field by an affine transformation. Since this idea is central to essentially all that is known in higher dimensions we will provide a reasonably detailed account.

We define the space $S(n) = SL(n, R)/SO(n)$. Thus $S(n)$ is the space of real, symmetric, positive definite $n \times n$ matrices with determinant equal to one. The general linear group $GL(n, R)$ acts transitively on the right of $S(n)$ via the rule

$$X[A] = |\det X|^{-2/n} X^t A X, \qquad X \in GL(n, R), \qquad A \in S(n).$$

The Riemannian metric,

$$ds^2 = (1/2)\sqrt{n}\,\mathrm{tr}\,(Y - 1dY)^2$$

on $S(n)$ gives rise to a metric distance which we denote by $d(A, B)$ for $A, B \in S(n)$. This metric is invariant under the right action of $GL(n, R)$ and makes $S(n)$ a globally symmetric Riemannian manifold, which is complete, simply connected, and of nonpositive sectional curvature, see [He] (we include the factor $\sqrt{n}/2$ so that when $n = 2$ we get hyperbolic space of constant curvature -1). Then

$$d(A) = d(\text{Identity}, A) = (1/2)\sqrt{n}\,((\ln \lambda_1)^2 + (\ln \lambda_2)^2 + \cdots + (\ln \lambda_n)^2)^{1/2},$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A, see [Maa]. Other distances can now be calculated from the transitivity of the $GL(n, R)$ action.

The *matrix dilatation* of a homeomorphism f of $\mathbb{S}^n$, which is differentiable with nonzero Jacobian matrix almost everywhere, is the measurable map $\mu_f: \mathbb{S}^n \to S(n)$ defined by

$$\mu_f(x) = |\det f'(x)|^{-2/n} f'(x)^t f'(x) = f'(x)[\mathrm{Id}]$$

where $f'(x)$ is the Jacobian matrix of f and Id denotes the identity.

A *conformal structure* on $\mathbb{S}^n$ is a pair $(\mathbb{S}^n, \mu)$, where $\mu: \mathbb{S}^n \to S(n)$ is a measurable mapping for which

$$\operatorname{ess\,sup}\{d(\mu(x)): x \in \mathbb{S}^n\} = d(\mu) < \infty.$$

We define $D(\mu) = \exp(d(\mu))$ and call this the dilatation of μ. The standard conformal structure is $\beta(x) \equiv \mathrm{Id}$. As we have said, one should think of a conformal structure as a measurable ellipse field on the tangent space to the sphere, such that the eccentricity of every ellipse is uniformly bounded. A homeomorphism of $\mathbb{S}^n$ viewed as a mapping between two conformal structures μ_1 and μ_2

$$f\colon (\mathbb{S}^n, \mu_1) \to (\mathbb{S}^n, \mu_2)$$

is called $D(\mu_1, \mu_2)$-quasiconformal if

(1) $f \in W^{1,n}_{\mathrm{loc}}$, *i.e.* f has locally L^n integrable first derivatives, and
(2) $D(f, \mu_1, \mu_2) = \operatorname{ess\,sup}\{\exp(d(\mu_1(x), f'(x)[\mu_2(f(x))])) : x \in \mathbb{S}^n\} < \infty$.

If $\mu_1 = \mu_2 = \beta$, the standard structure, we obtain the usual notion of quasiconformality and if no conformal structures are present we mean quasiconformal in this usual sense. In this case the quantity $\exp(d(f'(x)[\mathrm{Id}]))$ is often referred to as the Ahlfors-Earle dilatation of a quasiconformal mapping at a point x, see [Ah 1]. We call the essential supremum of $D(f, \mu_1, \mu_2)$ over $\mathbb{S}^n$ the D-dilatation of f and we will denote this quantity $D(f)$. Notice that by definition $D(f) = D(\mu_f)$. We say that f is *conformal* if $D(f, \mu_1, \mu_2) = 1$, that is

$$\mu_1(x) = f'(x)[\mu_2(f(x))]$$

almost everywhere.

For the basic facts concerning quasiconformal structures, see [Tu 1, Sec. D]. One should notice that a mapping which is quasiconformal in one structure is automatically quasiconformal in all other structures, however the dilation will vary from structure to structure. Recall too the Louiville theorem that a one-quasiconformal in the usual sense is conformal in the usual sense and so a Möbius transformation.

The notion of D-dilatation is especially useful in our situation. The usual notion (that is of our reference [Vä]) of K-quasiconformality of a homeomorphism f is

$$K(f) = \operatorname{ess\,sup}\{\log(\lambda_{\max}(x)), \log(1/\lambda_{\min}(x))\}^{n/2}$$

where f satisfies (1), (2) and $\lambda_{\max}(x)$ and $\lambda_{\min}(x)$ are the largest and smallest eigenvalues of the matrix $f'(x)[\mathrm{Id}]$. Here is the sharp relation between the D-dilatation and the K-dilatation, see [Tu 1, (D5)].

$$(3) \qquad \frac{K(f)}{\sqrt{n-1}} \leqslant D(f) \leqslant K(f) \quad (n \text{ even}) \quad \text{or} \quad \frac{K(f)}{\sqrt{\frac{n-1}{n}}} \quad (n \text{ odd}).$$

Notice the implication that

$$K(f) = D(f) \quad \text{when} \quad n = 2.$$

For a quasiconformal group G we set

$$D(G) = \sup \{D(g): g \in G\}.$$

Similarly we define $K(G)$. For a quasiconformal group G, a *G-invariant conformal structure* μ on $\mathbb{S}^n$ is a conformal structure $(\mathbb{S}^n, \mu)$ such that each g in G is conformal as a map

$$g: (\mathbb{S}^n, \mu) \to (\mathbb{S}^n, \mu).$$

We observe that if g is quasiconformal, then

$$h: (\mathbb{S}^n, \mu_h) \to (\mathbb{S}^n, \beta) \quad \text{and} \quad h^{-1}: (\mathbb{S}^n, \beta) \to (\mathbb{S}^n, \mu_h)$$

are conformal. Hence if μ is a G-invariant conformal structure and h is a homeomorphism satisfying (1) and (2) and $\mu_h(x) = \mu(x)$ almost everywhere, then the group $h \circ G \circ h^{-1}$ is a conformal group.

One can always construct a G-invariant quasiconformal structure for a quasiconformal group g. This is due to the $GL(n, R)$ invariance of the metric distance d in $S(n)$ and to the fact that $S(n)$ is nonpositively curved. In fact if $P(E)$ denotes the center of the smallest ball containing a bounded subset E of $S(n)$ (which is well defined and unique since $S(n)$ is nonpositively curved and simply connected), then

$$\mu(x) = P(\{\mu_g(x): g \in G\})$$

will be the desired G-invariant conformal structure, see [Tu. 1, Sections D and E]. A simple geometric argument based on the law of cosines in a negatively curved manifold implies that $D(\mu) \leqslant D(G)^{1/\sqrt{2}}$. In higher dimensions ($n \geqslant 3$) there is no measurable Riemann mapping theorem and we will have to use different techniques to produce a conjugating mapping other than appealing to an existence theorem.

The first idea is due to Gromov [Gr 1] and Tukia [Tu 1]. It is this: suppose that μ is a G-invariant conformal structure and that there is a limit point x_0 such that μ is continuous at x_0. We may assume that $x_0 = 0$ and that G acts on $\overline{\mathbb{R}}^n$. Let $\{g_i\}$ be a sequence in G such that $g_i \to 0$ locally uniformly in $\overline{\mathbb{R}}^n - \{y_0\}$ (recall G has the convergence property!). Suppose for the moment that $0 \neq y_0$ and so we can assume (without loss of generality) that $y_0 = \infty$. Let $\{\lambda_i\}$ be a sequence of real numbers such that $|g_i(\lambda_i e_1)| = 1$. Then $\{g_i\}$ forms a normal family and there is a quasiconformal homeomorphism h such that (for some subsequence) $g_i(\lambda_i x) \to h(x)$. The group $\lambda_i^{-1} G \lambda_i$ has the invariant conformal structure

$$\mu_i(x) = \mu\left(\frac{x}{\lambda_i}\right) \to A,$$

where A is some constant matrix. But

$$\lambda_i^{-1} G \lambda_i = \lambda_i^{-1} g_i^{-1} G g_i \lambda_i \to h^{-1} G h$$

and this last group is conjugate to a Möbius group since it has the constant matrix A as an invariant conformal structure. Consequently G is conjugate to a Möbius group.

Now of course one must tidy up all the details. A key result is Tukia's generalization of the good approximation theorem, see [Tu 1, Cor. D].

Theorem 5.6. (Good approximation theorem.) *Let $f_i: U \to \mathbb{S}^n$ be a sequence of K-quasiconformal embeddings. Suppose that $f_i \to f$ for some embedding $f: U \to \mathbb{S}^n$ and that $\mu_{f_i} \to \mu$ in measure for some measurable map $\mu: U \to S(n)$. Then f is K-quasiconformal and $\mu_f = \mu$ a.e. in U.*

If the invariant conformal structure is continuous at a limit point then the above outline works (the assumption that $x_0 \neq y_0$ is unnecessary). Of course we have produced the invariant conformal structure from general methods and it is quite unlikely that it will be continuous anywhere on the limit set (where the action is highly mixing). However the argument is sufficiently robust that all that is necessary is that μ be continuous in measure at a limit point x_0 where we can assert that there is a sequence $g_i \to x_0$ and $g_i^{-1} \to y_0$ with $x_0 \neq y_0$. Such a limit point is called a point of approximation or a *conical limit point*, see [MT]. This last term is because if the action extended to the ball then the orbit of the origin would meet some Stoltz cone based at x_0 infitely often. Since a measurable map is continuous almost everywhre we are more or less done if we assume that the limit set has positive measure (for instance if it is the whole sphere) and that there are plenty of conical limit points. If the action extended to the ball then the hypothesis that $\text{Int}(B^n)/G$ is compact easily implies every point is a conical limit point. Alternatively, if G acts on the triple space T_n so that T_n/G is compact then every limit point is a conical limit point [Tu 1, Cor. G]. We can put all this together to get the following theorem and its corollaries.

Theorem 5.7. *Let G be a quasiconformal group of $\mathbb{S}^n$ and μ a G-invariant conformal structure. If μ is continuous at a limit point of G or continuous in measure at a conical limit point of G, then g is quasiconformally conjugate to a Möbius group by a quasiconformal homeomorphism h satisfying $D(h) \leqslant D(G)^{1/\sqrt{2}}$.*

Corollary 5.8. *Let G be a quasiconformal group of $\mathbb{S}^n$ such that either*

(1) *G extends to B^{n+1} and $\operatorname{int}(B^{n+1})/G$ is compact, or*
(2) *G acts on the triple space T_n so that T_n/G is compact.*

Then G is quasiconformally conjugate to a Möbius group.

As an application of these ideas together with some results on the sort of domains in $\mathbb{R}^n$ that admit quasiconformal actions we find a generalisation of a theorem of Gromov [Gr 1, p. 209], see [Ma 3].

Corollary 5.9. *Let M be a compact n-dimensional manifold (or orbifold) whose universal cover quasiconformally embeds in $\mathbb{S}^n$ and has at least one locally quasiconformally flat boundary point. Then the universal cover of M is quasiconformally the ball and M has the homotopy type of a hyperbolic manifold (respectively orbifold).*

The example $\mathbb{S}^n \times \mathbb{S}^1$ is relevant to Corollary 5.9. The universal cover is $\mathbb{S}^n \times \mathbb{R}^1$ which embedds in $\mathbb{S}^{n+1}$ but $\mathbb{S}^n \times \mathbb{S}^1$ does not admit a hyperbolic structure. The point is that there are two natural ways to embedd $\mathbb{S}^n \times \mathbb{S}^1$ in $\mathbb{S}^{n+1}$. The first is as $\mathbb{S}^n \times (0, 1)$ which has smooth (and therefore locally quasiconformally flat boundary). This embedding cannot be quasiconformal. The only quasiconformal embedding has image $\mathbb{S}^{n+1} - \{0, \infty\}$ which does not have flat boundary.

Theorem 5.7 essentially says that as soon as the limit set of a quasiconformal group G is large enough, then G is quasiconformally conjugate to a Möbius group. We now turn to consider the case where the limit set is small, namely the elementary groups. We will see that smoothness (which played no role in the above) and algebraic simplicity imply quasiconformal conjugacy (as long as there is an element of infinite order).

The following results are from [Ma 2]. It is quite important to recall that although G may be a group of diffeomorphisms the G-invariant conformal structure is almost certainly not very nice. It is easy to construct a loxodromic quasiconformal transformation which is infinitely smooth at both its limit points and yet the invariant ellipse field above is not even approximately continuous there.

Definition 5.10. *A self homeomorphism f of $\mathbb{S}^n$ is said to be* ***affine*** *at x_0 if f is differentiable almost everywhere and there is a matrix $A \in S(n)$ such that*

$$d(\mu_f(x), A) \to 0 \quad \text{as} \quad x \to x_0 .$$

Notice that a diffeomorphism will always be affine at every point.

Theorem 5.11. *Let f be a parabolic quasiconformal homeomorphism of $\mathbb{S}^n$ which is affine at its fixed point. Then f is conjugate to a parabolic Möbius transformation by a K-quasiconformal mapping for which*

$$K(\langle f\rangle)^{1/2} \leqslant K \leqslant \sqrt{(n-1)}\, K(\langle f\rangle)^{1+1/\sqrt{2}}.$$

PROOF. Work in $\bar{\mathbb{R}}^n$ and assume the fixed point is infinity. Choose similarity maps $\beta_m = a_m x + b_m$, where a_m is real and b_m lies in $\mathbb{R}^n$, such that

$$f^n \circ \beta_m(0) = 0 \quad \text{and} \quad |f^m \circ \beta_m(e_1)| = 1.$$

Now since the mapping f^m is $K(\langle f\rangle)$-quasiconformal for all m and since every β_m is conformal, the sequence $\{f^m \circ \beta_m\}$ is normal and so contains a uniformly convergent subsequence converging to a $K(\langle f\rangle)$-quasiconformal group of affine transformations. To see this note that for all integers k

$$g^k = h^{-1} \circ f^k \circ h = \lim (\beta_m^{-1} \circ f^{-m} \circ f^k \circ f^m \circ \beta_m) = \lim (\beta_m^{-1} \circ f^k \circ \beta_m)$$

as $m \to \infty$, by the uniform convergence. Consequently g generates a $K(\langle f\rangle)$-quasiconformal group. We now show that g is affine. To do this we will compute its matrix dilatation and show that it is constant (note $g(\infty) = \infty$). Let $g_j = \beta_m^{-1} \circ f \circ \beta_m$ so

$$\mu_{g_j}(x) = \mu_{f \circ \beta_m}(x) = \beta_m'(x)[\mu_f(\beta_m(x))] = \mu_f(\beta_m(x)).$$

The assumption that f is affine at infinity now implies that there is an $A \in S(n)$ such that

$$\mu_{g_j}(x) = \mu_f(\beta_m(x)) \to A, \quad \text{as} \quad j \to \infty.$$

From the good approximation theorem we conclude that $\mu_g \equiv A$. That is, g has constant matrix dilatation. Finally the affine, uniformly quasiconformal group $\langle g\rangle$ is necessarily quasiconformally conjugate to a Möbius group by a linear mapping B whose dilatation is no more than $\sqrt{(n-1)}\, K(\langle g\rangle)^{1/\sqrt{2}}$. Thus $h \circ B$ is a $\sqrt{(n-1)}\, K(\langle f\rangle)^{1+\sqrt{2}}$-quasiconformal mapping conjugating f to a parabolic Möbius transformation.

A similar argument gives us the loxodromic case

Theorem 5.12. *Let f be a loxodromic quasiconformal transformation of $\mathbb{S}^n$ whose matrix dilatation is approximately continuous at a fixed point. Then f is conjugate to a loxodromic Möbius transformation by a K-quasiconformal mapping for which*

$$K(\langle f\rangle)^{1/2} \leqslant K \leqslant \sqrt{(n-1)}\, K(\langle f\rangle)^{1+1/\sqrt{2}}.$$

The algebraic idea we need in order to generalise this line of argument is the concept of admissibility

Definition 5.13. *Let G be an abstract group. If g and h are elements of G we define the* **commutator** *of g and h as $[g,h] = g^{-1}h^{-1}gh$. If H is a subset of G we define the commutator of G and H as $[G,H] = \{[g,h]: g \in G \text{ and } h \in H\}$. Realize that $[G,H]$ is just a subset of G and not a subgroup even if H is a subgroup. The* **center** *of G, $Z(G)$, is the largest subset of G such that $[G,Z(G)] = \{\text{Identity}\}$. We say that a subgroup H of G is* **virtually central** *if the set* [G, H] *is finite.*

Definition 5.14. *An abstract group G is called* **admissible** *if there is an infinite cyclic subgroup which is virtually central.*

Here is the sort of result we have been seeking [Ma 2, Thm. 4.5].

Theorem 5.15. *Let G be a discrete, admissible group of quasiconformal diffeomorphisms of $\mathbb{S}^n$. Then G is conjugate to a Möbius group by a K-quasiconformal homeomorphism for which*

$$K(G)^{1/2} \leqslant K \leqslant \sqrt{(n-1)}\, K(G)^{1+1/\sqrt{2}}.$$

PROOF. Since G is admissible there is $f \in G$ such that f is of infinite order and $\langle f \rangle$ is virtually central. Thus $F = [G, \langle f \rangle]$ is a finite set. Because G is a discrete quasiconformal group f is either parabolic or loxodromic and we may assume that f fixes infinity and that if f is loxodromic then infinity is the repulsive fixed point. We proceed as above to find conformal maps β_m such that (for some subsequence) $\beta_m \circ f^m$ converging uniformly to a $K(G)$-quasiconformal mapping h. As before $h^{-1} \circ f \circ h$ will be an affine mapping. Let $g \in G$. We show $h^{-1} \circ g \circ h$ is affine. Again, by the uniform convergence we see

$$g' = h^{-1} \circ g \circ h = \lim (\beta_{m(j)}^{-1} \circ f^{-m(j)} \circ g \circ f^{m(j)} \circ \beta_{m(j)}).$$

Now as $\langle f \rangle$ is virtually central we see that for all $m(j)$ there is $h_{m(j)} \in F$ such that the mapping

$$f^{-m(j)} \circ g \circ f^{m(j)} = g \circ h_{m(j)}.$$

Since F is a finite collection of mappings, we may pass to another subsequence so that

$$f^{-m(j)} \circ g \circ f^{m(j)} = g \circ h$$

for some fixed h in the set F. Then

$$g' = h^{-1} \circ g \circ h = \lim (\beta_{m(j)} \circ g \circ h \circ \beta_{m(j)}^{-1})$$

is a $K(G)$-quasiconformal homeomorphism of $\mathbb{R}^n$. Calculating the matrix dilatations of elements of this sequence and using the fact that the matrix dilatation of $g \circ h$ is continuous everywhere, and so in particular at infinity, then applying the good approximation theorem as we did in Theorem 5.11 we see as before that g' is a $K(G)$-quasiconformal affine mapping. Since g was arbitrary, the group $h^{-1} \circ G \circ h$ is affine and quasiconformal and thus conjugate to a conformal group via a linear mapping.

Corollary 5.16. *Let G be a finitely generated infinite discrete quasiconformal group of diffeomorphisms of $\mathbb{S}^n$. If G is abelian, then G is quasiconformally conjugate to a discrete Möbius group. Consequently the rank of G is at most n.*

We now turn to consider the hypotheses of Theorem 5.15 and to give examples to show that they are essentially all necessary. As we saw earlier the assumption that there is an element of infinite order in the group is necessary, when $n \geqslant 4$, due to the existence of smooth counterexamples to the generalized Smith conjecture. The examples of Freedman and Skora show that some geometric or algebraic restriction is necessary. We raise the following question though

Can the hypothesis of admissibility in Theorem 5.15 be replaced by the hypothesis that G is virtually abelian?

Notice that these two hypotheses are independent. It may be that this is not the case, it seems possible that there is a smooth quasiconformal group containing an exotic involution and isomorphic to the Dihedral group.

We recall Theorem 5.2 which was an example of a uniformly quasiconformal group of $\mathbb{S}^n$ isomorphic to $\mathbb{Z}^{n-1}$ (and so of course admissible) which is not quasiconformally conjugate to a Möbius group. As remarked in that paper [Ma 1], the group can be made smooth except at one point. It cannot be made smooth over this last point by our results above. By J. McKemie's results [McK] the dilatation of this group can be assumed arbitrarily close to one.

It is worth observing the following consequence of Theorem 5.15. Recall we saw earlier that fake elliptics can lie in some infinite discrete quasiconformal groups. However we see.

Proposition 5.17. *Let σ be a periodic diffeomorphism of $\mathbb{S}^n$ of period p. If σ is not topologically conjugate to an orthogonal transformation, then there is no discrete smooth $\mathbb{Z} \times \mathbb{Z}_p$ action on $\mathbb{S}^n$ which contains σ and is of bounded distortion.*

Actually of course, a periodic diffeomorphism which is not conjugate to an orthogonal transformation, can lie inside no finitely generated infinite abelian group of bounded distortion acting smoothly on $\mathbb{S}^n$.

In a different vein using more topological methods we conclude this section by recalling a result of P. Tukia and the author regarding the conjugacy of certain convergence groups in the three dimensional case. The main idea of the arguments used is that if G is a convergence group acting on $\bar{B}^3$ and which is of compact type, then the quotient three manifold can have no incompressible tori. It follows then (by Thurston's theorem, see [MB]) that the quotient manifold admits a hyperbolic structure if for instance the manifold has boundary or more generally is Haken. This produces the desired conjugacy to a Möbius group on the ordinary set. There is in general some problem in extending the conjugacy over the limit set, but this is automatic if G were for instance uniformly quasiconformal or if $L(G)$ were a Cantor set.

Theorem 5.18. *Let G be a torsion free uniformly quasiconformal group acting on $\bar{B}^3$ which is of compact type and of the second kind (or if G is of the first kind assume B^3/G is Haken). Then g is quasiconformally conjugate to a Möbius group.*

Theorem 5.19. *Let G be a torsion free convergence group acting on $\bar{B}^3$ which is of compact type and whose limit set is a Cantor set. Then G is topologically conjugate to a Möbius group.*

There is a suggested outline of how to extend the above results to the geometrically finite case in [MT]. The hope is to prove the general conjecture.

A uniformly quasiconformal group acting on B^3 is quasiconformally conjugate to a Möbius group.

Note that from the results of Section 4, the action on the boundary may be assumed conformal.

6. Leftovers

There is much that we have left unsaid. As an example one can define a Poincaré series for a discrete quasiconformal group acting on B^n [GM 2]. The con-

vergence and divergence of this series depends on the structure of the limit set. The radial limit set can be defined and was shown by Gehring, Garnett and Jones [GGJ] to satisfy the classical dichotomy of zero or full measure when $n \geqslant 2$. J. Garnett has raised the following question when $n = 1$ (recall that unlike quasiconformal mappings quasisymmetric mappings need not preserve Lebesgue null sets).

Let G be a quasisymmetric group of the circle. Is the measure of the conical limit set zero or full?

One can define what it means for a quasiconformal group to be geometrically finite and correspondingly solve Ahlfors' measure zero problem in this case using essentially no properties of analyticity.

Other questions regarding discreteness criteria (such as Jøgensen's inequality) and universal constraints (as are imposed by discreteness on Fuchsian groups) may be raised for quasiconformal groups in all dimensions.

Another project would be to more closely relate Gromov's theory of hyperbolic groups [Gr 2] to that of convergence groups. What are the consequences for convergence groups (or for hyperbolic groups)?

We hope that the questions and conjectures raised in this paper and others will motivate further investigation into convergence and quasiconformal groups.

References

[AM] Abikoff, W. and Maskit, B. Geometric Decomposition of Kleinian Groups, *Amer. J. Math.*, **99**(1974), 687-697.

[Ah 1] Ahlfors, L. V. A Somewhat New Approach to Quasiconformal Mappings in $\mathbb{R}^n$, *Lecture Notes in Math.* **599**, Springer-Verlag, (1977).

[Ah 2] —, Finitely Generated Kleinian Groups, *Amer. J. Math.*, **86**(1964), 413-429.

[Ah 3] —, Möbius Transformations in Several Dimensions, *Lecture notes. School of Math.*, Univ. of Minnesota (1981).

[AS] Ahlfors, L. V. and Sario, L. Riemann surfaces, Princeton Univ. Press, (1960).

[Bea] Beardon, A. The Geometry of Discrete Groups, Springer-Verlag, (1983).

[Be] Bers, L. On Ahlfors' Finiteness Theorem, *Amer J. Math.*, **89**(1967), 1078-1082.

[BC] Bestvina, M. and Cooper, D. A wild Cantor set as the limit set of a Schottky Group. *Proc. Amer. Math. Soc.* To appear.

[DE] Douady, A. and Earle, C. Conformally Natural Extensions of Homeomorphisms of the Circle. *Acta Math.*, **157**(1986), 23-48.

[Du] Dunwoody, M. J. Accessibility of finitely presented groups, *Invent. Math.*, **81**(1985), 449-457.

[EO] Eberlin, P. and O'Neil, B. Visibility Manifolds, *Pac. J. Math.*, **46**(1973), 45-109.

[Ei] Eilenberg, S. Sur les transformations périodiques de la surface du sphere, *Fund. Math.*, **22**(1934), 28-41.

[FM] Feighn, M. and McCullough, D. Finiteness Conditions for Three Manifolds with Boundary, *Amer. J. Math.*, **109**(1987), 1155-1169.

[Fr] Freedman, M. A Geometric Reformulation of Four Dimensional Surgery, *Top. Appl.*, **24**(1986).

[FS] Freedman, M. and Skora, R. Strange actions of groups on spheres. *J. Diff. Geometry*, **25**(1987), 75-98.

[Ga] Gardiner, F. P. Teichmüller Theory and Quadratic Differentials, *Pure and Appl. Math.*, Wiley Interscience (1987).

[Ge] Gehring, F. W. Rings and Quasiconformal Mappings in space, *Trans. Amer. Math. Soc.*, **103**(1962), 353-393.

[GGJ] Gehring, F. W., Garnett, J. and Jones, P. Quasiconformal Groups and the Conical Limit Set, *Holomorphic Functions and Moduli*, Springer-Verlag (1988), 59-68.

[GM 1] Gehring, F. W. and Martin, G. J. Discrete Quasiconformal Groups I. *Proc. London Math. Soc.*, **55**(1987), 331-358.

[GM 2] Gehring, F. W. and Martin, G. J. Discrete Quasiconformal Groups II, to appear.

[GM 2] —, Discrete Convergence Groups. *Lecture Notes in Math.* **1275**. Springer Verlag (1987), 158-168.

[GP] Gehring, F. W. and Palka, B. Quasiconformally Homogeneous Domains, *J. Analyse Math.*, **30**(1976), 172-199.

[Gi] Giffen, C. H. The Generalized Smith Conjecture. *Amer. J. Math.* **88**(1966), 187-198.

[Gre] Green, L. The Generalised Geodesic Flow, *Duke Math. J.*, **41**(1974), 115-126.

[Gr 1] Gromov, M. Hyperbolic Manifolds, Groups and Actions, *Ann. Math. Stud.* **97**. Princeton Univ. Press. 183-213.

[Gr 2] —, Hyperbolic groups, *Essays in Group Theory*, Springer-Verlag (1987).

[Gr 3] —, Structures Métriques pour les Variétés Riemanniannes. *Textes Mathématiques.* Cedric/Fernand Nathan (1981).

[GT] Gromov, M. and Thurston, W. Variable Negative Curvature and Constant Negative Curvature, *Invent. Math.*, to appear.

[He] Helgason, S. Differential Geometry, Lie Groups and Symmetric Spaces, Academic Press, (1978).

[Hi] Hinkkanen, A. Uniformly Quasisymmetric Groups. *Proc. London Math. Soc.*, **51**(1985), 318-338.

[Hi 2] —, Abelian and non-discrete convergence groups of the circle. To appear.

[HM] Hinkkanen, A. and Martin, G. J. Nondiscrete Convergence Groups in the Plane, to appear.

[HK] Homma, T. and Kinoshita, S. On a topological characterisation of the dilation in E^3, *Osaka Math. J.*, **6**(1954), 135-143.

[HP] Hirsch, M. and Pugh, C. Smoothness of Horocycle Foliations, *J. Diff. Geometry,* **10**(1975), 225-238.

[Hu 1] Husch, L. A Homotopy Theoretic Characterisation of the Translations in E^n, *Compositio Math.*, **24**(1972), 55-61.

[Hu 2] Husch, L. A Topological Theoretic Characterisation of the Dilation in E^n, *Proc. Amer. Math. Soc.*, **28**(1971), 234-236.

[Ker] Kerchoff, S. The Nielsen Realisation Problem, *Annals of Math.*, **117**(1983), 235-265.

[Ke 1] B. von Kerékjártó. Über die fixpunktfreien Abbildungen der Ebene, *Acta Litt. Sci. Szeged*, **6**(1934), 226-234.
[Ke 2] —, Topologische Charakterisierung der linearen Abbildungen, *Acta Litt. Sci. Szeged*, **6**(1934), 235-262.
[Kr] Kra, I. On the Cohomology of Kleinian Groups, *J. Analyse Math.*, **43** (1983/84), 51-87.
[Ku] Kulkarni, R. Infinite Regular Coverings and Kleinian Groups, *Proc. Symp. Pure Math.*, **32**(1978), 407-412.
[KS] Kulkarni, R. and Shalen, P. On Ahlfors' Finiteness Theorem, *Advances in Math.*
[Le] Lehto, O. Univalent Functions and Teichmüller Spaces, Springer-Verlag.
[Li] Linnell, P. A. On Accessibility of Groups, *J. Pure Appl. Alg.*, **30**(1983), 39-46.
[Maa] Maass, H. Siegel's Modular Forms and Dirichlet Series. *Lecture Notes in Math.*, **216**. Springer Verlag (1971).
[Ma 1] Martin, G. J. Discrete Quasiconformal Groups that are not the Quasiconformal Conjugates of Möbius Groups, *Ann Acad. Sci. Fenn. Ser.*, A.I., **11**(1986), 179-202.
[Ma 2] Martin, G. J. On Quasiconformal and Affine Groups, *J. Diff. Geometry*, **29**(1989), 427-448.
[Ma 3] Martin, G. J. Quasiconformal Actions on Domains in Space, *Holomorphic Functions and Moduli*, Springer-Verlag, (1988), 87-112.
[MS] Martin, G. J. and Skora, R. Group Actions on the Two Sphere, *Amer. J. Math.*, **111**(1989), 387-402.
[MT] Martin, G. J. and Tukia, P. Convergence and Möbius Groups, *Holomorphic Functions and Moduli*, Springer-Verlag, (1988), 113-140.
[Mas 1] Maskit, B. Kleinian Groups, Springer-Verlag, (1988).
[Mas 2] —, A Characterisation of Schottky Groups, *J. Analyse Math.*, **19**(1967), 227-230.
[McK] McKemie, J. Quasiconformal Groups with Small Dilatation, *Ann. Acad. Sci. Fenn. Ser. AI,* **12**(1987), 95-118.
[Me] Mess, G. Centers of 3-manifold groups and groups which are coarse quasi-isometric to planes. To appear.
[MZ] Montgomery, D. and Zipping, L. Examples of Transformation Groups, *Proc. Amer. Math. Soc.*, **5**(1954), 460-465.
[Mo] Moore, R. L. Concerning Upper-semicontinuous Collections of Continua, *Trans. Amer. Math. Soc.*, **27**(1925), 416-428.
[MB] Morgan, J. W. and Bass, H. The Smith Conjecture, Academic Press, (1984).
[Ne] Newman, M. H. A. A Theorem on Periodic Transformations of Spaces, *Quart. J. Math.*, **2**(1931), 1-8.
[Ru] Rushing, T. B. Topological Embeddings, Academic Press, (1973).
[Na] Nag, S. The Complex Analytic Theory of Teichmüller Spaces, *Canadian Math. Soc. Monographs*, Wiley Interscience (1988).
[Sc] Scott, G. P. The Geometries of Three Manifolds, *Bull. London Math. Soc.*, **15**(1983), 401-487.
[Se] Selberg, A. On Discontinuous Groups in Higher Dimensional Symmetric Spaces, *Int. Colloquium on Function Theory*. Tata Inst. Bombay, (1960).
[Sta] Stallings, J. Group Theory and Three Dimensional Manifolds. Yale Univ. Press., (1971).
[St] Steenrod, N. The Topology of Fiber Bundles, Princeton Univ. Press, (1951).

[Su 1] Sullivan, D. On the Ergodic Theory at Infinity of an Arbitrary Discrete Group of Hyperbolic Motions. *Ann. Math. Stud.* **97**. Princeton Univ. Press, 465-496.
[Su 2] —, Quasiconformal homeomorphisms in Dynamics, Topology and Geometry, *Proc. Int. Cong.*, Berkeley (1986).
[Su 3] —, Finiteness Theorems for Cusps, *Acta Math.*, **147**(1981), 289-299.
[Th] Thurston, W. The Geometry and Topology of Three Manifolds, Notes, Princeton Univ., (1981).
[Tu 1] Tukia, P. On Quasiconformal Groups, *J. Analyse Math.*, **46**(1986), 318-346.
[Tu 2] —, On Two Dimensional Quasiconformal Groups, *Ann. Acad. Sci. Fenn. Ser.*, A.I., **5**(1980), 73-78.
[Tu 3] —, Homeomorphic Conjugates of Fuchsian Groups, *J. Reine Angew. Math.* **391**(1988), 1-54.
[Tu 4] —, A Quasiconformal Group not Isomorphic to a Möbius Group, *Ann. Acad. Sci. Fenn. Ser.*, A.I., **6**(1981), 149-160.
[Vä] Väisälä, J. Lectures on n-Dimensional Quasiconformal Mappings, *Lecture Notes in Math.*, **229**, Springer-Verlag, (1971).
[Wa] Wall, C. T. C. Pairs of Relative Cohomological Dimension One, *J. Pure Appl. Alg.*, **1**(1971), 141-154.
[Wat] Waterman, P. L. Purely Elliptic Möbius Groups. *Holomorphic Functions and Moduli*, Springer-Verlag, (1988), 173-178.

Gaven J. Martin*
Yale University and
University of Auckland
U.S.A.

* Reseach partially supported by a U.S. National Science Foundation Grant.

Revista Matemática Iberoamericana
Vol. 4, n.os 3 y 4, 1988

Comparison Principles and Pointwise Estimates for Viscosity Solutions of Nonlinear Elliptic Equations

Neil S. Trudinger

Abstract

We prove comparison principles for viscosity solutions of nonlinear second order, uniformly elliptic equations, which extend previous results of P.L. Lions, R. Jensen and H. Ishii. Some basic pointwise estimates for classical solutions are also extended to continuous viscosity solutions.

1. Introduction

In this paper we are concerned with some comparison principles and local estimates for viscosity subsolutions and supersolutions of nonlinear uniformly elliptic differential equations of the form,

$$F[u] = F(x, u, Du, D^2u) = 0 \tag{1.1}$$

in domains Ω in Euclidean n-space, $\mathbb{R}^n$. Here the function $F \in C^0(\Gamma)$ where $\Gamma = \Omega \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{S}^n$, $\mathbb{S}^n$ denotes the linear space of real $n \times n$ symmetric

matrices and the operator F is at least *degenerate elliptic*, that is,

$$F(x, z, p, r) \leqslant F(x, z, p, r + \eta), \tag{1.2}$$

for all $x, z, p, r \in \Gamma$, $\eta \geqslant 0$, $\eta \in \mathbb{S}^n$. Following Lions [16], a function $u \in C^0(\Omega)$ is called a *viscosity subsolution* (*supersolution*) of equation (1.1) if for any function $\varphi \in C^2(\Omega)$ and local maximum (minimum) x_0 of $u - \varphi$, we have

$$F(x_0, u(x_0), D\varphi(x_0), D^2\varphi(x_0)) \geqslant 0, \qquad (\leqslant 0). \tag{1.3}$$

A *viscosity solution* is both subsolution and supersolution. Such a notion of solution is implicit in the classical Perron process but was invoked recently by Crandall and Lions [4] in the context of first order equations. Lions [16] treated the existence and uniqueness of viscosity solutions of second order equations in the case that (1.1) is a Bellman equation, arising from stochastic control theory. The first significant breakthrough in the general case was made by Jensen [9], who established comparison principles for Lipschitz solutions when the function F is essentially independent of the x variables. His result was extended to continuous solutions by Jensen, Lions and Souganidis [11] and the latter restriction removed independently by Ishii [7], Jensen himself [10] and the present author (for uniformly elliptic operators) using different methods. The main point of this paper is to demonstrate our approach which employs a combination of Jensen's ideas in [9] and a second derivative estimate for linear equations, due to Lin [15]. By adjoining the technique, used by Ishii [7], of adding more variables we are able to improve the results of Ishii and Jensen in the uniformly elliptic case. This latter technique was also utilized independently by us for regularity considerations [18], [19] and in this context it arose from a classical regularity argument which we supplied in [14, Theorem 1].

In order to describe our comparison principles we formulate structure conditions as follows:

*F*1 (*Uniform ellipticity*)

$$\lambda \operatorname{trace}(\eta) \leqslant F(x, z, p, r + \eta) - F(x, z, p, r) \leqslant \Lambda \operatorname{trace}(\eta);$$

*F*2 (*Lipschitz-continuity in p*)

$$|F(x, z, p, r) - F(x, z, q, r)| \leqslant \mu_1 |p - q|;$$

*F*3 (*Monotonicity in z*)

$$F(x, z, p, r) \leqslant F(x, z + t, p, r);$$

*F*4 (*Continuity in x*)

$$|F(x, z, p, r) - F(y, z, p, r)| \leqslant \mu_2 |x - y|^\delta |r| + \omega(|x - y|);$$

for all $x, y \in \Omega$, $|z| \leqslant K$, $|p|$, $|q| \leqslant K_1$, $r \in \mathbb{S}^n$, $\eta \geqslant 0$, $\eta \in \mathbb{S}^n$, $t, K, K_1 \in \mathbb{R}^+$, where λ, Λ, μ_1, μ_2, δ are fixed positive constants (depending possibly on K) and $\omega(a) \to 0$ as $a \to 0$.

We can then state the following comparison principle.

Theorem 1.1. *Let $u, v \in C^0(\bar{\Omega}) \cap C^{0,1}(\Omega)$ be respectively subsolution and supersolution of equation* (1.1) *in Ω, with $u \leqslant v$ on $\partial\Omega$. Then if the structure conditions $F1$, $F2$, $F3$, $F4$ are satisfied for $\delta = 1/2$, we have $u \leqslant v$ in Ω.*

In fact more general results hold. Namely one can allow $\delta < 1/2$, depending on n, Λ/λ, and if only u or $v \in C^0(\bar{\Omega}) \cap C^{0,\gamma}(\Omega)$, $0 \leqslant \gamma \leqslant 1$, (with $F1 - F4$ holding for all $p, q \in \mathbb{R}^n$), the conclusion holds for some $\delta < 1 - \gamma/2$. In the work of Jensen [10] and Ishii and Lions [8], which extends Ishii [7], the second inequality in $F1$ can be dispensed with but one must have $\delta > 1/2$. The further relaxation of $F4$ seems to be an open problem. Of course if $u, v \in W^{2,n}(\Omega)$, no such condition is required as the usual comparison principle then holds, (for example see [6]). Theorem 1.1 implies a uniqueness result for the Dirichlet problem, previously announced by us in [18, Remark 3.3].

The plan of this paper is as follows. In the following section we develop some auxiliary results for semi-convex subsolutions of linear eliptic equations. These are applied to yield comparison principles in Section 3, initially in the case $\delta = 1$ which we originally obtained. As consequences of Theorem 1.1 we obtain a strong comparison principle and a comparison principle for oblique boundary conditions. The results of Section 3 are extended to Hölder continuous solutions in Section 4. In Section 5 we note some of the pointwise estimates which also arise from the estimates of Section 2. In a further work [20] we show that viscosity solutions of uniformly elliptic equations are twice differentiable almost everywhere. Unless otherwise stated all notation in this paper follows the book [6].

The author is grateful to the Max Planck Institut für Mathematik in Bonn where the research for this paper was completed in 1987, and to P.L. Lions for informing him of related work [7], [8], [11], after we had proved the case $\delta = 1$ in Theorem 1.1.

2. Auxiliary Results for Linear Equations

Jensen's approach to viscosity solutions is based on their approximation by semi-convex subsolutions and semi-concave supersolutions. In this section we consider the appropriate extensions of the linear theory to such subsolutions and supersolutions. We first note that a function u on the domain Ω is called *semi-convex* (*semi-concave*) if $u = v + w$ for some $v \in C^{1,1}(\Omega)$ and convex

(concave) w. Such solutions are twice differentiable almost everywhere in Ω. We consider linear elliptic operators of the form

$$Lu = a^{ij}D_{ij}u + b^iD_iu \tag{2.1}$$

with coefficients a^{ij}, b^i, $i,j = 1, \ldots, n$ satisfying

$$\lambda I \leqslant [a^{ij}] < \Lambda I, \qquad |b^i| \leqslant \mu, \tag{2.2}$$

for positive constants λ, Λ, μ.

Lemma 2.1. *Let u be semi-convex in Ω and suppose $Lu + f \geqslant 0$ almost everywhere in Ω for some $f \in L^n(\Omega)$. We then have the following estimates*:

I. (*Aleksandrov maximum principle.*)

$$\sup_{\Omega} u \leqslant \sup_{\partial\Omega} u + C \operatorname{diam}\Omega \| f/\mathfrak{D}^* \|_{L^n(\Gamma^+)}, \tag{2.3}$$

where C depends on n, $\| b/\mathfrak{D}^* \|_n$ and $\mathfrak{D}^* = (\det [a^{ij}])^{1/n}$, Γ^+ is the upper contact set of u (see [6]);

II. (*Local maximum principle.*) *For any ball $B = B_R(y) \subset \Omega$, concentric subball $B_\sigma = B_{\sigma R}(y)$, $0 < \sigma < 1$, and $p > 0$,*

$$\sup_{B_\sigma} u \leqslant C\left\{\left(R^{-n}\int_B (u^+)^p\right)^{1/p} + R\| f/\lambda \|_{L^n(B)}\right\}, \tag{2.4}$$

where C depends on n, Λ/λ, $\mu R/\lambda$, σ and p;

III. (*Weak Harnack inequality.*) *For any ball $B = B_R(y) \subset \Omega$ and concentric subball B_σ, $0 < \sigma < 1$, there exists a positive number p, depending only on n, Λ/λ, such that*

$$\left(R^{-n}\int_{B_{\sigma R}} (M - u)^p\right)^p < C\left\{\inf_{B_{\sigma R}} (M - u) + R\| f/\lambda \|_{L^n(B)}\right\} \tag{2.5}$$

where $M = \sup_{B_R} u$ and C depends on n, Λ/λ, $\mu R/\lambda$ and σ.

IV. (*Derivative estimates.*) *There exists a positive number $\varkappa$, depending only on n, Λ/λ, such that, for any subdomain $\Omega' \subset\subset \Omega$,*

$$\left(\int_{\Omega'} |Du|^\varkappa + |D^2u|^\varkappa\right)^{1/\varkappa} \leqslant C\left\{\sup_{\Omega} |u| + \| f/\lambda \|_{L^n(\Omega)}\right\} \tag{2.6}$$

where C depends on n, Λ/λ, μ/λ, $\operatorname{diam}\Omega$, $\operatorname{dist}(\Omega', \partial\Omega)$.

The estimates I, II, III are established in the literature for subsolutions $u \in W^{2,n}(\Omega)$; (see for example [2], [6], [17]). The extension of the Aleksandrov estimate (2.3) follows automatically from his work [2] since the normal mapping of semi-convex functions will be absolutely continuous with respect to Lebesgue measure. Alternatively, all of the above estimates may be deduced, from their validity for classical subsolutions, by mollification since the Hessian matrices of the mollifications of a semi-concave function will be uniformly bounded from below.

The estimate IV was proved by Evans [5] and Lin [15], for the first and second derivatives respectively of solutions of the Dirichlet problem. To obtain a local estimate for subsolutions $u \in C^2(\Omega)$ we may first observe from Lin [15] that there exists a positive matrix function $[a^{ij}_*]$, depending on D^2u, such that

$$\frac{\lambda}{2}|D^2u| \leqslant (a^{ij}_* - a^{ij})D_{ij}u, \tag{2.7}$$

$$\frac{\lambda}{2}I \leqslant [a^{ij}_*] \leqslant 2\Lambda I.$$

For $\eta \in C_0^2(\Omega)$, $k < \inf u$ and

$$v = \eta^2(u-k)^2, \tag{2.8}$$

we now compute

$$\begin{aligned} a^{ij}_* D_{ij}v &= 2\eta^2(u-k)a^{ij}_* D_{ij}u + 2\eta^2 a^{ij}_* D_i u D_j u + 8\eta(u-k)a^{ij}_* D_i \eta D_j u \\ &\quad + (u-k)^2 a^{ij}_* D_{ij}\eta^2 \\ &\geqslant \frac{\lambda}{2}\eta^2\{(u-k)|D^2u| + |Du|^2\} \\ &\quad - C(u-k)\{\eta^2(u-k)|b|^2 + \eta^2|f| + \Lambda(u-k)(|D\eta|^2 + |\eta D^2\eta|\} \\ &= \frac{\lambda}{2}\eta^2\{(u-k)|D^2u| + |Du|^2\} - f_*, \end{aligned} \tag{2.9}$$

for some constant C. It then follows from Evans [5] (or Lin [15]) that there exists some positive $\varkappa$ depending only on n, Λ/λ for which

$$\left\{\int_\Omega \eta^{2\varkappa}[(u-k)|D^2u| + |Du|^2]^\varkappa\right\}^{1/\varkappa} \leqslant C\|f_*/\lambda\|_{L^n(\Omega)} \tag{2.10}$$

where C depends on n, Λ/λ and $\operatorname{diam}\Omega$. By appropriate choice of η and k we thus obtain the estimate (2.6). □

3. Comparison Principles for Lipschitz Solutions

In this section we prove Theorem 1.1, along with some applications. Following Jensen [9], we can approximate functions $u \in C^{0,1}(\Omega)$ by setting, for $\epsilon > 0$,

$$\Omega_\epsilon = \{x \in \Omega | \operatorname{dist}(x, \partial\Omega) > \epsilon\},$$

and letting $u_\epsilon^\pm \in C^{0,1}(\Omega_\epsilon)$ be the functions whose graphs have fixed distance ϵ from the graph of u and which lie respectively above and below the graph of u. The graphs of $u_\epsilon^\pm$ will satisfy lower and upper exterior sphere conditions respectively, with spheres of radius ϵ, and therefore, since the functions $u_\epsilon^\pm$ are Lipschitz, they will be semi-convex and semi-concave respectively. Moreover we have the estimates

$$|Du_\epsilon^\pm| \leqslant |Du|_0 \tag{3.1}$$

and

$$\pm D^2 u_\epsilon^\pm \geqslant -\frac{(1 + |Du|_0^2)^{1/2}}{\epsilon}$$

in the sense of distributions. A simple geometric consideration (together with condition $F3$) shows that if u is a viscosity subsolution (supersolution) of equation (1.1) in Ω, then the functions $u_\epsilon^+ - \epsilon$ ($u_\epsilon^+ + \epsilon$) will be a viscosity subsolution (supersolution) with the operator F replaced by F^+ (F^-) given by

$$F^\pm(x, z, p, r) = F(x^\pm, z, p, r), \tag{3.2}$$

where

$$x^\pm = x \pm \epsilon \nu^\pm, \qquad \nu^\pm = \frac{Du_\epsilon^\pm}{\sqrt{1 + |Du_\epsilon^\pm|^2}}$$

is the projection in $\mathbb{R}^n$ of the unit normal to the graph of $u_\epsilon^\pm$. Accordingly, (replacing $u_\epsilon^\pm$ by $u_\epsilon^\pm \mp \epsilon$), we have (see [9], [18], [7]) the differential inequalities,

$$\pm F^\pm[x, u_\epsilon^\pm, Du_\epsilon^\pm, D^2 u_\epsilon^\pm] \geqslant 0 \tag{3.3}$$

almost everywhere in Ω_ϵ.

We first prove Theorem 1.1 for the case $\delta = 1$, which was our original result. If u, v satisfy the hypotheses of Theorem 1.1, we have (almost everywhere) in the set where

$$w_\epsilon \equiv u_\epsilon^+ - v_\epsilon^- \geqslant 0 \tag{3.4}$$

the differential inequality,

$$
\begin{aligned}
(3.5)\quad Lw_\epsilon &\equiv a^{ij}D_{ij}w_\epsilon + b^iD_iw_\epsilon \\
&\geqslant F(x^-, v_\epsilon^-(x), Du_\epsilon^+(x), D^2u_\epsilon^+(x)) - F(x^+, v_\epsilon^-(x), Du_\epsilon^+(x), D^2u_\epsilon^+(x)) \\
&\geqslant -2\mu_2\epsilon|D_2u_\epsilon^+| - \omega(2\epsilon),
\end{aligned}
$$

by $F4$, with coefficients a^{ij}, b^i given by

$$
\begin{aligned}
(3.6)\qquad a^{ij} &= \int_0^1 F_{r_{ij}}^-(x, v_\epsilon^-, Dv_\epsilon^-, D^2(v_\epsilon^- + tw_\epsilon))\,dt, \\
b^i &= \int_0^1 F_{p_i}^-(x, v_\epsilon^-, D(v_\epsilon^- + tw_\epsilon), D^2u_\epsilon^+)\,dt.
\end{aligned}
$$

Using the structure conditions $F1$, $F2$ we see that

$$
\begin{aligned}
(3.7)\qquad \lambda I \leqslant [a^{ij}] &\leqslant \Lambda I, \\
|b| &\leqslant \mu_1.
\end{aligned}
$$

If the constant $\mu_2 = 0$, then Theorem 1.1 follows immediately from the Aleksandrov maximum principle and we have as a special case Jensen's result [9]. Here of course there is no need for the upper bound in (3.7) and the result is readily extended to the case of degenerate ellipticity and a strict monotonicity condition $F2$. To handle the second derivative term on the right hand side of (3.5), we now write the differential inequality (3.3) for u_ϵ^+ in the form,

$$
\begin{aligned}
(3.8)\qquad L_0u_\epsilon^+ &\equiv a_0^{ij}D_{ij}u_\epsilon^+ + b_0^iD_iu_\epsilon^+ \\
&\geqslant -F^+(x, u_\epsilon^+ - \epsilon, 0, 0) \\
&\geqslant -F_0
\end{aligned}
$$

for some constant F_0, where the coefficients a_0^{ij}, b_0^i, given by

$$
\begin{aligned}
(3.9)\qquad a_0^{ij} &= \int_0^1 F_{r_{ij}}^+(x, u_\epsilon^+, Du_\epsilon^+, tD^2u_\epsilon^+)\,dt, \\
b_0^i &= \int_0^1 F_{p_i}^+(x, u_\epsilon^+, tDu_\epsilon^+, 0)\,dt,
\end{aligned}
$$

also satisfy (3.7). From the second derivative estimate, Lemma 2.1, IV, we then have

$$
(3.10)\qquad \int_{\Omega'} |D^2u_\epsilon^+|^\varkappa \leqslant C
$$

for any $\Omega' \subset\subset \Omega$, where $\varkappa$ depends only on n, Λ/λ and C is independent of ϵ. Fixing $\Omega' \subset\subset \Omega_\epsilon$ we now apply the Aleksandrov maximum principle, Lemma 2.1, I, to the function w_ϵ, noting that on the upper contact set Γ^+ of w_ϵ, we have by virtue of (3.1),

$$
(3.11)\qquad |D^2u_\epsilon^+| \leqslant \frac{C}{\epsilon}.
$$

Accordingly, we obtain

$$\sup_{\Omega'} w_\epsilon \leqslant C\{\epsilon \| D^2 u_\epsilon^+ \|_{L^n(\Gamma^+)} + \omega(2\epsilon)\} + \sup_{\partial\Omega'} w_\epsilon^+ \tag{3.12}$$

$$\leqslant C\{\epsilon^{\varkappa/n} + \omega(2\epsilon)\} + \sup_{\partial\Omega'} w_\epsilon^+, \quad \text{by} \quad (3.10)$$

$$\to \sup_{\partial\Omega'} w_\epsilon^+, \quad \text{as} \quad \epsilon \to 0,$$

$$\to 0 \quad \text{as} \quad \Omega' \to \Omega.$$

Consequently $u \leqslant v$ in Ω and Theorem 1.1 is proved in the case $\delta = 1$.

To treat the general case, we extend w as a function of $2n$ variables (x, ξ) on the domain,

$$\bar{\Omega} = \Omega_{2\epsilon} \times B_0(1) \subset \mathbb{R}^{2n},$$

by defining

$$w_\epsilon(x, \xi) = u_\epsilon^+(x) - v_\epsilon^-(x + \epsilon\xi). \tag{3.13}$$

We then obtain (instead of 3.5),

$$Lw_\epsilon \geqslant -\mu_2(3\epsilon)^\delta |D^2 u_\epsilon^+| - \omega(3\epsilon), \tag{3.14}$$

where in the formulae (3.6) for the coefficients a^{ij}, b^i, the arguments of v_ϵ^- are $x + \epsilon\xi$ instead of x. Writing

$$g = |D^2 u_\epsilon^+|^{\varkappa/2n}, \tag{3.15}$$

we may then estimate

$$Lw_\epsilon \geqslant -C\epsilon^{\delta\chi} g |D^2 u_\epsilon^+| - o(\epsilon) \tag{3.16}$$

where $\chi = 1 + \varkappa/2n$ and δ is replaced by any smaller number. Now come some technical subtleties. Defining

$$\sigma^{ij} = \begin{cases} D_{ij} u_\epsilon^+ / |D^2 u_\epsilon^+| & \text{if} \quad D^2 u_\epsilon^+ \neq 0, \\ 0 & \text{if} \quad D^2 u_\epsilon^+ = 0, \end{cases}$$

we write (3.16) in the form

$$\begin{aligned} Lw_\epsilon &\geqslant -C\epsilon^{\delta\chi} g \sigma^{ij} D_{ij} u_\epsilon^+ - o(\epsilon) \\ &= -C\epsilon^{\delta\chi} g \sigma^{ij} (D_{ij} w + D_{ij} v_\epsilon^-) - o(\epsilon) \\ &= -C\epsilon^{\delta\chi} g \sigma^{ij} \left(D_{ij} w - \frac{1}{\epsilon} D_{i\xi_j} w \right) - o(\epsilon). \end{aligned} \tag{3.17}$$

To proceed further we need to write (3.17) as an *elliptic* differential inequality in the variables (x, ξ). First let us suppose that

$$C\epsilon^{\delta\chi} g \leqslant \frac{\lambda}{3} \tag{3.18}$$

to guarantee that inequality (3.17) is at least elliptic in the x variables (with minimum eigenvalue $\geqslant 2\lambda/3$). Setting, analogously to (3.9),

$$a_0^{ij} \equiv \int_0^1 F_{r_{ij}}^-(x, v_\epsilon^-, Dv_\epsilon^-, tD^2 v_\epsilon^-)\, dt, \tag{3.19}$$

we next have

$$a_0^{ij} D_{\xi_i \xi_j} w \geqslant -\epsilon^2 a_0^{ij} D^{ij} v_\epsilon^- \geqslant -C_1 \epsilon^2 \tag{3.20}$$

for a further constant C_1. Consequently

$$\begin{aligned} \tilde{L} w_\epsilon \equiv L w_\epsilon + C\epsilon^{\delta\chi} g \sigma^{ij} (D_{ij} w_\epsilon - \epsilon^{-1} D_{i\xi_j} w_\epsilon) + 3n(Cg\epsilon^{\delta\chi - 1}/\lambda)^2 a_0^{ij} D_{\xi_i \xi_j} w_\epsilon \\ \geqslant -C_1 g^2 \epsilon^{2\delta\chi} - o(\epsilon) \end{aligned} \tag{3.21}$$

for constants C, C_1 independent of ϵ, and the operator $\tilde{L}$ is elliptic if (3.18) is satisfied.

We next need to modify w_ϵ so that it cannot take its maximum for $|\xi| = 1$. Using the Lipschitz continuity of v and (3.1), we set

$$\tilde{w}_\epsilon(x, \xi) = w_\epsilon(x, \xi) - 2K\epsilon|\xi|^2, \tag{3.22}$$

where $K = |Du|_0$, to guarantee

$$\sup_{|\xi| = 1} \tilde{w}_\epsilon(x, \xi) \geqslant \tilde{w}_\epsilon(x, 0) - K\epsilon. \tag{3.23}$$

From (3.19) we then obtain

$$\begin{aligned} \tilde{L}\tilde{w}_\epsilon &\geqslant -C_1 g^2 \epsilon^{2\delta\chi - 1} - o(\epsilon) \\ &\equiv -\tilde{g}\ (\leqslant 0), \end{aligned} \tag{3.24}$$

for a further constant C_1, independent of ϵ. The inequality (3.23) prevents us from directly applying Lemma 2.1, as in the case $\delta = 1$, so we first need to convert (3.24) to a homogeneous inequality. To achieve this we fix a subdomain $\Omega' \subset\subset \Omega_{2\epsilon}$ and a ball $\tilde{B} \subset \Omega$ and extend $\tilde{g}$ to vanish in $\tilde{B} - \Omega'$. We then solve the following *generalized* Dirichlet problem for the Monge-Ampère equation,

$$\begin{aligned} \det D^2 G = h(x, DG) \quad &\text{in} \quad \tilde{B}, \\ G = 0 \quad &\text{on} \quad \partial\tilde{B}, \end{aligned} \tag{3.25}$$

where

$$h(x,p) = \left(\frac{2n}{\lambda}\right)^n (|p|^{n/(n-1)} + g_0^{n/(n-1)})^{n-1}(\mu^n + g_0^{-n}|\tilde{g}|^n),$$

and

$$g_0 = \frac{2}{\lambda}\|\tilde{g}\|_n.$$

If $\mu = 0$, we can simply take

$$h(x,p) = \left(\frac{2n\tilde{g}}{\lambda}\right)^n.$$

From the theory of Aleksandrov [1] and Bakel'man [3], there exists a unique convex generalized solution G satisfying (3.25) almost everywhere in Ω. Furthermore, we can estimate, (see also [6, Section 9.1]),

$$\begin{aligned} \inf G &\geqslant -C\|\tilde{g}\|_{L^n(\Omega')} \\ &\geqslant -C\epsilon^{2\delta\chi-1} - o(\epsilon), \end{aligned} \tag{3.26}$$

by virtue of (3.10), (3.15) and (3.24). Moreover, we now have

$$\tilde{L}(\tilde{w}_\epsilon + G) \geqslant 0, \tag{3.27}$$

almost everywhere on the subset of $\Omega' \times B$ where $w_\epsilon > 0$ and (3.18) holds. Since G is convex the inequality (3.11) continues to hold on the upper contact set of the function $\tilde{w}_\epsilon + G$, whence condition (3.18) is fulfilled there for sufficiently small ϵ. Applying the Aleksandrov maximum principle, Lemma 2.1, I, to the semi-convex function $\tilde{w}_\epsilon + G$, we find

$$\begin{aligned} \sup_{x\in\Omega'}(w_\epsilon + G) &\leqslant \sup_{\partial(\Omega'\times B)}(\tilde{w}_\epsilon + G)^+ + o(\epsilon) \\ &\leqslant \sup_{\partial\Omega'\times B} w_\epsilon^+ + o(\epsilon), \end{aligned} \tag{3.28}$$

by (3.23), and hence we conclude, by (3.26),

$$\begin{aligned} \sup_{\Omega'} w_\epsilon &\leqslant C\epsilon^{2\delta\chi-1} + o(\epsilon) + \sup_{\partial\Omega'} w_\epsilon^+ \\ &\to \sup_{\partial\Omega'}(u-v)^+ \quad \text{as} \quad \epsilon \to 0, \\ &\to \sup_{\partial\Omega}(u-v)^+ \quad \text{as} \quad \Omega' \to \Omega, \end{aligned} \tag{3.29}$$

provided $2\delta\chi - 1 > 0$, that is

$$\delta > \frac{1}{2 + \varkappa/n}. \tag{3.30}$$

We therefore have

$$\sup_{\Omega}(u - v) \leqslant \sup_{\Omega}(u - v)^{+} \tag{3.31}$$

and the proof of Theorem 1.1 is complete. □

Remark 3.1. The continuity condition $F4$ can be made more precise in terms of the value of $\varkappa$ in Lemma 2.1, IV. Namely, if we formulate $F4$ as

$$\tilde{F}4: \qquad |F(x, z, p, r) - F(y, z, p, r)| \leqslant \omega(|x - y|)(|x - y|^{\delta}|r| + 1),$$

then $\delta \geqslant 1/2(1 + \varkappa/n)$ suffices for the validity of Theorem 1.1. If only the first inequality in $F1$ holds, that is F is strictly elliptic rather than uniformly elliptic, then we must have $\varkappa = 0$, as in [10], and the proof of Theorem 1.1 simplifies greatly.

Remark 3.2. Following a standard argument, ([6, Section 3.2]), we can infer from the *weak* comparison principle, Theorem 1.1, a *strong* comparison principle and a *boundary* point inequality. Namely,

(i) (*Strong comparison principle.*) If u and $v \in C^{0,1}(\Omega)$ are respectively subsolution and supersolution of equation (1.1) in Ω, then the difference $w = u - v$ cannot take a positive maximum in Ω unless it is constant.

(ii) (*Boundary point lemma.*) If the above function w assumes a strict positive maximum, continuously at a boundary point x_0, where $\partial\Omega$ satisfies an interior sphere condition (with sphere S) then

$$D_{\beta}^{+} w(x_0) = \limsup_{t \to 0} \frac{w(x_0 + t\beta) - w(x_0)}{t} < 0 \tag{3.32}$$

for any $\beta \cdot \nu > 0$ where ν is the inner normal to S at x_0.

Remark 3.3. The previous remark yields comparison principles for oblique boundary conditions. Let G be a continuous function on $\Gamma_0 = \partial\Omega \times \mathbb{R} \times \mathbb{R}^n$ and suppose that an inner normal ν exists on $\partial\Omega$ in a reasonable sense, for example $\partial\Omega \in C^1$ (or more generally $\partial\Omega$ satisfies an interior cone condition). Then G is *degenerate oblique* (*oblique*) if

$$G(x, z, p + t\nu) - G(x, z, p) \geqslant 0 \qquad (> 0) \tag{3.33}$$

for all $x, z, p \in \Gamma_0$, $t > 0$. For $u \in C^0(\bar{\Omega})$, the boundary conditions $G[u] \geqslant 0$ ($\leqslant 0$) are then satisfied in the viscosity sense if for any $\varphi \in C^1(\bar{\Omega})$ and local maximum (minimum) $x_0 \in \partial\Omega$ of $u - \varphi$, we have

$$G(x_0, u(x_0), D\varphi(x_0)) \geqslant 0 \qquad (\leqslant 0). \tag{3.34}$$

We then have

Corollary 3.1. *Let $u, v \in C^{0,1}(\bar{\Omega})$ be respectively subsolution and supersolution of equation* (1.1) *in Ω satisfying $G[u] \geqslant 0$, $G[v] \leqslant 0$ on $\partial\Omega$, where $\partial\Omega$ satisfies an interior sphere condition and G is strictly oblique* (*with respect to an interior sphere normal at each boundary point*) *and non-increasing with respect to z. Then either $u \leqslant v$ in Ω or $u - v$ is a positive constant.*

PROOF. If $w = u - v$ takes a positive maximum at $x_0 \in \partial\Omega$, so also does $w^* = u^* - v^*$ where $u^*, v^* \in C^1(\bar{\Omega})$ satisfy $u^* \geqslant u$, $v^* \leqslant v$ in Ω,

$$u^*(x_0) = u(x_0),\ v^*(x_0) = v(x_0).$$

But then

$$\begin{aligned} G(x_0, v(x_0), Dv^*(x_0)) &\leqslant 0 \\ &\leqslant G(x_0, u(x_0), Du^*(x_0)) \\ &\leqslant G(x_0, v(x_0), Dv^*(x_0) + \nu \cdot Dw^*(x_0)) \end{aligned}$$

which contradicts (3.35) since $D_\nu w^*(x_0)$ can be chosen arbitrarily close to $D_\nu^+ w$ for $u, v \in C^{0,1}(\bar{\Omega})$. □

Remark 3.4. Finally we remark that if u or $v \in C^{1,\gamma}(\Omega)$, $\gamma > 0$, then by replacing ϵ by $\epsilon^{1+\gamma}$ in the approximations $u_\epsilon^{\pm}$ and (3.22) with $K = 2[Du]_\gamma$, we obtain the above results for

$$\delta > (1 - \gamma)/(2 + \varkappa/n).$$

If u or $v \in C^1(\Omega)$, the proof of Theorem 1.1 can be effected with $\varkappa = 0$.

4. Comparison Principles for Hölder Solutions

In this section we extend Theorem 1.1 to embrace solutions $u, v \in C^{0,\alpha}(\Omega)$ where possibly $\alpha \leqslant 1$. The reduced smoothness of u, v must be compensated by a strengthening of condition $F4$ and we must take $K_1 = \infty$ in conditions $F1$ to $F4$. We shall in fact modify $F4$ to extend $\bar{F}4$ as follows:

F^*4: $|F(x, z, p, r) - F(y, z, p, r)| \leqslant \omega(|x - y|)(|x - y|^\delta |r| + |x - y|^{\delta^*}|p| + 1)$

where δ, δ^* are non-negative constants to be specified. Throughout this section, unless otherwise stated, we assume $F1$, $F2$, $F3$, F^*4 to hold for all $x, y \in \Omega$, $|z| \leqslant K$, $p, q \in \mathbb{R}^n$, r, $\eta \geqslant 0$, $\eta \in \mathbb{S}^n$, $t, K \in \mathbb{R}^+$. We then have the following extension of Theorem 1.1.

Theorem 4.1. *Let $u, v \in C^0(\bar{\Omega}) \cap C^{0,\alpha}(\Omega)$, $0 \leqslant \alpha \leqslant 1$, be respectively subsolution and supersolution of equation* (1.1) *in Ω, with $u \leqslant v$ on $\partial\Omega$. Then if the structure conditions $F1$, $F2$, $F3$, F^*4 are satisfied for*

$$\delta \geqslant (1 - \alpha/2)/(1 + \varkappa/n), \qquad \delta^* \geqslant (1 - \alpha)/(1 + \varkappa/n), \tag{4.1}$$

we have $u \leqslant v$ in Ω.

PROOF. Although the approximations u_ϵ^+, v_ϵ^- are now not necessarily semi-convex, semi-concave, they will be in the upper contact set of the function w_ϵ in the proof of Theorem 1.1 and consequently the case when $\varkappa = 0$ ($g \equiv 1$) follows by only minor adjustment. As there appears some difficulty in extending the Lin estimate (3.10), we shall employ instead the corresponding approximations of Jensen, Lions and Souganidis [11] which are semi-convex, semi-concave everywhere. Accordingly, let us define for positive ϵ, the functions

$$u_\epsilon^+(x) = \sup_{y \in \Omega} \left\{ u(y) - \omega_0 \frac{|x - y|^2}{\epsilon^2} \right\}, \tag{4.2}$$

$$u_\epsilon^-(x) = \inf_{y \in \Omega} \left\{ u(y) + \omega_0 \frac{|x - y|^2}{\epsilon^2} \right\}$$

where $\omega_0 = \operatorname{osc}_\Omega u$. It is clear that the supremum and infimum will be achieved in (4.2) at points $x^\pm$ satisfying $|x - x^\pm| < \epsilon$ provided $x \in \Omega_\epsilon$ and moreover

$$|Du_\epsilon^\pm| \leqslant \frac{2\omega_0}{\epsilon}, \tag{4.3}$$

$$\pm D^2 u_\epsilon^+ \geqslant -\frac{2\omega_0}{\epsilon^2}$$

in the sense of distributions. Furthermore it is easily seen that if u is a viscosity subsolution (supersolution) of equation (1.1) in Ω, then u_ϵ^+ (u_ϵ^-) will be a viscosity subsolution (supersolution) with the operator F replaced by

$$F^\pm(x, z, p, r) = F^\pm(x^\pm, z, p, r).$$

We can now proceed as in the proof of Theorem 1.1 The dependence on p in condition F^*4 is handled by taking $g = (|D^2 u| + |Du|)^{\varkappa/2n}$ and augmenting $\tilde{L}$ by terms in $D_\xi w_\epsilon$. Instead of (3.22) we take

$$\tilde{w}_\epsilon(x, \xi) = w_\epsilon(x, \xi) - 2\sigma(\epsilon)|\xi|^2 \tag{4.4}$$

where

$$\sigma(\epsilon) = \sup_{y \in \Omega'} \left(\operatorname*{osc}_{B_{2\epsilon}(y)} u \right) \leqslant C\epsilon^\alpha. \tag{4.5}$$

Further details are left to the reader. □

Remark 4.1. The case $\alpha = 0$ was proved by Ishii [7] for subsolutions and supersolutions which need only be respectively upper-semicontinuous and lower-semicontinuous. A similar generalization holds here in that the conditions $u, v \in C^{0,\alpha}(\Omega)$ need only hold in a corresponding one-sided sense. Furthermore only u or v need lie in $C^{0,\alpha}(\Omega)$. The case $\varkappa = 0$ is also proved in [10], but by different methods and, as we pointed out previously, only strict ellipticity (or degenerate ellipticity and strict monotonicity in z) suffices in condition $F1$.

5. Pointwise Estimates

Using the approximations $u_\epsilon^\pm$ we can readily convert pointwise estimates for semi-convex subsolutions and semi-concave supersolutions to viscosity subsolutions and supersolutions. Consequently pointwise estimates such as the Hölder and Harnack estimates of Krylov and Safonov [13] can be extended to continuous viscosity solutions. Let us suppose that F is uniformly elliptic, satisfying $F1$ (with $K_1 = \infty$) and

$F5$: $$|F(x, z, p, 0)| < \mu_0 + \mu_2|p|^2$$

for all $x \in \Omega$, $|z| \leqslant K$, $p \in \mathbb{R}^n$, where μ_0, μ_2 are further constants, depending on K. If u is a subsolution of equation (1.1) then the functions

$$v_\epsilon^+ = \frac{1}{\mu_2}(e^{\mu_2 u_\epsilon^+} - 1) \tag{5.1}$$

satisfy

$$L_0 v_\epsilon^+ \equiv a_0^{ij} D_{ij} v_\epsilon^+ \geqslant -\mu_0 e^{\mu_2 |u|_0}$$

almost everywhere in Ω, for some $[a^{ij}]$ satisfying (2.2) and we can apply, for example, the estimates in Section 2. In particular we obtain the local estimates.

Theorem 5.1. *Let $u \in C^0(\Omega)$, F satisfy $F1$, $F5$, and*

$$B = B_R(y) \subset \Omega,$$
$$B_\sigma = B_{\sigma R}(y), \quad 0 < \sigma < 1.$$

I. (*Local maximum principle.*) *If u is a viscosity subsolution of* (1.1),

$$\sup_{B_\sigma} u \leqslant C\left\{\left(R^{-n}\int_B (u^+)^p\right)^{1/p} + \mu_0 R^2/\lambda\right\}, \tag{5.2}$$

for any $p > 0$, where C depends on n, Λ/λ, $\mu_2|u|_0/\lambda$, σ and p;

II. (*Weak Harnack inequality.*) *If u is a non-negative supersolution of* (1.1), *there exists a constant p, depending on n, Λ/λ such that*

$$\left(R^{-n}\int_{B_\sigma} u^p\right)^{1/p} < C\left\{\inf_{B_\sigma} u + \mu_0 R^2/\lambda\right\} \tag{5.3}$$

where C depends on n, Λ/λ, $\mu_2|u|_0/\lambda$ and σ.

III. (*Hölder estimate.*) *If u is a viscosity solution, then $u \in C^\alpha(\Omega)$ for some $\alpha > 0$, depending on n, Λ/λ and $\mu_2|u|_0/\lambda$ and*

$$\operatorname*{osc}_{B_{\sigma R}} u \leqslant C\sigma^\alpha\left\{\operatorname*{osc}_{B_R} u + \mu_0 R^2/\lambda\right\} \tag{5.4}$$

where C also depends on n, Λ/λ and $\mu_2|u|_0/\lambda$.

IV. (*Harnack inequality.*) *If u is a non-negative solution, then*

$$\sup_{B_\sigma} u \leqslant C\left\{\inf_{B_\sigma} u + \mu_0 R^2/\lambda\right\}, \tag{5.5}$$

where C depends on n, Λ/λ and $\mu_2|u|_0/\lambda$.

Remark 5.1. Boundary gradient estimates for nonlinear elliptic equations may be extended in a similar fashion. In particular if $u \in C^0(\bar{\Omega})$, $\partial\Omega \in C^{1,\gamma}$, $u = g$ on $\partial\Omega$, $g \in C^{1,\gamma}(\bar{\Omega})$ for some $\gamma > 0$ and u is a viscosity solution of (1.1) in Ω, then $u \in C^{0,1}(\bar{\Omega})$ and $v = (u - g)/\operatorname{dist}(x, \partial\Omega) \in C^\alpha(\bar{\Omega})$ for some $\alpha > 0$ depending on n, Λ/λ, $\mu_2|u|_0/\lambda$ and γ, (see [6], [12]). Further regularity properties of viscosity solutions are treated in [18], [19], [20].

Remark 5.2. After completion of this manuscript, we received work by Caffarelli [21] on regularity, which overlaps [18], [19] and also proves III and IV above.

References

[1] Aleksandrov, A. D. Dirichlet's problem for the equation det $\|Z_{ij}\| = \phi$, *Vestnik Leningrad Univ.* **13**(1958), 5-24.

[2] Aleksandrov, A. D. Uniqueness conditions and estimates for solutions of the Dirichlet problem, *Vestnik Leingrad Univ.* **18**(1963), 5-29; *Amer. Math. Soc. Transl.* (2) **68**(1968), 89-119.

[3] Bakel'man, I. Ya. The Dirichlet problem for equations of Monge-Ampère type and their n dimensional analogues, *Dokl. Akad. Nauk SSSR* **126**(1959), 923-926.

[4] Crandall, M. G. and Lions, P.L. Viscosity solutions of Hamilton-Jacobi equations, *Trans. Amer. Math. Soc.* **277**(1983), 1-42.

[5] Evans, L. C. Some estimates for non-divergence structure, second order elliptic equations. *Trans. Amer. Math. Soc.* **287**(1985), 701-712.

[6] Gilbarg, D. and Trudinger, N. S. Elliptic partial differential equations of second order, 2nd edition, Springer-Verlag, 1983.

[7] Ishii, H. On uniqueness and existence of viscosity solutions of fully nonlinear second order elliptic PDE's, *Comm. Pure Appl. Math.* **42**(1989), 15-45.

[8] Ishii, H. and Lions, P.L. Viscosity solutions of fully nonlinear second order elliptic partial differential equations, to appear.

[9] Jensen, R. The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations, *Arch. Rational Mech. Anal.*, **101**(1988), 1-27.

[10] Jensen, R. Uniqueness criteria for viscosity solutions of fully nonlinear elliptic partial differential equations. *Indiana Math. J.*, to appear.

[11] Jensen, R., Lions, P.L. and Souganidis, P. E. A uniqueness result for viscosity solutions of fully nonlinear second order partial differential equations, *Proc. Amer. Math. Soc*, **102**(1988), 975-978.

[12] Krylov, N. V. Nonlinear elliptic and parabolic equations of the second order, Reidel, Dordrecht (1987).

[13] Krylov, N. V. and Safonov, M. V. Certain properties of parabolic equations with measurable coefficients, *Izv. Akad. Nauk SSSR Ser. Mat.* **40**(1980), 161-175; *Math. USSR. Izv.* **16**(1981), 151-164.

[14] Lieberman, G. M. Global regularity of solutions of nonlinear second order elliptic and parabolic differential equations. *Math. Zeit.* **193**(1986), 331-346.

[15] Lin, F. H. Second derivative L^p estimates for elliptic equations of nondivergent type. *Proc. Amer. Math. Soc.* **96**(1986), 447-451.

[16] Lions, P.L. Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations II, *Comm. P.D.E.* **8**(1983), 1229-1276.

[17] Trudinger, N. S. Local estimates for subsolutions and supersolutions of general second order eliptic quasilinear equations, *Invent. Math.* **61**(1980), 67-79.

[18] Trudinger, N. S. Hölder gradient estimates for fully nonlinear elliptic equations, *Proc. Roy. Soc. Edinburgh*, **108**(1988), 57-65.

[19] Trudinger, N. S. On regularity and existence of viscosity solutions of nonlinear second order elliptic equations, Partial differential equations and the calculus of variations. Birkhouser (1989), 939-957.

[20] Trudinger, N. S. On the twice differentiability of viscosity solutions of nonlinear elliptic equations, *Bull. Aust. Math. Soc.* **39**(1989), 443-447.

[21] Caffarelli, L. Elliptic second order equations, Lecture Notes (1988).

Neil S. Trudinger
Centre for Mathematical Analysis
Australian National University
GPO Box 4 Canberra ACT 2601
AUSTRALIA

REVISTA MATEMÁTICA IBEROAMERICANA
VOL. 4, N.^{OS} 3 Y 4, 1988

Calderón's Problem for Lipschitz Classes and the Dimension of Quasicircles

Kari Astala

1. Introduction

In last years the mapping properties of the Cauchy integral

$$C_\Gamma f(z) = \frac{1}{2\pi i} \int_\Gamma \frac{f(\xi)}{\xi - z} d\xi$$

have been widely studied. The most important question in this area was Calderón's problem, to determine those rectifiable Jordan curves Γ for which C_Γ defines a bounded operator on $L^2(\Gamma)$. The question was solved by Guy David [Da] who proved that C_Γ is bounded on $L^2(\Gamma)$ (or on $L^p(\Gamma)$, $1 < p < \infty$) if and only if Γ is *regular, i.e.*

$$\mathcal{H}^1(\Gamma \cap B(z_0, R)) \leqslant CR \tag{1}$$

for every $z_0 \in \mathbb{C}$, $R > 0$ and for some constant C.

Once the L^p-cases are settled it is natural to ask when C_Γ is bounded on the other classical function spaces. In particular, it has been shown by Salaev [Sa], *cf.* also [Dy], that if Γ is regular, then C_Γ is a bounded operator on the Lipschitz classes

$$\Lambda^\alpha(\Gamma) = \left\{ f : \|f\|_{\Lambda^\alpha(\Gamma)} = \sup_{x,y \in \Gamma} \frac{|f(x) - f(y)|}{|x - y|^\alpha} < \infty \right\},$$

$0 < \alpha < 1$. Recently Zinsmeister [Z] made the interesting discovery that after a suitable reinterpretation, see Section 3, Calderón's problem makes sense for Lipschitz classes $\Lambda^\alpha(\Gamma)$ even on non-rectifiable curves Γ. His result was a follows.

Theorem 1.1. (Zinsmeister.) *If Γ is a bounded K-quasicircle, there is a constant $a \in [1, 40]$ such that the Cauchy operator*

$$C_\Gamma : \Lambda^\alpha(\Gamma) \to \Lambda^\alpha(\Gamma)$$

is bounded whenever $a(K) < \alpha < b(K)$ and $K^{2a} \leqslant (1 + \sqrt{5})/2$;

$$a(K) = \frac{K^{2a} - 1}{K^{2a} + 1},$$

$$b(K) = (2K^{4a} - 1)^{-1}.$$

Here a curve Γ is called a K-quasicircle if $\Gamma = \varphi(\{|z| = 1\})$ for some K-quasiconformal mapping φ of $\bar{\mathbb{C}}$. Similarly a domain D is called a K-quasidisk if ∂D is K-quasicircle. For the many different characterizations of quasicircles and —disks see, for instance, [L].

In this paper we shall obtain boundedness theorems for the Cauchy operator on all quasicircles and all $K \geqslant 1$. In fact, it turns out that for every quasicircle Γ there is a number $\alpha(\Gamma) < 1$, depending on the dimension rather than the dilatation of Γ, such that C_Γ is bounded on $\Lambda^\alpha(\Gamma)$ if $\alpha \in (\alpha(\Gamma), 1)$ and unbounded if $\alpha \in (0, \alpha(\Gamma)]$.

The best way to describe the behaviour of C_Γ is in terms of A_p weights of Muckenhoupt [M]. Recall that a function $w \geqslant 0$ is said to belong to the class A_p, $1 < p < \infty$, if

$$\left(\frac{1}{|B|}\int_B w(z)\, dm(z)\right)\left(\frac{1}{|B|}\int_B w(z)^{-1/(p-1)}\, dm(z)\right)^{p-1} \leqslant C$$

holds for a constant $C < \infty$ and for all disks $B \subset \mathbb{C}$; here $|B|$ denotes the Lebesgue measure of B. Further, $w \in A_1$ if

$$\frac{1}{|B|}\int_B w(z)\, dm(z) \leqslant C w(x) \quad \text{a.e.} \quad x \in B$$

and $w \in A_\infty$ if

$$\frac{1}{|B|}\int_B w(z)\, dm(z) \leqslant C \exp\left(\frac{1}{|B|}\int_B \log w(z)\, dm(z)\right)$$

hold for all disks B. Then $A_1 \subset A_p \subset A_\infty$ and A_∞ is the union of all A_p classes.

Theorem 1.2. *Let* Γ *be a bounded quasicircle and* $0 < \alpha < 1$. *Then the following conditions are equivalent*

(*a*) $C_\Gamma: \Lambda^\alpha(\Gamma) \to \Lambda^\alpha(\Gamma)$ *is a bounded operator.*
(*b*) $d(z, \Gamma) \in A_p$, $p = 1 + 1/(1 - \alpha)$.

Here $d(z, \Gamma)$ *denotes the euclidean distance from* Γ. *More precisely, if*

$$\alpha(\Gamma) = \inf \{\alpha: d(z, \Gamma) \in A_{1 + 1/(1-\alpha)}\} \tag{2}$$

then $0 \leqslant \alpha(\Gamma) < 1$ *and* C_Γ *is bounded on* $\Lambda^\alpha(\Gamma)$ *whenever* $\alpha(\Gamma) < \alpha < 1$ *and unbounded whenever* $0 < \alpha \leqslant \alpha(\Gamma)$.

The condition (*b*) can be replaced by

(*b'*) $d(z, \Gamma)^{\alpha - 1} \in A_1$

and also by

(*b''*) $d(z, \Gamma)^{\alpha - 1} \in A_\infty$,

in other words by $d(z, \Gamma)^{\alpha - 1} \in A_p$ for any p. As it is well known, the A_p-condition also characterizes the boundedness of the 2-dimensional Hilbert transform or the Beurling-Ahlfors transform

$$Hf(z) = \text{p.v.} \int_{\mathbb{C}} \frac{f(\xi)}{(\xi - z)^2}\, dm(\xi),$$

see [CF]. Hence we have

Corollary 1.3. *If* Γ *is a bounded quasicircle,* $0 < \alpha < 1$ *and* $1 < p < \infty$, *the following conditions are equivalent*

(*a*) $\|C_\Gamma f\|_{\Lambda^\alpha(\Gamma)} \leqslant M_1 \|f\|_{\Lambda^\alpha(\Gamma)}$, $f \in \Lambda^\alpha(\Gamma)$.

(*b*) $\int_{\mathbb{C}} |Hf(z)|^p d(z, \Gamma)^{\alpha - 1}\, dm(z) \leqslant M_2 \int_{\mathbb{C}} |f(z)|^p d(z, \Gamma)^{\alpha - 1}\, dm(z)$,
$f \in L^p_w(\mathbb{C})$, $w(z) = d(z, \Gamma)^{\alpha - 1}$.

In Theorem 1.2 the assumption that Γ is a quasicircle is not necessary, the proof works for a number of other curves, too. For example we obtain a proof for Salaev's theorem, *cf.* Corollary 3.9.

To see more clearly the geometric meaning of Theorem 1.2(*b*) we must introduce some notation. If E is a bounded subset of the complex plane and $0 < r \leqslant \operatorname{diam}(E)$, set

$$M^\beta(E; r) = \inf \left\{ n r^\beta: E \subset \bigcup_{i=1}^{n} B(x_i, r), n \in \mathbb{N} \right\}.$$

Then $\lim_{r\to 0} \sup M^\beta(E; r) = M^\beta(E)$ is the β—dimensional *Minkowski content* of E, *cf.* [MV]. Instead of the Minkowski content we need to use the following closely related quantity

$$h^\beta(E) = \sup\{M^\beta(E; r): 0 < r \leqslant \operatorname{diam}(E)\}.$$

If $\mathcal{H}^\beta$ denotes the β-dimensional Hausdorff measure, then clearly $\mathcal{H}^\beta(E) \leqslant M^\beta(E) \leqslant h^\beta(E)$.

We shall see in Lemma 2.2 below that a Jordan curve Γ is regular if and only if $h^1(\Gamma \cap B(z_0, R)) \leqslant CR$ for all $z_0 \in \mathbb{C}$ and $R > 0$. Therefore it is reasonable to say that Γ is *δ-regular* if there is a constant C such that

$$h^\delta(\Gamma \cap B(z_0, R)) \leqslant CR^\delta, \qquad z_0 \in \mathbb{C}, \qquad R > 0. \tag{3}$$

Theorem 1.4. *If Γ is a bounded quasicircle and if Γ is δ-regular, then C_Γ is bounded on $\Lambda^\alpha(\Gamma)$ whenever $\delta - 1 < \alpha < 1$. Conversely, if C_Γ is bounded on $\Lambda^\alpha(\Gamma)$, then Γ is $(1 + \alpha)$-regular. Thus*

$$\delta(\Gamma) = \inf\{\delta: \Gamma \text{ is } \delta\text{-regular}\} = 1 + \alpha(\Gamma).$$

To illustrate how these results describe the behavior of the Cauchy integral we mention that for the snowflake or Koch curve Γ, $\alpha(\Gamma) = \log(4/3)/\log 3$ and that Γ is $\delta(\Gamma)$-regular. In fact, $\delta(\Gamma) = 1 + \alpha(\Gamma) = \log 4/\log 3 = \dim_H(\Gamma)$, the Hausdorff dimension of Γ.

In the case of a general quasicircle Γ the Hausdorff dimension, the Minkowski dimension $\beta(\Gamma) = \inf\{\beta: M^\beta(\Gamma) < \infty\}$ and the degree of regularity $\delta(\Gamma)$ may be very different. However, the differences vanish if we look at the whole class of all K-quasicircles, *i.e.* as in Theorem 1.1 look for the estimate of $\alpha(\Gamma)$ in terms of the dilation K,

$$\alpha(K) = \sup\{\alpha(\Gamma): \Gamma \text{ is } K\text{-quasicircle}\}.$$

Theorem 1.5. *For each $K \geqslant 1$ the following quantities are equal*

(*a*) $d(K) = \sup\{\dim_H(\Gamma): \Gamma \text{ is } K\text{-quasicircle}\}$.
(*b*) $\beta(K) = \sup\{\beta(\Gamma): \Gamma \text{ is } K\text{-quasicircle}\}$.
(*c*) $\delta(K) = \sup\{\delta(\Gamma): \Gamma \text{ is } K\text{-quasicircle}\}$.

Moreover,

$$1 + \alpha(K) = d(K) = \beta(K) = \delta(K).$$

The above characterization yields numerical estimates for $\alpha(K)$: In a recent article Becker and Pommerenke [BP] estimated the Minkowski dimension of

quasicircles and proved that $\beta(K) \leqslant 2 - C_1 K^{-3,41}$ and that for K close to 1, $1 + 0{,}09\varkappa^2 \leqslant \beta(K) \leqslant 1 + 37\varkappa^2$, $\varkappa = (K-1)/(K+1)$.

Thus, if Γ is a K-quasicircle, then the Cauchy integral C_Γ is bounded on $\Lambda^\alpha(\Gamma)$ whenever $\alpha(K) < \alpha < 1$, where

$$\alpha(K) = d(K) - 1 \leqslant 1 - C_1 K^{-3,41}$$

and

$$0{,}09(K-1)^2/(K+1)^2 \leqslant \alpha(K) \leqslant 37(K-1)^2/(K+1)^2$$

for K near 1. The bound is best possible; if $\alpha < d(K) - 1$ we can find a K-quasicircle Γ such that C_Γ is not bounded on $\Lambda^\alpha(\Gamma)$.

It is our pleasure to express our gratitude to M. Zinsmeister for pointing out mistakes in the first version of this paper and, especially, for his help in constructing the correct proof for Lemma 3.4, which is now based on a suggestion of him.

2. Preliminaries

Following the terminology of Väisälä [V] we call a set A *porous* if there is a constant $0 < \lambda < 1$ such that every disk $B(z_0, R)$ in $\mathbb{C}$ contains a disk $B(z, \lambda R)$ with $A \cap B(z, \lambda R) = \varnothing$. We show first that for porous curves Γ the conditions (b), (b') and (b'') in Theorem 1.2 are equivalent.

Lemma 2.1. *Let Γ be a Jordan curve and $0 < \alpha < 1$.*

(a) If there is a constant $C < \infty$ such that

$$\int_{B(z_0, R)} d(z, \Gamma)^{\alpha - 1}\, dm(z) \leqslant C R^{1+\alpha} \tag{4}$$

for all $z_0 \in \mathbb{C}$ and $R > 0$, then $d(z, \Gamma)^{\alpha-1} \in A_1$.

(b) If Γ is porous and $d(z, \Gamma)^{\alpha-1} \in A_\infty$, then (4) holds for all $z_0 \in \mathbb{C}$, $R > 0$.

(c) If $p = 1 + 1/(1-\alpha)$, then $d(z, \Gamma) \in A_p$ if and only if $d(z, \Gamma)^{\alpha-1} \in A_{p'}$, $p' = p/(p-1)$.

PROOF. Fix $z_0 \in \mathbb{C}$ and $R > 0$ and denote $B = B(z_0, R)$. In (a) if Γ intersects $B(z_0, 2R)$, then $d(z, \Gamma) \leqslant 3R$ for $z \in B$. Thus

$$\frac{1}{|B|} \int_B d(x, \Gamma)^{\alpha-1}\, dm(x) \leqslant C_1 d(z, \Gamma)^{\alpha-1}, \qquad z \in B.$$

If Γ does not intersect $B(z_0, 2R)$, then $d(z_0, \Gamma)/2 \leqslant d(z, \Gamma) \leqslant 2d(z_0, \Gamma)$ for every $z \in B$ and hence

$$\int_B d(x,\Gamma)^{\alpha-1}\,dm(x) \leqslant 4^{1-\alpha}|B|d(z,\Gamma)^{\alpha-1},\ z\in B.$$

To prove (*b*) choose a disk $B(x,\lambda R)\subset B(z_0,R)$ which does not intersect Γ. As $d(z,\Gamma)^{\alpha-1}$ belongs to A_∞, $d(z,\Gamma)^{\alpha-1}\in A_p$ for some $p<\infty$ and so

$$\frac{1}{|B|}\int_B d(z,\Gamma)^{\alpha-1}\,dm(z) \leqslant C\left(\frac{1}{|B|}\int_B d(z,\Gamma)^{(1-\alpha)/(p-1)}\,dm(z)\right)^{1-p}$$
$$\leqslant C|B|^{p-1}(\lambda R/2)^{\alpha-1}|B(x,\lambda R/2)|^{1-p}$$

which gives (4). Finally (*c*) follows from the fact that $w\in A_p$ if and only if $w^{-1/(p-1)}\in A_{p'}$. □

Every quasicircle is porous [V]. There are, of course, many other examples. For instance, it is not difficult to see that regular Jordan curves, or even δ-regular curves with $\delta<2$, are all porous.

Lemma 2.2. *A Jordan curve Γ is regular if and only if it is 1-regular in the sense of* (3).

PROOF. Since $\mathcal{H}^1(E)\leqslant h^1(E)$, 1-regularity implies the usual regularity. Conversely, if Γ is regular and if Γ intersects $B(z_0,R)$, let $\rho\leqslant \operatorname{diam}(\Gamma\cap B(z_0,R))$. Then there exist points $x_j\in\Gamma\cap B(z_0,R)$, $1\leqslant j\leqslant m$, such that $\Gamma\cap B(z_0,R)\subset \bigcup_1^m B(z_j,\rho)$ and each point of $\Gamma\cap B(z_0,R)$ is contained in at most M of the balls $B(x_j,\rho)$ (see, for example, [St]). Here M is an absolute constant. As $x_j\in\Gamma$,

$$\begin{aligned} m\rho &\leqslant m\mathcal{H}^1(\Gamma\cap B(x_j,\rho))\\ &\leqslant M\mathcal{H}^1(\Gamma\cap\textstyle\bigcup_1^m B(x_j,\rho))\\ &\leqslant M\mathcal{H}^1(\Gamma\cap B(z_0,3R))\leqslant CR.\end{aligned}$$

Consequently, $h^1(\Gamma\cap B(z_0,R))\leqslant CR$ and Γ is 1-regular. □

Lemma 2.3. *Let Γ be a porous Jordan curve and $0<\alpha<1$. If $d(z,\Gamma)^{\alpha-1}\in A_1$, then Γ is $(1+\alpha)$-regular, and if Γ is δ-regular $d(z,\Gamma)^{\alpha-1}\in A_1$ whenever $1+\alpha>\delta$. In particular, $\delta(\Gamma)=1+\alpha(\Gamma)$.*

PROOF. Assume first that Γ is δ-regular and denote $B(t)=B(0,t)$. If $t\leqslant R$, it follows from basic covering theorems, *cf.* [MV, Lemma 3.1], that

$$\begin{aligned} t^{\delta-2}|\{z\in B(z_0,R): d(z,\Gamma)<t\}| &\leqslant |\Gamma\cap B(z_0,2R)+B(t)|/t^{2-\delta}\\ &\leqslant C_1 h^\delta(\Gamma\cap B(z_0,2R))\\ &\leqslant C_2 R^\delta.\end{aligned}$$

Integrating this we have for $\alpha \in (\delta - 1, 1)$

$$\begin{aligned}\int_{B(z_0,R)} d(z,\Gamma)^{\alpha-1}\,dm(z) &= (1-\alpha)\int_0^\infty |\{z \in B(z_0,R): d(z,\Gamma) < t\}|\,t^{\alpha-2}\,dt \\ &\leqslant C_2 R^\delta \int_0^R t^{\alpha-\delta}\,dt + (1-\alpha)\pi R^2 \int_R^\infty t^{\alpha-2}\,dt \\ &\leqslant C_3 R^{1+\alpha},\end{aligned}$$

$C_3 = \pi + C_2/(1 + \alpha - \delta)$. According to Lemma 2.1 $d(z,\Gamma)^{\alpha-1} \in A_1$.

On the other hand, in case $d(z,\Gamma)^{\alpha-1} \in A_1$, we may apply as above [MV, Lemma 3.1] and Lemma 2.1 to obtain

$$\begin{aligned}h^{1+\alpha}(\Gamma \cap B(z_0,R)) &\leqslant C \sup_{0<\rho<2R} |\Gamma \cap B(z_0,R) + B(\rho)|/\rho^{1-\alpha} \\ &\leqslant C \sup_{0<\rho<2R} \rho^{\alpha-1} |\{x \in B(z_0,3R): d(x,\Gamma)^{\alpha-1} > \rho^{\alpha-1}\}| \\ &\leqslant C \int_{B(z_0,3R)} d(z,\Gamma)^{\alpha-1}\,dm(z) \\ &\leqslant C_4 R^{1+\alpha}. \quad \square\end{aligned}$$

3. The Cauchy Integral

Let Γ be first a rectifiable Jordan curve and let D be the bounded component of $\mathbb{C}\backslash\Gamma$. If F is a C^∞-function with compact support and if $f = F|_\Gamma$, it then follows from Stokes' theorem that

$$C_\Gamma f(z) = \frac{1}{2\pi i}\int_\Gamma \frac{f(\xi)}{\xi - z}\,d\xi = -\frac{1}{\pi}\int_{\mathbb{C}\backslash D} \frac{\bar\partial F(\xi)}{\xi - z}\,dm(\xi), \qquad z \in D.$$

Here the latter expression is well defined even if Γ is not rectifiable. Hence we can take it as the definition of $C_\Gamma f$ in case of a general or non-rectifiable Γ:

$$(5)\quad C_\Gamma f(z) = -\frac{1}{\pi}\int_{\mathbb{C}\backslash D} \frac{\bar\partial F(\xi)}{\xi - z}\,dm(\xi), \qquad z \in D, f = F|_\Gamma \text{ and } F \in C_0^\infty(\mathbb{C}).$$

Note that by applying the generalized Cauchy integral formula we see immediately that $C_\Gamma f$ as defined in (5), does, indeed, depend only on f and not on the specific extension F. Furthermore, it is easily seen that $C_\Gamma f$ extends continuously to $\bar D$ (in fact, formula (5) defines a function continuous at each $z \in \mathbb{C}$ when $F \in C_0^\infty(\mathbb{C})$).

Definition 3.1. *Let Γ be a Jordan curve bounding the domain $D \subset \mathbb{C}$ and let $0 < \alpha < 1$. We say that the Cauchy operator C_Γ is bounded on $\Lambda^\alpha(\Gamma)$, if*

(6) $\|C_\Gamma f\|_{\Lambda^\alpha(\Gamma)} \leqslant M\|f\|_{\Lambda^\alpha(\Gamma)}$ *for every* $f = F|_\Gamma$, $F \in C_0^\infty(\mathbb{C})$,

for some constant M independent of f.

Remark 3.2. In Definition 3.1 the continuity requirement is minimal or the weakest possible and so the definition is the most general one. If C_Γ is bounded in the above sense, then C_Γ extends a priori only to $\Lambda_0^\alpha(\Gamma)$, the closure of C^∞ in $\Lambda^\alpha(\Gamma)$,

$$\Lambda_0^\alpha(\Gamma) = \{f \in \Lambda^\alpha(\Gamma): |f(x) - f(y)|/|x - y|^\alpha = o(|x - y|)\}.$$

However, the next two lemmas show that if Γ is a quasicircle and C_Γ is bounded on $\Lambda^\alpha(\Gamma)$ in the sense of Definition 3.1, then necessarily *every* $f \in \Lambda^\alpha(\Gamma)$ has an extension F to $\mathbb{C}$ with $\bar\partial F \in L^1(\mathbb{C}\backslash D) \cap C^\infty(\mathbb{C}\backslash D)$. In addition, then $C_\Gamma f(z)$ is well defined via formula (5) for each $f \in \Lambda^\alpha(\Gamma)$ and $z \in D$, $C_\Gamma f(z)$ is analytic in D and it has a continuous extension to $\partial D = \Gamma$ such that (6) holds.

Lemma 3.3. *Let D be a bounded Jordan domain and let* $0 < \alpha < 1$. *Suppose further that* $v \in C_0^\infty(\mathbb{C})$, $\operatorname{supp} v \subset B(z_0, R)$, $z_0 \in \Gamma = \partial D$, $\bar\partial v(z) \geqslant 0$ *for* $z \notin D$ *and that* $\|v\|_{\Lambda^\alpha(\Gamma)} = 1$. *Then, if the Cauchy operator* C_Γ *is bounded on* $\Lambda^\alpha(\Gamma)$,

$$\int_{\mathbb{C}\backslash D} \bar\partial v(z)\, dm(z) \leqslant C_1 R^{1+\alpha}.$$

PROOF. Choosing points $w, w' \in \Gamma$ such that $|w - z_0| = 5R$ and $|w' - z_0| = 10R$ we can estimate

$$\int_{\mathbb{C}\backslash D} \bar\partial v(z)\, dm(z) \leqslant 20R \left| \int_{\mathbb{C}\backslash D} \frac{\bar\partial v(z)}{z - w}\, dm(z) - \int_{\mathbb{C}\backslash D} \frac{\bar\partial v(z)}{z - w'}\, dm(z) \right|.$$

Here the latter expression is equal to $20R|C_\Gamma v(w) - C_\Gamma v(w')|$ and since the Cauchy operator is bounded on $\Lambda^\alpha(\Gamma)$,

$$|C_\Gamma v(w) - C_\Gamma v(w')| \leqslant C_0 |w - w'|^\alpha \leqslant 15^\alpha C_0 R^\alpha.$$

Consequently

$$\int_{\mathbb{C}\backslash D} \bar\partial v(z)\, dm(z) \leqslant C_1 R^{1+\alpha},$$

$C_1 = 20 \cdot 15^\alpha C_0$. □

Lemma 3.4. *Let D be a bounded K-quasidisk and let* $0 < \alpha < 1$. *If the Cauchy operator is bounded on* $\Lambda^\alpha(\Gamma)$, $\Gamma = \partial D$, *then*

(7) $$\int_B d(z, \Gamma)^{\alpha - 1}\, dm(z) \leqslant C R^{1+\alpha}$$

for each disk B of radius R.

PROOF. Let $B = B(z_0, R)$ and $2B = B(z_0, 2R)$. If B does not intersect Γ, (7) follows trivially since then $d(z, \Gamma)^{\alpha-1} \leqslant d(z, \partial B)^{\alpha-1}$ for each $z \in B$. Thus we may assume that $z_0 \in \Gamma$. Moreover, by a similar reasoning it suffices to study only the case $R \leqslant \operatorname{diam}(\Gamma)$.

To prove (7) we shall find a $v \in C_0^\infty(\mathbb{C})$, satisfying the assumptions of Lemma 3.3 with $\operatorname{supp} v \subset B(z_0, 4R)$, such that

$$\int_B d(z, \Gamma)^{\alpha-1}\, dm(z) \leqslant C_2(K) \int_{\mathbb{C}\setminus D} \bar\partial v(z)\, dm(z). \tag{8}$$

Indeed, it is enough to find such a v in the special case $R = 1$ since otherwise we may change variables and set $v_R(z) = R^\alpha v(z_0 + (z - z_0)/R)$ for $R \neq 1$.

Assuming that $R = 1$ choose for each $n \in \mathbb{N}$ a maximal set of points $x_i = x_i^n \in \Gamma \cap 2B$, $1 \leqslant i \leqslant k_n$, such that

$$|x_i^n - x_j^n| \geqslant 2^{2-n}, \qquad i \neq j.$$

From the basic distortion properties of quasiconformal mappings we see that Γ is porous in the following slightly stronger sense: There exists a constant $\lambda = \lambda(K)$ such that $B(x_j^n, 2^{-n-1})$ contains a point $w_j^n \in D$ with $d(w_j^n, \Gamma) \geqslant \lambda 2^{-n}$.

We can now construct the function v. Define first $g(x) = x - 1$ if $\lambda \leqslant x \leqslant 1$, $g(x) = (\lambda - 1)(x/\lambda)^2$ if $0 \leqslant x \leqslant \lambda$ and $g(x) = 0$ if $x \in \mathbb{R}\setminus[0, 1]$. Next set $\varphi(z) = g(|z|)/z$, $z \in \mathbb{C}$. Clearly $\operatorname{supp}\varphi \subset B(1)$, $|\varphi(z) - \varphi(w)| \leqslant \lambda^{-2}|z - w|$ and

$$\bar\partial\varphi(z) = 1/2|z| \quad \text{if} \quad \lambda < |z| < 1. \tag{9}$$

After these preparations let

$$v(z) = \sum_{n=0}^{m} u_n(z), \qquad u_n(z) = \sum_{j=1}^{k_n} 2^{-n\alpha}\varphi(2^n(z - w_j^n)),$$

where $m \in \mathbb{N}$ will be chosen later.

It follows easily that the support of v is contained in $B(z_0, 4)$ and, by (9), that $\bar\partial v(z) \geqslant 0$ if $z \notin D$. Since a standard smoothening gives a function $v \in C_0^\infty(\mathbb{C})$ with the same properties, it remains to show that, for all m, $\|v\|_{\Lambda^\alpha(\Gamma)} \leqslant C_3(K) < \infty$ and that (8) holds when m is large enough.

We start with the Lipschitz estimate. Suppose $z, w \in \Gamma$ and $2^{-p} \leqslant |z - w| < 2^{-p+1}$. Since for a fixed n the disks $B(w_j^n, 2^{-n})$ are disjoint, $|u_n(z) - u_n(w)| \leqslant |u_n(z)| + |u_n(w)| \leqslant 2 \cdot 2^{-n\alpha}$ for $n \geqslant p$ and $|u_n(z) - u_n(w)| \leqslant \lambda^{-2}2^{n-n\alpha}|z - w|$ for $n < p$. Hence

$$\begin{aligned} |v(z) - v(w)| &\leqslant \sum_{n=0}^{p-1} \lambda^{-2}2^{n(1-\alpha)}|z - w| + 2\sum_{n=p}^{m} 2^{-n\alpha} \\ &\leqslant C_3(|z - w|2^{p(1-\alpha)} + 2^{-p\alpha}) \\ &\leqslant C_4|z - w|^\alpha. \end{aligned}$$

Here C_4 depends only on α and K (or λ).

Thus we are left with the proof of (8). Since the disks $B(x_i^n, 2^{2-n})$, $1 \leqslant i \leqslant k_n$, cover $\Gamma \cap 2B$,

$$\begin{aligned} k_n 2^{(1+\alpha)(2-n)} &\geqslant M^{1+\alpha}(\Gamma \cap 2B; 2^{2-n}) \\ &\geqslant C_5 |\Gamma \cap 2B + B(2^{-n})| 2^{(1-\alpha)n}. \end{aligned}$$

In the latter inequality we used again [MV, 3.1]. Beacuse $B(x_i^n, 2^{-n-1}) \subset B(w_i^n, 2^{-n})$ we can find a disk $B(y_i^n, \lambda 2^{-n})$ inside $B(w_i^n, 2^{-n}) \cap (\mathbb{C}\backslash D)$ and as $\bar{\partial} u_n(z) \geqslant 0$ in $\mathbb{C}\backslash D$, (9) implies

$$\int_{\mathbb{C}\backslash D} \bar{\partial} u_n(z)\, dm(z) \geqslant \sum_{i=1}^{k_n} \int_{B(y_i^n, \lambda 2^{-n})} \bar{\partial} u_n(z)\, dm(z) \geqslant \lambda^2 k_n 2^{-n\alpha - n}.$$

Since Lemma 3.3 yields

$$\sum_{n=0}^{m} \int_{\mathbb{C}\backslash D} \bar{\partial} u_n(z)\, dm(z) = \int_{\mathbb{C}\backslash D} \bar{\partial} v(z)\, dm(z) \leqslant C_1 \cdot C_4 \cdot 4^{1+\alpha} < \infty$$

for all $m \in \mathbb{N}$, we may estimate

$$\begin{aligned} \int_B d(z, \Gamma)^{\alpha-1}\, dm(z) &= \int_0^\infty |\{z \in B: d(z, \Gamma)^{\alpha-1} > t\}|\, dt \\ &\leqslant \sum_{n=0}^{\infty} 2^{n(1-\alpha)} |\{z \in B: d(z, \Gamma) < 2^{-n}\}| \\ &\leqslant C_6 \sum_{n=0}^{\infty} \int_{\mathbb{C}\backslash D} \bar{\partial} u_n(z)\, dm(z) \\ &\leqslant 2C_6 \int_{\mathbb{C}\backslash D} \bar{\partial} v(z)\, dm(z) \end{aligned}$$

as soon as m is large enough. Here $C_6 = 8 \cdot 2^\alpha \cdot C_5^{-1} \cdot \lambda^{-2}$. The inequality (7) follows now from Lemma 3.3. □

With Lemma 3.4 and the following corollary to the classical Whitney extension theorem, see [St, p. 174], we can fulfil the promise made in Remark 3.2.

Theorem 3.5. *Let Γ be a bounded Jordan curve in the complex plane and $0 < \alpha < 1$. Then every $f \in \Lambda^\alpha(\Gamma)$ has an extension $F \in \Lambda^\alpha(\mathbb{C})$ such that $\|F\|_{\Lambda^\alpha(\mathbb{C})} < M_0 \|f\|_{\Lambda^\alpha(\Gamma)}$ where M_0 is independent of f, $F \in C^\infty(\mathbb{C}\backslash\Gamma)$, F is compactly supported and $|\operatorname{grad} F(z)| \leqslant M_1 \|f\|_{\Lambda^\alpha(\Gamma)} d(z, \Gamma)^{\alpha-1}$.*

Indeed, if the Cauchy operator is bounded on $\Lambda^\alpha(\Gamma)$ according to Definition 3.1 and Γ is a quasicircle, then $d(z, \Gamma)^{\alpha-1}$ is locally integrable in $\mathbb{C}$ by Lemma

3.4. Thus Whitney's theorem gives for each $f \in \Lambda^\alpha(\Gamma)$ an extension F such that $\bar\partial F(\xi)(\xi - z)^{-1} \in L^1(\mathbb{C}\backslash D)$ whenever $z \in D$. In particular, the expression

$$C_\Gamma f(z) = -\frac{1}{\pi}\int_{\mathbb{C}\backslash D} \frac{\bar\partial F(\xi)}{\xi - z}\, dm(\xi), \qquad z \in D, \tag{10}$$

is well defined for every $f \in \Lambda^\alpha(\Gamma)$. And it is easily seen that this expression does not depend on the particular choice of the admissible extension F and, furthermore, that in case Γ is rectifiable or $f \in C^\infty$, (10) reduces to the standard definition of the Cauchy integral.

It remains to show that in our situation $C_\Gamma f$ has also boundary values in $\Lambda^\alpha(\Gamma)$ with

$$\|C_\Gamma f\|_{\Lambda^\alpha(\Gamma)} \leqslant M\|f\|_{\Lambda^\alpha(\Gamma)} \quad \text{for all} \quad f \in \Lambda^\alpha(\Gamma). \tag{11}$$

This leads us to the sufficiency of the condition (*b*) in Theorem 1.2; the inequality (11) will then be a consequence of Lemmas 3.4, 2.1 and Corollary 3.7.

Lemma 3.6. *Let σ be a complex measure with compact support $K \subset \mathbb{C}$. If $0 < \alpha < 1$ and $|\sigma|(B(z_0, R)) \leqslant MR^{1+\alpha}$ for all $z_0 \in \mathbb{C}$, $R > 0$, then the Cauchy transform*

$$\hat\sigma(z) = \int_{\mathbb{C}} \frac{d\sigma(\xi)}{\xi - z} \tag{12}$$

is holomorphic and α-Hölder continuous with $\|\hat\sigma\|_{\Lambda^\alpha(G)} \leqslant C(\alpha)M$ in each component G of $\mathbb{C}\backslash K$.

Lemma 3.7 is due to Dolzhenko [Do] but under a different formulation. However, the same proof gives the above result; see also [G, Theorem III.4.4] and its proof.

Corallary 3.7. *Let Γ be a bounded and porous Jordan curve and denote by D the bounded component of $\mathbb{C}\backslash\Gamma$. If $0 < \alpha < 1$ and $d(z, \Gamma)^{\alpha - 1} \in A_1$ then every $f \in \Lambda^\alpha(\Gamma)$ has an extension F such that*

$$C_\Gamma f(z) = -\frac{1}{\pi}\int_{\mathbb{C}\backslash D} \frac{\bar\partial F(\xi)}{\xi - z}\, dm(\xi), \qquad z \in D,$$

is well defined and holomorphic with

$$\begin{aligned}\|C_\Gamma f\|_{\Lambda^\alpha(\Gamma)} &= \|C_\Gamma f\|_{\Lambda^\alpha(D)} \\ &\leqslant C\|f\|_{\Lambda^\alpha(\Gamma)}.\end{aligned}$$

PROOF. If $f \in \Lambda^\alpha(\Gamma)$ is given let F be its Whitney extension as in Theorem 3.5. According to Lemma 2.1 the measure

$$\sigma(A) = \int_A \chi_{\mathbb{C}\setminus D}(\xi)\, \bar{\partial} F(\xi)\, dm(\xi)$$

satisfies the growth condition of 3.6. Hence $\|C_\Gamma f\|_{\Lambda^\alpha(D)} \leqslant C\|f\|_{\Lambda^\alpha(\Gamma)}$. The equality $\|C_\Gamma f\|_{\Lambda^\alpha(D)} = \|C_\Gamma f\|_{\Lambda^\alpha(\Gamma)}$ follows now from [GHH]. □

Finally we collect the above steps to the

PROOF OF THEOREM 1.2. If $0 < \alpha < 1$, if Γ is a quasicircle and if the Cauchy operator C_Γ is bounded on $\Lambda^\alpha(\Gamma)$, then $d(z, \Gamma) \in A_{1+1/(1-\alpha)}$ by Lemmas 3.4 and 2.1, *i.e.*, (*a*) implies (*b*). Conversely, if $d(z, \Gamma) \in A_{1+1/(1-\alpha)}$, then $d(z, \Gamma)^{\alpha-1} \in A_1$, the Cauchy integral $C_\Gamma f(z)$, formula (10), is well defined not only for $f \in C^\infty$ but for all $f \in \Lambda^\alpha(\Gamma)$ and $z \in D$ and by Corollary 3.7 C_Γ is a bounded operator with $\|C_\Gamma f\|_{\Lambda^\alpha(\Gamma)} = \|C_\Gamma f\|_{\Lambda^\alpha(D)} \leqslant C\|f\|_{\Lambda^\alpha(\Gamma)}$. Thus (*b*) implies (*a*).

It follows form the work of Gehring and Väisälä [GV], either via the original proof or via Theorem 1.5, that every quasicircle is δ-regular for some $\delta < 2$. Hence, according to Lemma 2.3, $\alpha(\Gamma) = \inf\{\alpha: d(z, \Gamma)^{\alpha-1} \in A_1\} < 1$. Moreover, if w is a weight in the A_1-class, then by Jensen's inequality $w^\beta \in A_1$ whenever $0 < \beta < 1$. Thus C_Γ is bounded on $\Lambda^\alpha(\Gamma)$ for each α in the open interval $(\alpha(\Gamma), 1)$. If C_Γ were bounded in $\Lambda^\alpha(\Gamma)$ for some positive $\alpha \leqslant \alpha(\Gamma)$, then $w(z) = d(z, \Gamma)^{\alpha(\Gamma)-1} \in A_1$ and by Muckenhoupt's theorem [M, p. 214] $w^{1+\varepsilon} \in A_1$ for some $\varepsilon > 0$. But that is clearly impossible as $(\alpha(\Gamma) - 1)(1 + \varepsilon) < \alpha(\Gamma) - 1$. The proof of Theorem 1.2 is complete. □

PROOF OF THEOREM 1.4. If Γ is δ-regular and $\delta < \alpha + 1 < 2$, then $d(z, \Gamma)^{\alpha-1} \in A_1$ and $C_\Gamma: \Lambda^\alpha(\Gamma) \to \Lambda^\alpha(\Gamma)$ by Lemma 2.3 and Theorem 1.2. Conversely, if C_Γ is bounded on $\Lambda^\alpha(\Gamma)$, $d(z, \Gamma)^{\alpha-1} \in A_1$ and Γ is $(1 + \alpha)$-regular. □

Remark 3.8. The proofs described here for Theorems 1.2 and 1.4 remain valid, in addition to the quasicircles Γ, also to a number of other Jordan curves. In fact, the only property of quasicircles we used was that they were «biporous»: There is a constant λ such that whenever $x_0 \in \Gamma$ and $R < \operatorname{diam}(\Gamma)$, then both $D \cap B(x_0, R)$ and $(\mathbb{C}\setminus D) \cap B(x_0, R)$ contain a disk of radius λR. Consequently, Theorems 1.2 and 1.4 hold for all biporous Jordan curves.

The above approach gives also a proof for Salaev's theorem in a generalized form.

Corollary 3.9. *If Γ is a δ-regular Jordan curve and $\delta < 2$, then $C_\Gamma: \Lambda^\alpha(\Gamma) \to \Lambda^\alpha(\Gamma)$ for each $\delta - 1 < \alpha < 1$.*

PROOF. Since δ-regular curves, $\delta < 2$, are porous $d(z, \Gamma)^{\alpha - 1} \in A_1$ by Lemma 2.3. The claim follows therefore from Corollary 3.7. □

There are many other ways to see that $d(z, \Gamma)^{\alpha - 1} \in A_1$ for Γ regular and $0 < \alpha < 1$. For example, one can show directly that $d(z, \Gamma) \in A_p$ for $p > 2$ (M. Zinsmeister, private communication) or we can use the Hardy-Littlewood maximal function $M\mu(x)$ of the arclength measure μ on Γ,

$$M\mu(x) = \sup_{x \in B} \frac{1}{|B|} \int_B d\mu = \sup_{x \in B} \frac{l(\Gamma \cap B)}{|B|};$$

here the supremum is taken over all disks B containing x. Indeed, by regularity $C_1 d(z, \Gamma)^{-1} \leqslant M\mu(z) \leqslant C_2 d(z, \Gamma)^{-1}$ and according to a theorem of Coifman and Rochberg [CR] $(M\mu)^\epsilon$ belongs to the class A_1 whenever $0 < \epsilon < 1$.

Similar arguments yield the correct estimates of the boundedness of the Cauchy integral on many other curves, too. For instance, if Γ is the standard Koch curve or the snowflake curve, then by [H] Γ carries a natural measure μ such that $C_1 d(z, \Gamma)^{\beta - 2} \leqslant M\mu(z) \leqslant C_2 d(z, \Gamma)^{\beta - 2}$ where $\beta = \log 4/\log 3$ is the Hausdorff dimension of Γ. Hence C_Γ is bounded on $\Lambda^\alpha(\Gamma)$ if $\log(4/3)/\log 3 < \alpha < 1$. Conversely, $\beta \leqslant \delta(\Gamma) = 1 + \alpha(\Gamma)$ and thus C_Γ is not bounded on $\Lambda^\alpha(\Gamma)$ if $0 < \alpha \leqslant \log(4/3)/\log 3$. We also note that by combining these estimates with the proof of Lemma 2.3 one can show that the snowflake Γ is $(\log 4)/(\log 3)$-regular. More generally, if Γ is any (porous) Jordan curve which supports a positive measure μ such that

$$C_1 R^d \leqslant \mu(B(z_0, R)) \leqslant C_2 R^d$$

whenever $z_0 \in \Gamma$ and $R < \operatorname{diam}(\Gamma)$, then

$$d = \dim_H(\Gamma) = \beta(\Gamma) = \delta(\Gamma) = 1 + \alpha(\Gamma)$$

and Γ is d-regular. In particular, *cf.* [MV, 4.19], this holds for all selfsimilar fractal curves satisfying the open set condition of [H, p. 735].

4. The Hausdorff Dimension

In this last section we prove Theorem 1.5, the relation between $\alpha(K)$ and the upper bound for the Hausdorff dimension $d(K) = \sup\{\dim_H(\Gamma)$: Γ is K-quasicircle$\}$. According to Theorem 1.4 it will be enough to show that $\delta(K) = d(K)$. For this some lemmas are needed.

Lemma 4.1. *For each* $K \geqslant 1$, $d(K) = \lim_{\epsilon \to 0^+} d(K + \epsilon)$.

PROOF. All bounded $(K+\epsilon)$-quasicircles Γ are of the form $\Gamma = \lambda\varphi(\{|z| = 1\}) + \mu$ where $\lambda, \mu \in \mathbb{C}$ and φ is $(K+\epsilon)$-quasiconformal on $\bar{\mathbb{C}}$ with $\varphi(0) = 0$, $\varphi(1) = 1$ and $\varphi(\infty) = \infty$. Moreover φ admits the factorization $\varphi = \varphi_1 \circ \varphi_2$ where φ_1, φ_2 fix $0, 1, \infty$ and have dilatations $K(\varphi_1) = (K+\epsilon)/K$, $K(\varphi_2) = K$, *cf.* [L, p. 29].

According to Mori's classical distortion theorem φ_1 is $1/K(\varphi_1)$-Hölder continuous on compact subsets of $\mathbb{C}$. Hence

$$\dim_H(\Gamma) \leqslant K(\varphi_1)\dim_H(\varphi_2\{|z| = 1\}) \leqslant (1 + \epsilon/K)d(K).$$

Since Γ was arbitrary, $d(K+\epsilon) \leqslant (1+\epsilon/K)d(K)$. □

The next lemma is a standard deformation argument. No proof, however, seems to appear in the literature and hence we sketch the details.

Lemma 4.2. *Let φ be a K-quasiconformal mapping on $\bar{\mathbb{C}}$ fixing $0, 1$ and ∞. Then for each $\epsilon > 0$ there is a number $\rho = \rho(K, \epsilon) \in (0, 1/2)$ and a $(K+\epsilon)$-quasiconformal mapping ϕ on $\mathbb{C}$ such that*

(a) $\phi(z) = \varphi(z)$ *if* $1/2 \leqslant |z|$
(b) $\phi(z) = z$ *if* $|z| \leqslant \rho$.

PROOF. Assume first that φ is conformal in the unit disk $B(1)$. If $\lambda = \varphi'(0)$, then $1/M \leqslant |\lambda| \leqslant M$ and $|\varphi(z) - \lambda z| \leqslant M\rho^2$, $|\varphi'(z) - \lambda| \leqslant M\rho$ for $|z| \leqslant \rho < 1/2$ with a constant M depending only on K. Given a C^∞-function v such that $v(z) = 0$ for $|z| \geqslant 2$ and $v(z) = 1$ for $|z| \leqslant 1$, set

$$g(z) = \varphi(z) + (\lambda z - \varphi(z))v(z/\rho).$$

Then g is quasiconformal on $\mathbb{C}$ and $K(g|_{B(1/2)}) \leqslant 1 + C\rho$ for ρ small. Finally, we replace g by $g(z)(|g(z)|/\lambda\rho)^{\pm\epsilon}$ in an annulus $\rho_1 \leqslant |z| \leqslant \rho$ and obtain a mapping $\tilde{g}$ with the properties: $\tilde{g}(z) = \varphi(z)$ if $|z| > 1/2$, $\tilde{g}(z) = z$ if $|z| < \rho_1$ and $K(\tilde{g}|_{B(1/2)}) \leqslant 1 + \epsilon$, $\rho_1 = \rho_1(K, \epsilon)$.

The general case follows from the above. Indeed, we may factorize $\varphi = k^{-1} \circ h$, where h is conformal in $B(\rho_2)$ and k is conformal outside $\varphi B(\rho_2)$ and deform h and k so that $\varphi(z) = h(z)$ for $|z| > 1/2$ and $h(z) = z$ for $|z| < \rho$. □

Lemma 4.3. *If $d(K) < \delta$, there is a constant*

$$C_0 = C_0(K, \delta)$$

such that

$$h^\delta(\varphi[0, 1]) \leqslant C_0$$

for each K-quasiconformal mapping φ on $\bar{\mathbb{C}}$ fixing $0, 1$ and ∞.

PROOF. If the claim is not true, we can find a sequence $\{\varphi_n\}_1^\infty$ of K-quasiconformal mappings on $\bar{\mathbb{C}}$, each fixing 0, 1 and ∞, such that

$$h^\delta(\varphi_n[0,1]) > n, \qquad n \in \mathbb{N}. \tag{13}$$

Using Lemma 4.2 we shall then construct for every $\epsilon > 0$ a new mapping Φ_ϵ on $\bar{\mathbb{C}}$ with $K(\Phi_\epsilon) \leqslant K + \epsilon$ and $\dim_H(\Phi_\epsilon[0,1]) > \delta$. By the Möbius invariance of quasiconformal mappings

$$d(K) = \sup\{\dim_H(\varphi[0,1]): \varphi \text{ is } K\text{-quasiconformal on } \bar{\mathbb{C}},\ \varphi(\infty) = \infty\}$$

and hence we obtain $d(K+\epsilon) \geqslant \delta > d(K)$ for all $\epsilon > 0$. This, however, contradicts Lemma 4.1. Therefore to prove Lemma 4.3 it is enough to find the mappings Φ_ϵ.

Now, assuming the existence of the sequence (13), choose for each n a radius $r_n < 1$ such that $M^\delta(\varphi_n[0,1]; r_n) \geqslant n$. Then choose a maximal set of points $z_i = z_i^n \in [0,1]$, $1 \leqslant i \leqslant k_n$ such that

$$|\varphi_n(z_i) - \varphi_n(z_j)| \geqslant r_n, \qquad i \neq j. \tag{14}$$

Clearly, the union of the balls $B(\varphi_n(z_i), r_n)$ covers $\varphi_n[0,1]$ and thus

$$k_n(r_n)^\alpha \geqslant M^\delta(\varphi_n[0,1]; r_n) \geqslant n. \tag{15}$$

By (14) the disks $B(\varphi_n(z_i), r_n/2)$ are disjoint. Let $B(z_i, \lambda_i)$ be the largest disk, with center z_i, contained in $\varphi_n^{-1}B(\varphi_n(z_i), r_n/2)$.

Next, we deform φ_n and create «holes» at the disks $B(z_i, \lambda_i)$. We shall then fill the holes by similarity-copies of φ_n and as a result obtain a selfsimilar set E, $\dim_H(E) > \delta$, contained in a $(K+\epsilon)$-quasicircle.

To be more precise, we fix $\epsilon > 0$ and apply Lemma 4.2 to find a number $\rho = \rho(K, \epsilon) \in (0, 1/2)$ and for each $n \in \mathbb{N}$ a $(K+\epsilon)$-quasiconformal mapping ϕ_n such that the following conditions hold.

(16*a*) $\phi_n(z) = \varphi_n(z)$ if $|z| < 2$ and $z \notin \bigcup_i B(z_i, \lambda_i)$

(16*b*) $\phi_n(z) = z$ if $1/\rho < |z|$.

(16*c*) In $B_i \equiv B(z_i, \rho\lambda_i)$ $\phi_n = \tau_i$, a similarity with $\tau_i(z_i) = \varphi_n(z_i)$ and $\tau_i(z_i + \lambda_i) = \varphi_n(z_i + \lambda_i)$.

Note that here τ_i, λ_i and B_i depend also on n. From the distortion properties of quasiconformal mappings we deduce

$$\frac{r_n}{C_1\lambda_i} \leqslant |\tau_i'| \leqslant \frac{r_n}{2\lambda_i} \tag{17}$$

where $C_1 = C_1(K)$.

If B_i is as in (16c) let u_i be the similarity $u_i(z) = z_i \pm \lambda_i \rho^2 z$ with $u_i B(0, 1/\rho) = B_i$ and $u_i[0, 1] \subset [0, 1]$. Then the holes B_i can be filled in by defining new quasiconformal mappings $\phi_n^{(k)}$ as follows: set $\phi_n^{(1)} = \phi_n$,

$$\phi_n^{(k)}(z) = \phi_n \circ u_i \circ \phi_n^{(k-1)} \circ u_i^{-1}(z), \quad \text{if} \quad z \in B_i, \tag{18}$$

and $\phi_n^{(k)}(z) = \phi_n^{(k-1)}(z)$ otherwise. It is easily seen that each $\phi_n^{(k)}$ is $(K + \epsilon)$-quasiconformal on $\bar{\mathbb{C}}$. In fact, ϕ_n is a similarity on B_i and $u_i \circ \phi_n^{(k-1)} \circ u_i^{-1}$ the identity outside B_i. Consequently, as $k \to \infty$ the $\phi_n^{(k)}$ converge uniformly on $\bar{\mathbb{C}}$ to a $(K + \epsilon)$-quasiconformal mapping Φ_n.

Finally, from (16c) and (18) we have

$$\Phi_n \circ u_i(z) = \tau_i \circ u_i \circ \Phi_n(z), \qquad |z| < 1/\rho. \tag{19}$$

The similarities u_i are contractions and by Hutchinson's theorem [H, 3.2] there is a unique compact set E_u such that

$$\Sigma_u(E_u) = E_u, \qquad \Sigma_u(A) = \bigcup_{i=1}^{k_n} u_i(A).$$

Since these similarities map the unit interval into itself, $E_u \subset [0, 1]$. On the other hand, the similarities $\tau_i \circ u_i$ are also contractions, $r_n \rho^2 / C_1 \leqslant |(\tau_i \circ u_i)'| \leqslant r_n \rho^2 / 2$ by (17). Hence we have a unique compact E for which

$$\Sigma_\tau(E) = E, \qquad \Sigma_\tau(A) = \bigcup_{i=1}^{k_n} \tau_i \circ u_i(A)$$

and it is easily seen from (19) that $\Phi_n(E_u) = E$.

Lastly, we have to estimate the Hausdorff dimension of E. Because the disks B_i are disjoint, E_u and hence E satisfy the open set condition of Hutchinson [H, 5.2]. According to [H, 5.3], $\dim_H(E)$ is then the unique number s for which

$$\sum_{i=1}^{k_n} |(\tau_i \circ u_i)'|^s = 1.$$

But $|(\tau_i \circ u_i)'| \geqslant r_n \rho^2 / C_1$ and when n is large, (15) yields $k_n (r_n \rho^2 / C_1)^\delta \geqslant n \rho^{2\delta} / C_1^\delta > 1$. Therefore

$$\dim_H(\Phi_n[0, 1]) \geqslant \dim_H(E) > \delta, \qquad n \geqslant n_0. \quad \square$$

PROOF OF THEOREM 1.5. We must show that $d(K) = \delta(K)$. Since $d(K) < \delta(K)$ trivially, it is enough to prove that for each $\delta > d(K)$ and each bounded K-quasicircle Γ there is a constant $C < \infty$ with

$$h^\delta(B(z_0, R) \cap \Gamma) \leqslant C R^\delta, \qquad z_0 \in \mathbb{C}, \qquad R > 0.$$

By Lemma 4.3 and the Möbius invariance of quasiconformal mappings, if $\delta > d(K)$ then

$$h^{\delta}(\Gamma) \leqslant C(K, \delta) \operatorname{diam}(\Gamma)^{\delta}$$

for all bounded K-quasicircles Γ. If z_0, R are given, take a point $w_0 \in B(z_0, R)$ such that $d(w_0, \Gamma) \geqslant C_0(K)R$ and let $\phi(z) = (z - w_0)^{-1}$. As $|\phi(z) - \phi(z')| \geqslant |z - z'|(2R)^{-2}$ whenever $z, z' \in B(z_0, R) \cap \Gamma$,

$$\begin{aligned} h^{\delta}(B(z_0, R) \cap \Gamma) &\leqslant (2R)^{2\delta} h^{\delta}(\phi\Gamma) \\ &\leqslant (2R)^{2\delta} C \operatorname{diam}(\phi\Gamma)^{\delta}. \end{aligned}$$

However, $\operatorname{diam}(\phi\Gamma)^{\delta} \leqslant (2/C_0R)^{\delta} = C_1 R^{-\delta}$ and hence the claim is proved. The equalities $1 + \alpha(K) = d(K) = \beta(K) = \delta(K)$ follow now from Theorem 1.4. □

References

[BP] Becker, J. and Pommerenke, Ch. On the Hausdorff Dimension of Quasicircles. *Ann. Acad. Sci. Fenn. Ser. A I Math.*, **12**(1988).

[CF] Coifman, R. R. and Fefferman, C. C. Weighted Norm Inequalities for Maximal Functions and Singular Integrals. *Studia Math.*, **51**(1974), 241-250.

[CR] Coifman, R. R. and Rochberg, R. Another Characterization of BMO. *Proc. Amer. Math. Soc.*, **79**(1980), 249-254.

[Da] David, G. Opérateurs intégraux singuliers sur certaines courbes du plan complexe. *Ann. Sc. E.N.S.*, **17**(1984), 157-189.

[Do] Dolzhenko, E. P. The Removability of Singularities of Analytic Functions. *Uspeki Mat. Nauk.*, **18**(1963), (112), 135-142 (Russian).

[Dy] Dynkin, E. M. On the Smoothness of Cauchy Type Integrals. *Sem. Leningrad. Otdel. Mat. Inst. Steklov.*, **92**(1979), 115-133.

[G] Garnett, J. Analytic Capacity and Measure. Springer-Verlag, 1972.

[GHH] Gehring, F. W., Hayman, W. K. and Hinkkanen, A. Analytic Functions Satisfying Hölder Conditions on the Boundary. *J. Approx. Theory*, **35**(1982), 243-249.

[GV] Gehring, F. W. and Väisälä, J. Hausdorff Dimension and Quasiconformal Mappings. *J. London Math. Soc.*, **6**(1973), 504-512.

[H] Hutchinson, J. E. Fractals and Selfsimilarity. *Indiana Univ. Math. J.*, **30** (1981), 713-747.

[L] Lehto, O. Univalent Functions and Teichmüller Spaces. Springer-Verlag, 1987.

[MV] Martio, O. and Vuorinen, M. Whitney Cubes, p-Capacity and Minkowski Content. *Expo. Math.*, **5**(1987), 17-40.

[M] Muckenhoupt, B. Weighted Norm Inequalities for the Hardy Maximal Function. *Trans. Amer. Math. Soc.*, **165**(1972), 207-226.

[Sa] Salaev, V. V. Direct and Inverse Estimates for a Singular Cauchy Integral Along a Closed Curve. *Math. Zametki*, **19**(1976), 365-380 (Russian).

[St] Stein, E. M. Singular Integrals and Differentiability Properties of Functions. Princeton Univ. Press, 1970.
[V] Väisälä, J. Porous Sets and Quasisymmetric Maps. *Trans. Amer. Math. Soc.*, **299**(1987), 525-533.
[Z] Zinsmeister, M. Problèmes de Dirichlet, Neumann, Calderón dans les quasi-disques pour les classes hölderiennes. *Revista Mat. Iberoamericana*, **2**(1986), 319-332.

Kari Astala
Department of Mathematics
University of Helsinki
SF-00100 Helsinki 10
FINLAND