Boletín de la

SOCIEDAD MATEMÁTICA MEXICANA

Tercera Serie

Volumen 12
Número 1
Abril de 2006

Contenido

Continúa/Continued on back cover

SOCIEDAD MATEMATICA MEXICANA

Bol. Soc. Mat. Mexicana (3) Vol. 12, 2006

DINÁMICA DE LOS POLINOMIOS CUADRÁTICOS

GAMALIEL BLÉ GONZÁLEZ Y ROGELIO VALDEZ DELGADO

ABSTRACT. En este artículo presentamos los elementos básicos de la dinámica holomorfa y enunciamos los resultados y las conjeturas más importantes que involucran a las funciones racionales y en particular a los polinomios cuadráticos.

ABSTRACT. We present the basic elements in the theory of holomorphic dynamics and state the main results and conjectures in the field concerning rational maps and particular, about the family of quadratic polynomials.

1. Introducción

El comienzo de la dinámica holomorfa se remonta a los trabajos del siglo XIX de E. Schroeder, G. Koenigs, L. E. Böttcher y a las memorias sobre iteración de aplicaciones racionales de G. Julia y P. Fatou, alrededor de 1920. Después de ellos, el área estuvo en hibernación, pasando algunas décadas sin grandes avances, salvo los trabajos de H. Cremer en 1932, C. L. Siegel en 1942 y de H. Brolin en 1965.

En la década de los 80's el estudio de la dinámica holomorfa regresó de manera explosiva al primer plano de la investigación. Este resurgimiento o renacimiento de la dinámica holomorfa se debió en gran parte a los avances en la graficación por computadora y la introducción de una nueva herramienta teórica, las aplicaciones quasiconformes. Esta nueva herramienta fue utilizada por D. Sullivan en 1982, para demostrar la conjetura de Fatou, acerca de dominios no errantes, resolviendo uno de los problemas principales que Fatou había dejado abierto.

Este acontecimiento marcó una nueva etapa en el estudio de la dinámica de las aplicaciones racionales, destacándose los trabajos de A. Douady y J. H. Hubbard en 1982. La teoría de aplicaciones quasiconformes, ha dado respuesta a muchos de los interrogantes que se han presentado en la iteración de las aplicaciones racionales.

En este trabajo se presenta un panorama general sobre los avances que se han tenido en las últimas décadas en el estudio de la iteración de las aplicaciones racionales.

2000 *Mathematics Subject Classification*: Primary: 37F10; Secondary: 37F15, 37F25, 37F45, 37F50.

Keywords and phrases: Rational maps, holomorphic dynamics, Julia set, Mandelbrot set, renormalization.

El primer autor fue apoyado por el proyecto de CONACYT, 42249 y el segundo por PROMEP, UAEMOR-PTC-166. Los autores agradecen al Instituto de Matemáticas de la UNAM, Unidad Cuernavaca, por su generosa hospitalidad durante la realización de este trabajo.

Dentro de las posibles familias de aplicaciones racionales o polinomiales, hemos decidido detenernos en los detalles que conciernen a la familia cuadrática $f_c(z) = z^2 + c$, ya que ésta presenta un espectro completo de los comportamientos dinámicos.

2. Clasificación de órbitas periódicas

Uno de los problemas principales en el estudio de la iteración de una aplicación racional $R\colon \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$, es mostrar la convergencia de la sucesión $\{z_n\}$, generada a partir de una condición inicial $z_0 \in \widehat{\mathbb{C}}$ y definida recursivamente como $z_n = R(z_{n-1})$. Notemos que si la sucesión $\{z_n\}$ tiene un límite L, entonces L es un punto fijo de R, ya que

$$R(L) = R(\lim_{n\to\infty} z_n) = \lim_{n\to\infty} R(z_n) = \lim_{n\to\infty} z_{n+1} = L.$$

Por lo que los puntos fijos y los puntos periódicos de R juegan un papel muy importante en la dinámica global de R, ya que si en lugar de tomar toda la sucesión $\{z_n\}$, tomamos la subsucesión generada por los múltiplos de k, para un k fijo, es decir, la subsucesión $\{z_{kn}\}$, entonces, cuando ésta converja, por el argumento anterior, lo hará a un punto periódico de periodo l, donde $l|k$.

Definición (2.1). Un punto $z \in \widehat{\mathbb{C}}$ es un punto periódico de periodo k de una función f si $f^k(z) = z$ y $f^j(z) \neq z$ para $j < k$, donde f^k denota la composición de f consigo misma k veces.

Dado $z \in \widehat{\mathbb{C}}$, al conjunto

$$\mathcal{O}_f(z) = \{w \in \mathbb{C} : w = f^k(z) \text{ para alguna } k \in \mathbb{N} \cup \{0\}\}$$

se le llama la *órbita* de z bajo f. En el caso que z es periódico, este conjunto es finito y se le llama órbita periódica.

Definición (2.2). Dadas f y g funciones analíticas, se dice que son topológicamente (analíticamente) conjugadas en el abierto $U \subset \mathbb{C}$, si existe un homeomorfismo (bi-holomorfismo) $\varphi\colon U \to \varphi(U)$ tal que $\varphi \circ f(z) = g \circ \varphi(z)$ para $z \in U$. Esto es, el siguiente diagrama conmuta,

$$\begin{array}{ccc} U & \xrightarrow{f} & U \\ {\scriptstyle\varphi}\downarrow & & \downarrow{\scriptstyle\varphi} \\ \varphi(U) & \xrightarrow[g]{} & \varphi(U). \end{array}$$

Observemos que si f y g son dos funciones topológicamente conjugadas por medio de φ y z es un punto fijo de f, entonces $\varphi(z)$ es un punto fijo de g, ya que $g(\varphi(z)) = \varphi(f(z)) = \varphi(z)$. De hecho, aplicando inducción sobre n se puede mostrar que esto también es cierto para los puntos periódicos de f en U.

PROPOSICIÓN (2.3). *Para cada polinomio cuadrático $P(z) = a_0 + a_1 z + a_2 z^2$, existe $c \in \mathbb{C}$, tal que P y $f_c(z) = z^2 + c$ son analíticamente conjugados. De hecho, la conjugación φ es un bi-holomorfismo del plano complejo, es decir, una función afín de la forma $az + b$, con $a, b \in \mathbb{C}$, $a \neq 0$.*

Demostración. Sean $P(z) = a_0 + a_1 z + a_2 z^2$, $f_c(z) = z^2 + c$ y $\varphi(z) = az + b$, con a_0, a_1, a_2, a, b, $c \in \mathbb{C}$, tales que $aa_2 \neq 0$.
Resolviendo la ecuación $\varphi \circ P(z) = f_c \circ \varphi(z)$ se obtiene: $a = a_2$, $b = \frac{a_1}{2}$ y $c = a_0 a_2 + \frac{a_1}{2} - \frac{a_1}{2}^2$. □

Esta proposición nos muestra que el estudio dinámico de los polinomios cuadráticos lo podemos restringir a los polinomios de la forma f_c, con la ventaja de que estos últimos están parametrizados por el campo $\mathbb{C}$. De aquí, encontrar los puntos fijos de P se reduce a encontrar los puntos fijos de f_c, es decir, las raíces del polinomio $z^2 - z + c$, las cuales son:

$$z_{1,2} = \frac{1 \pm \sqrt{1-4c}}{2}.$$

Estas raíces son dos, salvo en el caso $c = \frac{1}{4}$, donde el único punto fijo de f_c es $\frac{1}{2}$.

Definición (2.4). Sea f una función analítica y z un punto periódico de f de periodo k con multiplicador $\lambda = Df^k(z)$, donde $Df^k(z)$ denota la derivada compleja de f^k en z. Decimos que:

1. z es atractor si $\mid \lambda \mid < 1$; si $\lambda = 0$ diremos que z es super-atractor,
2. z es repulsor si $\mid \lambda \mid > 1$, y
3. z es indiferente si $\mid \lambda \mid = 1$.

Si $\{z_1, z_2, \ldots, z_k\}$ es una órbita periódica de f, entonces aplicando la regla de la cadena, se obtiene que $Df^k(z_j)$ no depende de la j; en consecuencia, la definición anterior dada para puntos fijos, también se aplica a las órbitas periódicas.

Para mostrar la importancia que tiene esta clasificación de los puntos periódicos enunciaremos sin demostración algunos teoremas que pueden ser revisados en [CG], [B], [M1].

(2.5) Atractores y Repulsores.

TEOREMA (2.5.1) (Koenigs-1884). *Sea f una función analítica con un punto fijo en z_0 y cuyo multiplicador es λ. Si $0 < |\lambda| < 1$ ó $|\lambda| > 1$, entonces existen U, V vecindades de z_0 y de 0, respectivamente, y un biholomorfismo $\varphi\colon U \to V$ que conjuga analíticamente f con $g(z) = \lambda z$. Además, esta conjugación es única, módulo multiplicación por un escalar real.*

Este teorema nos dice que cerca de los puntos fijos atractores o repulsores las funciones se comportan como multiplicación por λ. En particular, si z_0 es atractor, las órbitas de los puntos en una vecindad de z_0 convergen a z_0.
Este teorema también es válido para órbitas periódicas atractoras o repulsoras de periodo k, sustituyendo f por f^k.

(2.6) Super-atractores.

TEOREMA (2.6.1) (Böttcher-1904). *Sean f una aplicación analítica y z_0 un punto fijo super-atractor. Si $f(z) = z_0 + a_p(z - z_0)^p + \ldots$, con $a_p \neq 0$, entonces existen U, V vecindades de z_0 y 0, respectivamente, y un biholomorfismo $\varphi\colon U \to V$ que conjuga $f(z)$ y $g(z) = z^p$. Esta conjugación es única módulo multiplicación por una raíz $(p-1)$-ésima de la unidad.*

De igual manera que en el teorema de Koenigs, este teorema es válido para órbitas periódicas super-atractoras.

(2.7) Indiferente. Si una función analítica f en z_0, tiene un punto fijo en z_0 con multiplicador λ de módulo uno, entonces pueden ocurrir dos cosas: que exista una vecindad U de z_0 donde f sea conjugada a la rotación λz o, que no exista tal vecindad; si ocurre el primer caso, decimos que f es *linealizable* en z_0.

Cuando λ es una raíz de la unidad, es fácil mostrar que f no es linealizable [CG]. Sin embargo, cuando $\lambda = e^{2\pi i\theta}$ y θ es irracional, el problema de linealización ha resultado ser complicado y antes de dar los resultados que se han obtenido, será necesario dar algunas definiciones de teoría de números que pueden ser consultadas en [HW].

Notemos que si $t \in [0,1)$, entonces lo podemos desarrollar en fracción continua y obtener una sucesión de números racionales

$$\frac{p_n}{q_n} = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ldots + \cfrac{1}{a_n}}}} := [a_1, \ldots, a_n]$$

que converge a t. A $\dfrac{p_n}{q_n}$ se le llama n-ésima reducida o n-ésima aproximación a t. De hecho, la sucesión $\{a_n\}_{n\geq 1}$ está formada por números enteros no negativos y se obtiene usando el algoritmo de la división. En general, se puede obtener la sucesión $\{a_n\}_{n\geq 0}$ para todo $t \in \mathbb{R}$ y, salvo a_0, todos son enteros no negativos. Además, la sucesión es finita cuando t es racional e infinita cuando t es irracional.

Definición (2.7.1). Sea $t \in \mathbb{R} \setminus \mathbb{Q}$, $t = [a_0, a_1, a_2, \ldots]$. Decimos que t es de tipo *acotado* si la sucesión $\{a_n\}$ es acotada.

Definición (2.7.2). Decimos que $t \in \mathbb{R}$ es diofantino de exponente k si existe $C > 0$, tal que

$$\left|\theta - \frac{p}{q}\right| \geq \frac{C}{q^k} \qquad \text{para todo} \quad \frac{p}{q} \in \mathbb{Q}.$$

PROPOSICIÓN (2.7.3). *Sea $t \in \mathbb{R} \setminus \mathbb{Q}$ tal que $t = [a_0, a_1, a_2, \ldots]$. Entonces, t es de tipo acotado si y sólo si t es diofantino de exponente dos.*

Como todos los números algebraicos son de tipo acotado, este resultado nos dice que los números algebraicos son mal aproximados por racionales.

TEOREMA (2.7.4) (Siegel-1942). *Sea f una función analítica en z_0 tal que $f(z_0) = z_0$ y $\lambda = f'(z_0) = e^{2\pi i\theta}$. Si θ es diofantino, entonces f es linealizable en z_0.*

Al dominio máximo de linealización Δ de f se le llama *disco de Siegel* y a θ se le llama número de rotación de f en Δ.

A partir de este teorema, se tiene la siguiente clasificación de los puntos indiferentes en términos de θ.

Definición (2.7.5). Sea f una función analítica en z_0 tal que z_0 es un punto fijo de f con multiplicador $\lambda = e^{2\pi i\theta}$. Decimos que el punto fijo z_0 es:

1. Parabólico, si θ es racional.
2. Siegel, si θ es irracional y f es linealizable en una vecindad de z_0.
3. Cremer, si θ es irracional y f no es linealizable en z_0.

TEOREMA (2.7.6) (Brjuno-1965). *Sea f una función analítica en z_0 tal que z_0 es un punto fijo de f con multiplicador $\lambda = e^{2\pi i\theta}$. Si $\frac{p_n}{q_n}$ denota la n-ésima aproximación a θ y*

$$\sum_{n=1}^{\infty} \frac{\log q_{n+1}}{q_n} < \infty,$$

entonces f es linealizable en z_0.

J.C. Yoccoz demostró en 1988 que para la familia de polinomios cuadráticos, la condición de Brjuno es necesaria para tener linearización [Y], [Br], pero en general no se tiene una condición necesaria.

3. Conjuntos de Julia y conjuntos de Fatou

Para definir los conjuntos de Julia y de Fatou de una aplicación racional recordemos la definición de familia normal.

Definición (3.1). Sea $U \subset \widehat{\mathbb{C}}$ un conjunto abierto y conexo. Sea $\mathcal{F} = \{f : U \to \widehat{\mathbb{C}}\,\}$ una familia de funciones analíticas en U. La familia $\mathcal{F}$ es normal en $z_0 \in U$ si para toda sucesión $\{f_n\} \subset \mathcal{F}$ existe una subsucesión $\{f_{n_k}\}$ que converge uniformemente en subconjuntos compactos de U, en una vecindad de z_0, a una función f_0.

Definición (3.2). Sea R una aplicación racional. Decimos que $z_0 \in \widehat{\mathbb{C}}$ pertenece al conjunto de Fatou de R, F_R, si la familia $\{R^n\}$ es normal en una vecindad de z_0. El conjunto de Julia J_R se define como el complemento de F_R.

Obsérvese que si P es un polinomio de grado $d > 1$, entonces el infinito es un punto fijo super-atractor y por el teorema de Böttcher, existe una vecindad U del infinito donde P es analíticamente conjugado a la función z^d. En consecuencia, la órbita de cualquier punto $z \in U$ converge al infinito y por lo tanto, $z \in F_P$.

Definimos el *dominio de atracción del infinito*

$$A_P(\infty) = \{z \in \mathbb{C} : \lim_{n\to\infty} P^n(z) = \infty\,\} \subset F_P,$$

y el *conjunto de Julia lleno* de P

$$K_P = \{z \in \mathbb{C} : \text{ la órbita } \mathcal{O}_P(z) \text{ es acotada }\}.$$

Como el interior de K_P está contenido en F_P, se tiene que J_P es igual a la frontera de K_P y también es igual a la frontera de $A_P(\infty)$.

TEOREMA (3.3) (Fatou-1919). *Sea P un polinomio. El conjunto K_P es conexo si y sólo si la órbita de cada punto crítico de P es acotada.*

Si R es una aplicación racional, se tienen las siguientes propiedades de los conjuntos J_R y F_R que pueden ser consultadas en [B], [CG].

Propiedades:

1. El conjunto J_R es compacto, perfecto y diferente del vacío.
2. Los conjuntos J_R y F_R son completamente invariantes, es decir, $R^{-1}(J_R) = R(J_R) = J_R$ y de igual manera para F_R.
3. Si f denota la k-ésima iterada de R para alguna $k \in \mathbb{N}$, es decir $f = R^k$, entonces $J_R = J_f = J_{R^k}$ y $F_R = F_f = F_{R^k}$.
4. Si $z \in J_R$, entonces el conjunto $\bigcup_{n=1}^{\infty} R^{-n}(z)$ es denso en J_R.
5. Sea z un punto periódico de periodo k de R.
   1. Si z es atractor, entonces $z \in F_R$.
   2. Si z es repulsor, entonces $z \in J_R$.
6. Los puntos periódicos repulsores de R son un conjunto denso en J_R, es decir, $J_R = \overline{\{\text{puntos periódicos repulsores de } R\}}$.

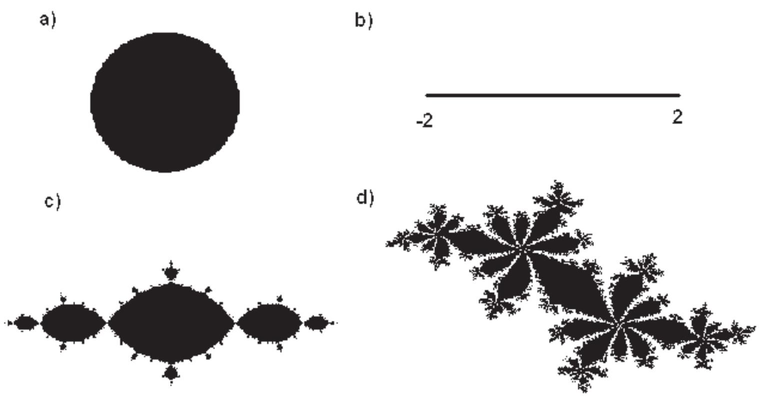


Figure 1. Conjuntos de Julia lleno K_c para: a) c=0, b) c=-2, c) c=-1 y d) c=-0.36+0.62i

Definición (3.4). Sea R una aplicación racional. El *dominio de atracción* $A_R(z_0)$ de un punto fijo atractor z_0 es el conjunto

$$A_R(z_0) = \{z \in \mathbb{C} : \lim_{n\to\infty} R^n(z) = z_0\}.$$

En el caso que $\zeta = \{z_1, z_2, \ldots, z_k\}$ es una órbita atractora de periodo k, entonces z_j es un punto fijo de R^k para cada $j = 1, ..., k$ y el dominio de atracción de ζ es la unión de los dominios de atracción $A_{R^k}(z_j)$ de cada z_j con respecto a R^k, es decir,

$$A_R(\zeta) = \bigcup_{j=1}^{k} A_{R^k}(z_j).$$

El dominio inmediato de atracción del ciclo ζ, denotado por $A_R^*(\zeta)$, es la unión de las k componentes de $A_R(\zeta)$ que contienen al ciclo.

TEOREMA (3.5) (Fatou-1919). *Si z_0 es un punto periódico atractor de R, entonces el dominio inmediato de atracción $A_R^*(z_0)$ contiene al menos un punto crítico.*

4. Clasificación de componentes de Fatou

La dinámica de una aplicación racional en su conjunto de Fatou fue completamente descrita por D. Sullivan. Como el conjunto de Fatou es completamente invariante, la imagen de una de sus componentes conexas es otra componente conexa. Una componente conexa U del conjunto de Fatou es *periódica* si existe un entero $n > 0$ tal que $R^n(U) = U$. Fatou y Julia habían entendido completamente la dinámica de R^n en una componente periódica U, para la cual existen cinco posibilidades.

i) Existe un punto fijo atractor de R^n en U y U está contenido en el dominio de atracción del punto fijo.

ii) Existe un punto fijo super-atractor de R^n en U y U está contenido en el dominio de atracción del punto fijo.

iii) Existe un único punto fijo de R^n en la frontera de U, este punto fijo es racional indiferente y U está contenido en el dominio de atracción de este punto fijo.

iv) La restricción de R^n a U es conformemente conjugada a una rotación irracional $R_\lambda : \mathbb{D} \to \mathbb{D}$, $R_\lambda(z) = \lambda z$.

v) La restricción de R^n a U es conformemente conjugada a una rotación irracional R_λ de un anillo $A_r = \{z : r < |z| < 1\}$ en si mismo.

La clasificación de componentes periódicas del conjunto de Fatou está contenida en los trabajos de Fatou y Julia, pero la existencia de dominios del tipo iv) y v) fue demostrado posteriormente por Siegel y Herman, respectivamente, [D2], [Si].

Un conjunto X se dice *errante* si $R^n X \cap R^m X = \emptyset$ para toda $n > m \geq 0$. La posible existencia de componentes errantes del conjunto de Fatou F_R fue la principal dificultad para investigar la dinámica de R. Sullivan quitó este obstáculo a principios de los años 80.

TEOREMA (4.1) (Sullivan-1985). *El conjunto de Fatou F_R de una aplicación racional no tiene componentes errantes. Es decir, cada componente de F_R es eventualmente periódica.*

Este teorema nos dice que la órbita de cualquier componente de Fatou de una aplicación racional termina en una componente periódica. Sin embargo,

I. N. Baker mostró que los conjuntos de Fatou de algunas funciones enteras en $\mathbb{C}$ tienen dominios errantes, [Ba].

El conjunto postcrítico

$$\mathcal{P}_R = \overline{\{R^n(z_0) : z_0 \text{ es punto crítico de } R \text{ y } n \in \mathbb{N}\}},$$

es decir, la cerradura de la unión de las órbitas de todos lo puntos críticos de una aplicación racional R, está estrechamente ligado a la dinámica de R, como resumimos en el siguiente resultado, consecuencia de los trabajos de Fatou, [CG].

PROPOSICIÓN (4.2). *El conjunto postcrítico $\mathcal{P}_R$ contiene, los ciclos atractores de R, los ciclos indiferentes que pertenecen al conjunto de Julia y la frontera de cada disco de Siegel y anillo de Herman.*

5. Hiperbolicidad

Una idea central en dinámica, que fue desarrollada en los años 1960s y 1970s (por Smale, Anosov, Sinai y muchos otros), es la idea de hiperbolicidad. En esta sección daremos la definición de aplicaciones racionales hiperbólicas y mostraremos las propiedades que hacen a este tipo de aplicaciones bien comportadas.

Definición (5.1). Sea $R\colon \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ una aplicación racional de grado $d > 1$. La aplicación R es *hiperbólica* si las órbitas de los puntos críticos convergen a los ciclos periódicos atractores de R.

De entre todas las funciones racionales, las hiperbólicas son las mejor comportadas, ya que cuando R es hiperbólica, existe un conjunto finito $A \subset \widehat{\mathbb{C}}$ que atrae a un subconjunto abierto de $\widehat{\mathbb{C}}$ de medida total. El siguiente resultado caracteriza la propiedad de ser hiperbólico.

TEOREMA (5.2) (Caracterización de Hiperbolicidad). *Sea R una aplicación racional de grado mayor que uno. Las siguientes condiciones son equivalentes:*

1. *El conjunto postcrítico $\mathcal{P}_R$ (esto es, la cerradura de las órbitas de los puntos críticos de R) es ajeno al conjunto de Julia J_R.*
2. *No hay puntos críticos o ciclos parabólicos en el conjunto de Julia.*
3. *Cada punto crítico de R converge a un ciclo atractor bajo iteración positiva.*
4. *Existe una métrica conforme suave ρ definida en una vecindad del conjunto de Julia tal que $||R'(z)||_\rho > C > 1$ para toda $z \in J_R$.*
5. *Existe un entero $n > 0$ tal que R^n expande estrictamente la métrica esférica en el conjunto de Julia.*

Demostración. Si $|\mathcal{P}_R| = 2$, entonces R es conjugada a z^n y se verifican todas las condiciones del teorema. Supongamos entonces que $|\mathcal{P}_R| > 2$.

Si $\mathcal{P}_R \cap J_R = \emptyset$, entonces no hay puntos críticos o parabólicos en el conjunto de Julia (cada punto parabólico atrae un punto crítico). Por la proposición (4.2), si no hay puntos críticos o puntos parabólicos en el conjunto de Julia, no hay dominios parabólicos, ni discos de Siegel o anillos de Herman y por lo tanto cada punto crítico bajo iteración converge a un ciclo atractor. Esta última condición implica que $\mathcal{P}_R \cap J_R = \emptyset$. Luego, $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 1$.

Si la condición 3 es cierta, se tiene que $\mathcal{P}_R \cap J_R = \emptyset$; además, $\mathcal{P}_R$ y $Q_R = R^{-1}(\mathcal{P}_R)$ son conjuntos numerables con sólo un número finito de puntos límite. Luego, $\widehat{\mathbb{C}} \backslash \mathcal{P}_R$ y $\widehat{\mathbb{C}} \backslash Q_R$ son conexos y

$$R\colon \widehat{\mathbb{C}} \backslash Q_R \to \widehat{\mathbb{C}} \backslash \mathcal{P}_R$$

es una aplicación cubriente por lo tanto, una isometría con respecto de las correspondientes métricas hiperbólicas. Como $|\mathcal{P}_R| > 2$, $Q_R \backslash \mathcal{P}_R$ es distinto del vacío, de donde, la inclusión

$$i\colon \widehat{\mathbb{C}} \backslash Q_R \to \widehat{\mathbb{C}} \backslash \mathcal{P}_R$$

es una contracción. Luego, R expande la métrica hiperbólica en $\widehat{\mathbb{C}} \backslash \mathcal{P}_R$, y la expansión es estricta en el conjunto de Julia ya que J_R es un subconjunto compacto de $\widehat{\mathbb{C}} \backslash \mathcal{P}_R$, así se tiene $3 \Rightarrow 4$.

Cualesquiera dos métricas conformes definidas cerca del conjunto de Julia son quasi-isométricas y el factor de expansión de R^n es mayor que la constante de quasi-isometría cuando n es suficientemente grande, [McM1]. Luego, $4 \Rightarrow 5$.

Finalmente, si R^n expande una métrica conforme en el conjunto de Julia, entonces J_R no contiene puntos críticos o ciclos parabólicos, luego $5 \Rightarrow 2$. $\square$

A las aplicaciones racionales hiperbólicas algunas veces se les llama *expansivas* o que satisfacen el *Axioma A* de Smale.

De la propiedad 5 de hiperbolicidad se tiene el siguiente resultado:

TEOREMA (5.3). *El conjunto de Julia de una aplicación racional hiperbólica tiene medida de Lebesgue cero.*

De hecho, la dimensión de Hausdorff del conjunto de Julia de una aplicación racional hiperbólica es estrictamente menor que dos, [S2].

Uno de los problemas centrales en dinámica holomorfa es la siguiente conjetura que se remonta de alguna forma a los tiempos de Fatou.

Conjetura (5.4) (Densidad de hiperbolicidad). El conjunto de aplicaciones racionales hiperbólicas es abierto y denso en el espacio Rat_d de todas las aplicaciones racionales de grado d.

La propiedad de que el conjunto de aplicaciones racionales hiperbólicas sea abierto es consecuencia del teorema de la función implícita, pero la propiedad de densidad es difícil de mostrar y es conocida únicamente para familias muy particulares, [GS].

6. Estabilidad

La estabilidad es otra de las ideas básicas en el estudio dinámico de las aplicaciones racionales y, como veremos, está fuertemente relacionada con el concepto de hiperbolicidad.

(6.1) Movimientos Holomorfos.

Definición (6.1.1). Sea X una variedad compleja, conexa.

i) Una *familia holomorfa de aplicaciones racionales*, parametrizada por X, es una aplicación holomorfa $f\colon X \times \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$, tal que $f_\lambda\colon \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ es una aplicación racional para cada $\lambda \in X$. Denotemos esta aplicación por $f_\lambda(z)$, donde $\lambda \in X$ y $z \in \widehat{\mathbb{C}}$.

ii) Sea x un punto base en X. Un *movimiento holomorfo* de un conjunto $E \subset \widehat{\mathbb{C}}$ parametrizado por (X, x), es una familia de funciones

$$\phi \colon X \times E \to \widehat{\mathbb{C}},$$

tal que

1. ϕ_λ es inyectiva para cada $\lambda \in X$.
2. $\phi_\lambda(z)$ es una función holomorfa de λ para cada z fija.
3. $\phi_x =$id.

Una propiedad fundamental de movimientos holomorfos es la siguiente.

TEOREMA (6.1.2) (λ-Lema). *Un movimiento holomorfo de E tiene una única extensión a un movimiento holomorfo de $\bar{E}$. La extensión de este movimiento genera una función continua $\phi \colon X \times \bar{E} \to \widehat{\mathbb{C}}$. Para cada λ, la función $\phi_\lambda \colon E \to \widehat{\mathbb{C}}$ se extiende a una función quasiconforme de $\widehat{\mathbb{C}}$ en si misma.*

Dada una familia holomorfa de funciones racionales f_λ, se dice que la correspondiente familia de conjuntos de Julia $J_\lambda \subset \widehat{\mathbb{C}}$ *se mueve holomorfamente*, si existe un movimiento holomorfo

$$\phi_\lambda \colon J_x \to \widehat{\mathbb{C}}, \quad (x \neq \lambda)$$

tal que $\phi_\lambda(J_x) = J_\lambda$ con

$$\phi_\lambda \circ f_x(z) = f_\lambda \circ \phi_\lambda(z)$$

para toda $z \in J_x$. Esto es, ϕ_λ es una conjugación entre f_x y f_λ en sus respectivos conjuntos de Julia. El movimiento ϕ_λ es único, si existe, por la densidad de los ciclos periódicos en J_x. Los conjuntos de Julia se mueven holomorfamente en x si ellos se mueven holomorfamente en alguna vecindad U de x en X. Finalmente, un punto periódico z de f_x de periodo n es *persistentemente indiferente* si existe una vecindad U de x y una función holomorfa $w \colon U \to \widehat{\mathbb{C}}$, tal que $w(x) = z$, $f_\lambda^n(w(\lambda)) = w(\lambda)$ y $|(f_\lambda^n)'(w(\lambda))| = 1$ para todo λ en U, [MSS], [L1].

Definición (6.1.3). Sea f_λ una familia holomorfa de aplicaciones racionales parametrizada por X y sea x un punto en X. Decimos que la familia es estable en x si el conjunto de Julia se mueve holomorfamente en x.

El siguiente resultado, consecuencia del λ-lema, es una caracterización de estabilidad y su demostración puede ser consultada en [McM1], página 54.

TEOREMA (6.1.4) (Caracterización de estabilidad). *Sea f_λ una familia holomorfa de aplicaciones racionales parametrizada por X y sea x un punto en X. Entonces, las siguientes condiciones son equivalentes:*

1. *El número de ciclos atractores de f_λ es localmente constante en x.*
2. *El máximo periodo de un ciclo atractor de f_λ es localmente acotado en x.*
3. *El conjunto de Julia se mueve holomorfamente en x.*
4. *Para toda y suficientemente cercana a x, cada punto periódico de f_y es atractor, repulsor o persistentemente indiferente.*

5. *El conjunto de Julia J_λ depende continuamente en λ (en la topología de Hausdorff de subconjuntos compactos en $\widehat{\mathbb{C}}$) en una vecindad de x.*

Definición (6.1.5). Al conjunto abierto $X^{estable} \subset X$ donde las condiciones anteriores se satisfacen, se le llama conjunto de *parámetros J-estables* de la familia f_λ.

TEOREMA (6.1.6) ([MSS]). *El conjunto $X^{estable}$ de parámetros J-estables es un conjunto abierto y denso de X.*

Como consecuencia se obtiene el siguiente resultado.

TEOREMA (6.1.7). *En cualquier familia holomorfa de aplicaciones racionales, los parámetros hiperbólicos forman un subconjunto abierto y cerrado del conjunto de parámetros J-estables.*

De esta forma se dice que una aplicación racional R de grado d es J-estructuralmente estable si es J-estable en la familia de todas las funciones racionales de grado d.

Conjetura (6.1.8). Una aplicación racional J-estructuralmente estable de grado d es hiperbólica.

De los teoremas (6.1.6) y (6.1.7) se sigue que las conjeturas (5.4) y (6.1.8) son equivalentes.

7. Familia cuadrática

En esta sección vamos a restringir el estudio a la familia $f_c(z) = z^2 + c$ para $c \in X = \mathbb{C}$. Denotemos por $K_c = K_{f_c}$ y $J_c = J_{f_c}$. Del teorema de Fatou (3.3) tenemos:

COROLARIO (7.1). *El conjunto J_c es conexo si y sólo si la órbita de cero es acotada.*

COROLARIO (7.2). *El conjunto J_c es un conjunto de Cantor si y sólo si la órbita de cero converge a infinito, es decir, $0 \in A_c(\infty)$.*

(7.3) Conjunto de Mandelbrot. Definimos el conjunto de Mandelbrot M como,

$$M = \{c \in \mathbb{C} : \ J_c \text{ es conexo}\} = \{c \in \mathbb{C} \ : \ \text{la órbita } \mathcal{O}_c(0) \text{ es acotada }\}.$$

Notemos que si $c = 0$, entonces K_c es el disco unitario cerrado centrado en cero y por lo tanto es conexo. De aqui tenemos que M es diferente del vacío. En esta sección mostraremos más propiedades de M y su relación con la dinámica de f_c.

PROPOSICIÓN (7.3.1). *Sea $S = \max\{2, \mid c \mid\}$. Si $\mid z \mid > S$, entonces*

$$\lim_{n\to\infty} f_c^n(z) = \infty.$$

Observación (7.3.2). Si $\mid c \mid > 2$, entonces

$$\lim_{n\to\infty} f_c^n(0) = \infty.$$

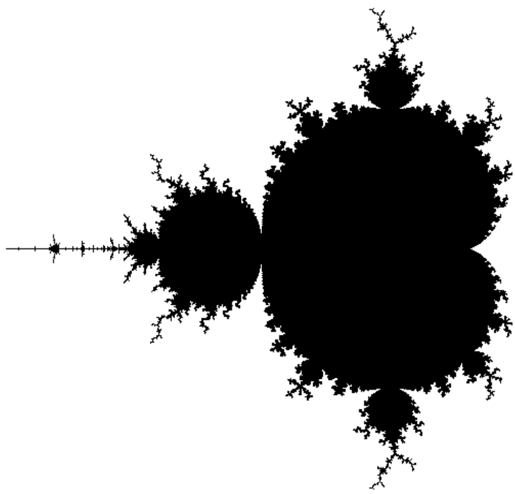


Figure 2. Conjunto de Mandelbrot

COROLARIO (7.3.3). *El conjunto M está contenido en el disco*

$$\overline{D_2(0)} = \{c \in \mathbb{C} \ : \mid c \mid \leqslant 2\}.$$

TEOREMA (7.3.4) (Douady-Hubbard-1982). *El conjunto M es conexo y compacto. Además, $\widehat{\mathbb{C}} \setminus M$ es conexo.*

TEOREMA (7.3.5) (Shishikura-1992). *La frontera del conjunto M tiene dimensión de Hausdorff dos.*

Este resultado proporcionó el primer ejemplo de un conjunto considerado fractal y cuya dimensión es un número entero.

TEOREMA (7.3.6) (McM1). *La frontera del conjunto de Mandelbrot M es igual al conjunto de las $c \in \mathbb{C}$ tales que las funciones $\{c \mapsto f_c^n(0) : n = 1, 2, 3, ...\}$ no forman una familia normal cerca de c. Esto es, $X^{estable} = \mathbb{C} \setminus \partial M$, donde $X^{estable}$ denota el conjunto de parámetros J-estables de la familia f_c.*

Como consecuencia de la definición de hiperbolicidad, se tiene el siguiente resultado.

TEOREMA (7.3.7). *Para c en el conjunto de Mandelbrot, $f_c(z) = z^2 + c$ es hiperbólico si y sólo si f_c tiene un ciclo atractor en $\mathbb{C}$.*

Demostración. Si f_c es hiperbólico y $c \in M$, entonces el punto crítico $z = 0$ converge a una órbita periódica atractora, que debe estar en $\mathbb{C}$ ya que la órbita del punto crítico es acotada. Reciprocamente, si f_c tiene una órbita periódica atractora finita, por el teorema de Fatou (3.5), esta órbita atrae al punto crítico

$z = 0$, luego $c \in M$. Además, como su otro punto crítico $z = \infty$ es un punto fijo super-atractor, f_c es hiperbólico. □

Definición (7.3.8). Una componente U del interior del conjunto de Mandelbrot M es hiperbólica si f_c es hiperbólico para alguna c en U.

Por el teorema (6.1.7), si U es hiperbólica, entonces f_c es hiperbólico para toda c en U. Además, toda componente hiperbólica W del interior de M, es isomorfa al disco $\mathbb{D}$; el isomorfismo está dado por el multiplicador de la órbita atractora y se extiende continuamente a la frontera de W, lo que permite definir el argumento interno o ángulo interno, de todo punto c en la frontera de W, [D1].

De los teoremas (6.1.4) y 7.3 se tiene el siguiente resultado.

TEOREMA (7.3.9). *Si f_c tiene un ciclo indiferente, entonces c pertenece a la frontera del conjunto de Mandelbrot.*

La conjetura (5.4) se traduce para la familia cuadrática en la siguiente:

Conjetura (7.3.10). El conjunto de las c para las cuales $z^2 + c$ es hiperbólico forma un conjunto abierto y denso de $\mathbb{C}$.

Es claro que f_c es hiperbólico cuando c no pertenece a M, porque el punto crítico converge al punto fijo super-atractor en infinito. Así, una formulación equivalente de la conjetura (7.3.10) es la siguiente:

Conjetura (7.3.11). Cada componente del interior del conjunto de Mandelbrot es hiperbólica.

Conjetura (7.3.12) (MLC). La frontera del conjunto de Mandelbrot es localmente conexa.

La importancia de la conjetura MLC radica en el hecho de que Douady y Hubbard mostraron en [DH1] que MLC implica la conjetura (7.3.11) y, en consecuencia, implica la conjetura de Fatou en la familia cuadrática. Por esta razón esta conjetura ha sido y sigue siendo el centro de estudio de la dinámica holomorfa. Si unicamente consideramos parámetros $c \in \mathbb{R}$, se tiene la familia cuadrática real y para esta subfamilia, la conjetura (7.3.10) fue demostrada independientemente por Graczyk-Świątek y Lyubich [GS], [L], mostrando que las componentes hiperbólicas son abiertas y densas en $\mathbb{R}$.

8. Rayos externos y Equipotenciales

Dado que el complemento del conjunto de Mandelbrot es simplemente conexo, por el teorema de la aplicación conforme de Riemann, existe un isomorfismo ϕ entre $\widehat{\mathbb{C}} \setminus M$ y $\mathbb{D}$. Además, por el teorema de Carathéodory, ϕ se extiende continuamente a la frontera de $\mathbb{D}$ si y sólo si la frontera de M es localmente conexa [McM1], [D1]. De aquí, la conjetura MLC es equivalente a mostrar que el isomorfismo ϕ se puede extender continuamente a la frontera. Para analizar la extensión a la frontera, en primer lugar daremos más información del isomorfismo e introduciremos el concepto de rayos externos a M.

Como el polinomio f_c tiene un punto fijo super-atractor en el infinito, por el teorema de Böttcher existe una vecindad U del infinito donde el polinomio f_c es analíticamente conjugado a la función z^2. Denotemos por ϕ_c al bi-holomorfismo que realiza la conjugación, deja fijo al infinito y es tangente a la identidad en el infinito. Si U es el conjunto máximo donde ϕ_c conjuga f_c a z^2, entonces tenemos dos casos:

1. Si $c \in M$, $U = \widehat{\mathbb{C}} \setminus K_c$.

2. Si $c \notin M$, entonces U es una vecindad del infinito que contiene al valor crítico c.

A partir del bi-holomorfismo ϕ_c se puede definir la función

$$\begin{array}{rcl} \Phi_M \colon \widehat{\mathbb{C}} \setminus M & \to & \widehat{\mathbb{C}} \setminus \overline{\mathbb{D}} \\ c & \mapsto & \phi_c(c). \end{array}$$

Douady y Hubbard demostraron que esta función es un bi-holomorfismo y relaciona el espacio dinámico con el espacio de parámetros. Para entender el comportamiento de Φ_M en la frontera vamos a definir los rayos externos a M y a J_c. Si $\theta \in \mathbb{T} = \mathbb{R}/\mathbb{Z}$, entonces el rayo externo a M de ángulo θ es el conjunto

$$R_M(\theta) = \Phi_M^{-1}(\{z \in \mathbb{C} \,:\, z = re^{2\pi i\theta},\ 1 < r < \infty\}).$$

Si el $\lim_{r\to 1} R_M(\theta) = c$, se dice que el rayo de ángulo θ aterriza en c y que c tiene a θ como argumento externo.

Se define el equipotencial $E_M(r)$ de M, de radio $r > 1$, como

$$E_M(r) = \Phi_M^{-1}(\{z \in \mathbb{C} \,:\, z = re^{2\pi i\theta},\ \theta \in \mathbb{T}\}).$$

Ambas definiciones son válidas para los conjuntos de Julia lleno conexos, si sustituimos a Φ_M por ϕ_c.

TEOREMA (8.1). *(Douady-Hubbard-1982) Sea c un parámetro en la frontera de una componente hiperbólica W de M y con ángulo interno $t \in \mathbb{T}$.*

1. *Si t es racional y $c \neq 1/4$, entonces c tiene dos argumentos externos, es decir, hay dos ángulos θ_1, θ_2 tales que los rayos externos $R_M(\theta_i)$ aterrizan en c, para $i = 1, 2$. Además, los rayos $R_c(\theta_i)$ aterrizan en un punto de la frontera de la componente del interior de K_c que contiene a c y son adyacentes a ésta.*

2. *Si t es irracional, entonces existe un único ángulo θ tal que $R_M(\theta)$ aterriza en c.*

Además, Douady y Hubbard demostraron que todos los rayos externos de ángulo racional θ aterrizan en la frontera de M y de hecho mostraron que si θ es periódico bajo la función 2θ, entonces $R_M(\theta)$ aterriza en un parámetro c parabólico (f_c tiene una órbita parabólica) y en caso contrario, $R_M(\theta)$ aterriza en un parámetro de Misiurewicz c, (cero es pre-periódico). Por otro lado, Yoccoz demostró que se tiene conexidad local en todos los parámetros que se encuentran en la frontera de una componente hiperbólica de M, [H].

9. Aplicaciones de tipo cuadrático

Comenzaremos esta sección con una nota sobre el teorema de Sullivan, ya que éste resolvió una conjetura importante y aportó una nueva herramienta en el estudio de sistemas dinámicos. Sullivan había estudiado el trabajo de Ahlfors en grupos Kleinianos; la técnica de Ahlfors hizo uso de la teoría de aplicaciones quasiconformes y Sullivan se dió cuenta que la misma técnica podía ser usada en dinámica holomorfa. Desde el punto de vista de Sullivan, existe un diccionario que relaciona grupos Kleinianos con dinámica holomorfa y la prueba de la conjetura de Fatou, demostraba el poder de este programa. Cuando Douady y Hubbard estudiaron el método utilizado por Sullivan, comprendieron el poder que las funciones quasiconformes podian tener en dinámica holomorfa y comenzaron a utilizarlas. Actualmente, la teoría de aplicaciones quasiconformes, es una de las principales herramientas en dinámica holomorfa.

(9.1) Aplicaciones Quasiconformes. Las aplicaciones quasiconformes son suficientemente regulares para ser objeto de análisis, pero por otra parte, son bastante irregulares que producen objetos geométricos fractales, por ejemplo, curvas de Jordan con dimensión Hausdorff mayor que uno.

Definición (9.1.1). Un homeomorfismo $f\colon X \to Y$ entre superficies de Riemann X, Y es K-*quasiconforme*, con $K \geq 1$, si para todo anillo topológico $A \subset X$,

$$\frac{1}{K}\,\mathrm{mod}(A) \leq \mathrm{mod}(f(A)) \ \leq K\,\mathrm{mod}(A).$$

Existen otras definiciones que son equivalentes a la anterior [LV]. Para mencionar una de ellas, vamos a usar la siguiente notación: $dz = dx + idy$, $d\bar{z} = dx - idy$ y

$$f_z \equiv \frac{\partial f}{\partial z} = \frac{1}{2}(f_x - if_y), \quad f_{\bar{z}} \equiv \frac{\partial f}{\partial \bar{z}} = \frac{1}{2}(f_x + if_y).$$

Un homeomorfismo $f\colon X \to Y$ es K-quasiconforme si localmente tiene derivadas distribucionales en $L^2(X)$ y si la dilatación compleja μ, dada localmente por

$$\mu(z)\frac{d\bar{z}}{dz} = \frac{f_{\bar{z}}}{f_z} = \frac{\partial f/\partial \bar{z}}{\partial f/\partial z}\frac{d\bar{z}}{dz},$$

satisface

$$|\mu| \leq \frac{K-1}{K+1}, \quad \text{casi dondequiera con la medida de Lebesgue, [LV]}.$$

La dilatación μ es también llamada *coeficiente de Beltrami* de f y a la ecuación $f_{\bar{z}} = \mu f_z$ se le llama la *ecuación de Beltrami*.

Notemos que $|\mu| < 1$ si f preserva la orientación y que $\mu = 0$ si y sólo si f es conforme. De aquí, f es 1-quasiconforme si y sólo si f es conforme.

La gran flexibilidad de las aplicaciones quasiconformes viene del hecho de que cualquier diferencial de Beltrami μ, con supremo esencial menor que 1, es realizable por una aplicación quasiconforme [A], [LV].

TEOREMA (9.1.2) (Ahlfors-Bers). *Para cualquier diferencial de Beltrami μ en $L^\infty(\mathbb{C})$ con $\|\mu\|_\infty < 1$, existe una única aplicación quasiconforme $\phi\colon \widehat{\mathbb{C}} \longrightarrow \widehat{\mathbb{C}}$, tal que ϕ fija 0, 1, ∞ y la dilatación compleja de ϕ es μ.*

En particular, este hecho es bastante útil si consideramos una aplicación quasiconforme f que deja invariante una diferencial de Beltrami μ, esto es, $f^*\mu = \mu$ y $f_*\mu = \mu$. Luego, conjugando f con la solución dada por el teorema de Ahlfors-Bers, obtenemos una función holomorfa. Esta es la propiedad principal de funciones quasiconformes utilizada en dinámica holomorfa.

Una de las herramientas en dinámica holomorfa que usa aplicaciones quasiconformes es la cirugía quasiconforme, la cual consiste en tomar dos sistemas dinámicos actuando en diferentes partes del plano y construir un nuevo sistema dinámico que combine la dinámica de ambos, [BD], [BF], [DH2], [H2], [Pe], [PeZa], [Sh2].

(9.2) Definición de aplicación de tipo cuadrático. En las imágenes de M generadas por computadora, podemos observar muchas copias pequeñas del conjunto de Mandelbrot que forman parte de M. Para explicar estas copias pequeñas de M e interesados en la auto-similaridad parcial del conjunto M, Douady y Hubbard introdujeron el concepto de *renormalización compleja*, [DH1]. Sin embargo el espacio de polinomios cuadráticos no resultó invariante bajo este operador de renormalización y fue necesario definir un espacio más grande, el de las aplicaciones de tipo cuadrático[1], [DH2].

Una aplicación holomorfa $f\colon U \to U'$ es llamada *aplicación de tipo cuadrático* si es una aplicación cubriente ramificada de grado 2 entre dos discos topológicos U, U' en $\mathbb{C}$ tal que la cerradura de U es un subconjunto compacto de U', $U \Subset U'$. Esta aplicación tiene sólo un punto crítico y podemos suponer que está en el origen 0.

El conjunto de Julia lleno de f se define como el conjunto de puntos que *no escapan* bajo iteración: $K(f) = \{z : f^n z \in U,\ n = 0, 1, ...\}$ y su frontera es llamada el conjunto de Julia, $J(f) = \partial K(f)$. De igual manera que en los polinomios cuadráticos, los conjuntos $K(f)$ y $J(f)$ son conexos si y sólo si $0 \in K(f)$. En caso contrario, estos conjuntos son Cantor.

El anillo fundamental A de una aplicación de tipo cuadrático $f\colon U \to U'$, es el anillo topológico entre el dominio y el rango de f, $A = U' \backslash \bar{U}$. Una aplicación de tipo cuadrático $f\colon U \to U'$ es *real* si los dominios U y U' son $\mathbb{R}$-simétricos y f conmuta con $z \mapsto \bar{z}$.

(9.3) Espacio de gérmenes de aplicaciones de tipo cuadrático. Una aplicación de tipo cuadrático $g\colon V \to V'$ es un *ajuste* de otra aplicación de tipo cuadrático $f\colon U \to U'$ si $V \subset U$, $g = f|_V$ y $\partial V' \subset U' \backslash \bar{U}$. (En particular, podemos restringir f a $V = f^{-1}U$, con la condición $f(0) \in U$). Decimos que dos aplicaciones de tipo cuadrático f y $\tilde{f}$ representan el mismo germen de tipo cuadrático si existe una sucesión de aplicaciones de tipo cuadrático $f = f_0, f_1, ..., f_n = \tilde{f}$, tal que f_{i+1} se obtiene por un ajuste de f_i o viceversa. Un germen de tipo cuadrático tiene un conjunto de Julia lleno bien definido [McM1], §5.5.

Cualquier polinomio cuadrático f_c determina un germen de tipo cuadrático, cuando se restringe f_c a la preimagen $f_c^{-1}(\mathbb{D}_r)$ de un disco de radio suficientemente grande $\mathbb{D}_r$ y a estos gérmenes se les llama "polinomios cuadráticos".

Denotemos por $\mathcal{Q}$ al espacio de gérmenes de tipo cuadrático y por $\mathcal{C}$ a su lugar de conexidad, esto es, el subconjunto de gérmenes con conjunto de Julia conexo. Lyubich en [L4] dotó a $\mathcal{Q}$ con una topología y estructura analítica compleja modelada en una familia de espacios de Banach $\mathcal{B}_V$, [L]. Este espacio junto con su estructura analítica ha sido fundamental en el estudio de la familia cuadrática; por ejemplo, ha permitido a Lyubich demostrar varias de

[1]Quadratic-like

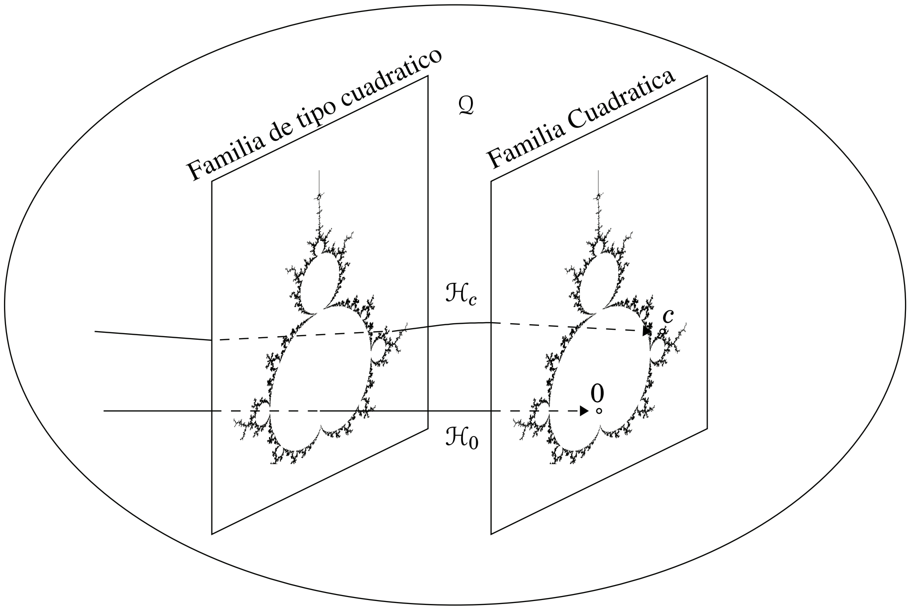


Figure 3. El espacio de gérmenes de aplicaciones de tipo cuadrático.

las conjeturas de Milnor acerca del conjunto de Mandelbrot, entre ellas, la autosimilaridad del conjunto de Mandelbrot alrededor de ciertos parámetros como el Feigenbaum y los de combinatoria acotada. Así como la densidad en pequeñas escalas[2], [L3].

(9.4) Conjugación y clases híbridas. Dos aplicaciones de tipo cuadrático $f: U \to U'$ y $\tilde{f}: V \to V'$ son *topológicamente (quasiconformemente) conjugadas* si existe un homeomorfismo (quasiconforme) $h: (U', U) \to (V', V)$ tal que $h(fz) = \tilde{f}(hz)$, $z \in U$. Dos gérmenes de tipo cuadrático f y $\tilde{f}$ son topológicamente (quasiconformemente) conjugados si existe una elección de aplicaciones de tipo cuadrático, representantes de f y $\tilde{f}$, topológicamente (quasiconformemente) conjugadas. Si $\partial h/\partial\bar{z} = 0$ casi dondequiera en el conjunto de Julia lleno, entonces f y $\tilde{f}$ son *híbridos equivalentes*. Sea $\mathcal{H}(f)$ la clase híbrida de $f \in \mathcal{Q}$. La relación entre elementos de $\mathcal{C}$ y la familia cuadrática está dada por el siguiente teorema fundamental en dinámica holomorfa.

TEOREMA (9.4.1) (Rectificación). [3] [DH2] *Si f es una aplicación de tipo cuadrático con conjunto de Julia conexo, entonces su clase híbrida $\mathcal{H}(f)$ contiene un único polinomio cuadrático $P: z \mapsto z^2 + \chi(f)$, donde $c = \chi(f)$ es un punto del conjunto de Mandelbrot M.*

El teorema define una aplicación $\chi: \mathcal{C} \to M$ llamada la rectificación. Además, el polinomio P y la función quasiconforme h, que conjuga P y f,

[2]Hairness
[3]Straightening

están determinados unicamente por la elección de una función quasiconforme equivariante

$$H : \mathbb{C} \backslash U \to \mathbb{C} \backslash \mathbb{D}_r, \quad \text{tal que} \quad H(fz) = P_0(Hz) \quad \text{para } z \in \partial U\,,$$

donde $P_0(z) = z^2$. A esta función H se le llama el *entubamiento*[4] del anillo fundamental $U' \backslash U$ y al polinomio cuadrático P se le conoce como la *rectificación* de f. Entonces, dada una aplicación de tipo cuadrático con conjunto de Julia conexo, podemos definir *rayos externos* y *equipotenciales* cerca del conjunto de Julia lleno, conjugando la aplicación a un polinomio y transfiriendo las curvas correspondientes.

9.4.2. *Laminación Híbrida.* Por el teorema de rectificación, cada clase híbrida $\mathcal{H}(f)$ en $\mathcal{C}$ intersecta a la familia cuadrática en un sólo punto $c = \chi(f)$ del conjunto de Mandelbrot M. Estas clases se denotan como $\mathcal{H}_c$, $c \in M$.

Lyubich demostró que las clases híbridas $\mathcal{H}_c$, $c \in M$, son subvariedades holomorfas de $\mathcal{Q}$, conexas y de codimensión-uno. Además, la familia cuadrática es una subvariedad compleja de $\mathcal{Q}$ de dimensión uno y transversal a dichas subvariedades [L3].

10. Renormalización compleja

¿Cómo entender un sistema dinámico f a pequeña escala? Nosotros debemos tomar una pequeña pieza del espacio dinámico, considerar la *aplicación de primer retorno* a esta pieza, y finalmente ajustarla a su *tamaño original*. El nuevo sistema dinámico es llamado la *renormalización* Rf del sistema original. Puede suceder que Rf sea "similar" a f, y entonces podemos tratar de repetir este procedimiento, y construir la segunda renormalización $R^2 f$, etc. Propiedades asintóticas de esta sucesión de renormalizaciones refleja la micro-estructura del sistema original. Como se mencionó anteriormente, la noción de renormalización compleja fue introducida por Douady y Hubbard [DH2] para explicar la aparición de pequeñas copias del conjunto de Mandelbrot observadas por computadora.

Sea f una aplicación de tipo cuadrático. Supongamos que podemos encontrar discos topológicos $U \Subset U'$ alrededor de 0 y un entero $p > 0$ tal que $g = f^p : U \to U'$ es una aplicación de tipo cuadrático con conjunto de Julia conexo. Supongamos que los pequeños conjuntos de Julia $f^k J(g)$, $k = 0, ..., p-1$, son ajenos dos a dos excepto, tal vez, intersectándose en sus puntos fijos β. Entonces a la función f se le llama *renormalizable* de periodo p y a la función g se le llama *pre-renormalización*. Al germen de la aplicación de tipo cuadrático g, considerada módulo una conjugación afín, se le llama la renormalización Rf de f, [McM1].

(10.1) La Renormalización de la familia cuadrática. A cada parámetro super-atractor $c \neq 0$ en M, se le puede asociar un subconjunto de M que es homeomorfo a M y se le conoce como el conjunto de Mandelbrot *modulado*[5] por c, $c * M$; este conjunto es una pequeña copia de M [M2], [H1]. La raíz de

[4] Tubing
[5] Tuned

$c * M$, r_c, es el punto correspondiente a la cúspide $1/4$ y el centro es el punto c. Decimos que $c * M$ es real si c es real.

Una pequeña copia del conjunto de Mandelbrot es *primitiva*, si su raíz no pertenece a otra componente hiperbólica; en caso contrario se le llama *satélite*. Para cada copia $c * M$, existe $p > 1$ tal que para cualquier $c' \in c * M$ (excepto posiblemente la raíz) y cualquier $f \in \mathcal{H}(c')$, existe un dominio $U \ni 0$ tal que $f^p|_U$ es una aplicación de tipo cuadrático. Entonces, la función $f^p|_U$ es una pre-renormalización(compleja) de f y se dice que f es renormalizable de periodo p. Notemos que $f^p|_U$ no es un polinomio, aún cuando f lo sea, de modo que el proceso de renormalización automáticamente nos lleva a la clase de aplicaciones de tipo cuadrático. Esta pre-renormalización siempre es simple, es decir, las iteradas de $J(f^p|_U)$ bajo f son ajenas ó se intersectan sólo a lo largo de la órbita de los puntos fijos de tipo β, [McM1].

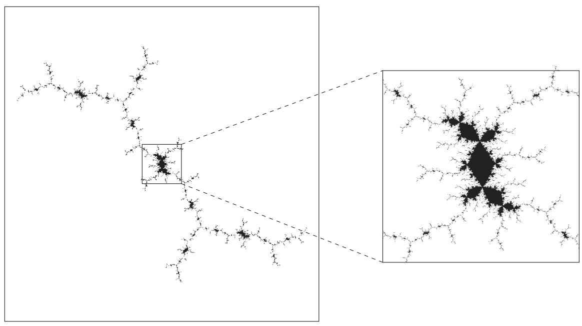


Figure 4. Ejemplo de renormalización de un polinomio cuadrático

El periodo de la copia, $p(c * M)$, es el máximo de tales p y decimos que $c * M$ es máximo si existe sólo una de tales p o, equivalentemente, si no pertenece a ninguna otra copia, excepto M mismo. Esas copias son ajenas y cualquier otra copia, excepto M mismo, pertenece a una única copia máxima.

Todas las copias máximas son primitivas excepto las que se generan en la primera bifurcación de la cardioide principal.

Decimos que $c * M$ es real si c es real. La única M-copia real y maximal para la cual el punto raíz no es renormalizable es la copia de periodo dos $M^{(2)}$. También, todas las copias máximas reales son primitivas excepto por la copia de periodo dos $M^{(2)}$. Definamos $\mathcal{H}(c * M)$ como el conjunto de funciones renormalizables $f \in \chi^{-1}(c * M)$.

Todas las copias $M' \neq M$ son obtenidas de M por iteración de la *modulación*

$$M' = c_l * ... * c_1 * M,$$

donde c_k es el centro de la copia máxima $M_k = M_{c_k}$. Esto es, cualesquiera dos M-copias son o ajenas o una está contenida dentro de la otra.

Sea $c*M$ una M-copia máxima con periodo p y supongamos que $f \in \mathcal{H}(c*M)$. Si $f^p|_U, f^p|_{U'}$ son dos pre-renormalizaciones, entonces decimos que $f^p|_U$ y $f^p|_{U'}$ están en la misma clase. De aquí podemos definir la renormalización $R(f)$ como el germen normalizado de cualquier pre-renormalización de periodo p. Esto

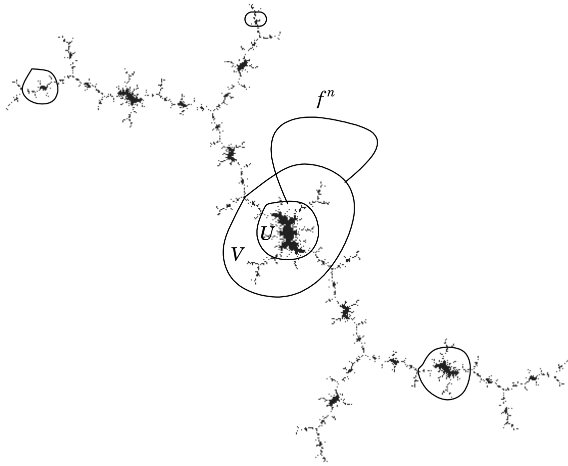


Figure 5. Aplicación de tipo cuadrático generada por la renormalización en la figura 4

es, la renormalización de un germen $R([f])$ es la renormalización de un representante. Una aplicación de tipo cuadrático f es infinitamente renormalizable si $R^n(f)$ está definida para toda $n \geq 0$, i.e., $\chi(f)$ está contenida en infinitas M-copias. El invariante de modulación[6] de una función infinitamente renormalizable f es

$$\tau(f) = \{M_0, M_1, M_2, ...\}$$

donde M_n es la M-copia máxima que contiene a $\chi(R^n(f))$. Decimos que f tiene combinatoria real si todas las M-copias en $\tau(f)$ son reales y una función f infinitamente renormalizable es de tipo acotado, si todos los periodos $p(M_n)$ son acotados. En el caso en que todos los periodos sean iguales, se dice que f es Feigenbaum.

11. Campos de líneas invariantes y conexidad local

Definición (11.1). Sea E un subconjunto de medida positiva del conjunto de Julia J_R de una aplicación racional R. Un campo de líneas en el conjunto de Julia J_R es la asignación de una línea real a través del origen en el espacio tangente en z, para cada z en E, tal que la pendiente es una función medible de z. Un campo de líneas es invariante si $R^{-1}(E) = E$ y si R' manda la línea que pasa por z a la línea que pasa por $R(z)$.

Un acercamiento a las siguientes conjeturas fue desarrollado en [MSS] y [McS], usando aplicaciones quasiconformes. Este punto de vista tiene la

[6]Tuning

ventaja de cambiar el punto de atención de una familia de funciones a la dinámica de una sola función.

Definición (11.2). Sea $X = \mathbb{C}/\Lambda$ un toro complejo. Decimos que $f\colon \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ es cubierta doblemente por un endormofismo del toro, si existen $\alpha \in \mathbb{C}$ y $\mathcal{P}\colon X \to \widehat{\mathbb{C}}$ tales que $|\alpha| > 1$, $\alpha\Lambda \subset \Lambda$, $\mathcal{P}$ es función par, cubriente ramificada de grado 2 de $\widehat{\mathbb{C}}$ y el siguiente diagrama conmuta

$$\begin{tikzcd} \mathbb{C}/\Lambda \arrow[r, "z\mapsto\alpha z"] \arrow[d, "\mathcal{P}"'] & \mathbb{C}/\Lambda \arrow[d, "\mathcal{P}"] \\ \widehat{\mathbb{C}} \arrow[r, "f"'] & \widehat{\mathbb{C}}. \end{tikzcd}$$

Conjetura (11.3) (No existen campos de líneas invariantes). Una aplicación racional R no acepta un campo de líneas invariantes en su conjunto de Julia, excepto cuando R es cubierta doblemente por un endomorfismo del toro.

En el caso de la familia cuadrática obtenemos:

Conjetura (11.4). Un polinomio cuadrático no acepta campos de líneas invariantes en su conjunto de Julia.

Un avance en la solución de estas conjeturas es el siguiente resultado, consecuencia del λ-lema y demostrado por McMullen en 1994, ([McM1], página 61).

TEOREMA (11.5) (Campos de líneas e hiperbolicidad). *Un punto c pertenece a una componente U no hiperbólica del interior del conjunto de Mandelbrot si y sólo si el conjunto de Julia J_c tiene medida positiva y acepta un campo de líneas invariante.*

COROLARIO (11.6). *La familia de polinomios cuadráticos hiperbólicos es densa en la familia cuadrática si y sólo si no existe polinomio cuadrático con una campo de líneas invariante en su conjunto de Julia.*

Esto muestra que la conjetura (11.4) es equivalente a las conjeturas (7.3.10) y (7.3.11).
Yoccoz demostró MLC para parámetros $c \in \partial M$ que pertenecen a la frontera de una componente hiperbólica o que son a los más finitamente renormalizables y no tienen puntos periódicos indiferentes, [H], [M3]. A partir de este resultado, McMullen en [McM1] mostró lo siguiente.

TEOREMA (11.7). *Un polinomio cuadrático que acepta un campo de líneas invariantes en su conjunto de Julia es infinitamente renormalizable.*

TEOREMA (11.8). *El conjunto de Julia de un polinomio cuadrático real no acepta un campo invariante de líneas.*

En particular, cada componente del interior del conjunto de Mandelbrot que intersecta el eje real es hiperbólica.

Por el teorema (11.7), los puntos de la frontera de M donde se desconoce la conexidad local de M son parámetros infinitamente renormalizables. Antes del trabajo de Yoccoz, Douady y Hubbard demostraron MLC para parámetros parabólicos y Misiurewicz (el punto crítico es pre-periódico), [DH1]. Actualmente, la conjetura MLC ha sido ligada a la teoría de renormalización, [L5]. En [L2], Lyubich la demuestra para cierta clase de parámetros infinitamente renormalizables, que forman un conjunto denso en la frontera de M, por medio de un análisis geométrico de los correspondientes conjuntos de Julia. En particular, Lyubich da los primeros ejemplos de parámetros infinitamente renormalizables de tipo acotado donde se tiene MLC; sin embargo, la conjetura permanece sin resolver.

(11.9) Conexidad local de los conjuntos de Julia. En la búsqueda de una demostración de la conjetura MLC, se ha demostrado la conexidad local de los conjuntos de Julia J_c, para una colección grande de parámetros $c \in M$. Douady y Hubbard demostraron la conexidad local de los conjuntos de Julia de polinomios cuadráticos hiperbólicos, de polinomios con una órbita parabólica o de polinomios con el punto crítico pre-periódico (de Misiurewicz) [DH1]. Además, Douady demostró la existencia de conjuntos de Julia que no son localmente conexo, tal es el caso de los conjuntos de Julia de polinomios cuadráticos con un punto de Cremer o con un disco de Siegel Δ cuya frontera no contiene al punto crítico, [D2], [He].

En 1998, Levi y van Strien, e independientemente Lyubich y Yampolsky, demostraron la conexidad local, de todos los conjuntos de Julia que corresponden a parámetros reales en M, [LvS], [LY].

En lo que concierne a los polinomios cuadráticos con disco de Siegel, en 1994 Petersen demostró que si el número de rotación es de tipo acotado, entonces el conjunto de Julia es localmente conexo y de medida de Lebesgue cero [Pe]. En 2004, Petersen y Zakeri, generalizaron este resultado para aquellos polinomios cuadráticos con discos de Siegel, cuyo número de rotación $\theta = [a_1, a_2, ...]$, satisface que la sucesión $\{\log(a_n)\}$ crece del orden $\sqrt{n}$, [PeZa].

En relación a la medida de Lebesgue de los conjuntos de Julia J_c, es conocido que los polinomios cuadráticos hiperbólicos, los parabólicos y los que cumplen la condición de Petersen-Zakeri, tienen un conjunto de Julia de medida cero; salvo este conjunto grande de parámetros, se desconoce la medida de Lebesgue de los J_c.

Conjetura (11.9.1). Existe $c \in M$ tal que la medida de Lebesgue de J_c es mayor que cero.

Un resultado que vino a fortalecer esta conjetura de Douady, fue el de Shishikura, quién demostró en 1992 la existencia de un conjunto residual de parámetros c en la frontera de M, para los cuales J_c tiene dimensión de Hausdorff dos, [Sh], [Sh1], es decir, existen parámetros para los cuales la frontera de K_c es lo suficientemente complicada como para tener área. Recientemente, Buff y Cheritat han anunciado la existencia de parámetros $c \in M$ tal que J_c tiene medida de Lebesgue positiva, resolviendo la conjetura (11.9.1), [BC].

En resumen, presentamos el siguiente diagrama que muestra las diferentes relaciones que se tienen entre las principales conjeturas que se han mencionado hasta ahora. La conjetura central en la teoría de iteración de aplicaciones racionales es la conjetura de Fatou (5.4); sin embargo observemos que de la conjetura (11.3) se deducen todas las demás, excepto la conjetura MLC. Note que hemos denotado una flecha con la etiqueta “Cuadr” para denotar la restricción de las conjeturas (5.4) y (11.3) a la familia cuadrática.

$$
\begin{tikzcd}
\text{Conj. (11.3)} \arrow[r, dash, "\Rightarrow"] \arrow[d, dash, "\text{Cuadr}"'] & \text{Conj. (5.4)} \arrow[r, dash, "\Leftrightarrow"] \arrow[d, dash, "\text{Cuadr}"] & \text{Conj. (6.1.8)} \\
\text{Conj. (11.4)} \arrow[r, dash, "\Leftrightarrow"] \arrow[d, dash, "\Leftrightarrow"'] & \text{Conj. (7.3.10)} \arrow[dl, dash, "\Leftrightarrow"'] & \\
\text{Conj. (7.3.11)} \arrow[r, dash, "\Leftarrow"] & \text{MLC.} &
\end{tikzcd}
$$

Received August 17, 2005

Final version received March 30, 2006

Gamaliel Blé González
División Académica de Ciencias Básicas
Universidad Juárez Autónoma de Tabasco
Km. 1 Carr. Cunduacán-Jalpa
86690 Cunduacán, Tabasco
México
gble@ujat.mx

Rogelio Valdez Delgado
Facultad de Ciencias
Universidad Autónoma del Estado de Morelos
Av. Universidad 1001. Col. Chamilpa
62209 Cuernavaca, Morelos
México
rogelio@matcuer.unam.mx

References

[A] L. Ahlfors, Lectures on Quasiconformal Mappings. D. Van Nostrand, Princeton, NJ, 1966.

[Ba] I. N. Baker, *An entire function which has wandering domains*. J. Austral. Math. Soc. (Series A). **22**, (1976), 173-176.

[B] A. Beardon, Iteration of Rational Functions. Springer Verlag, New York, 1991.

[BD] B. Branner y A. Douady, *Surgery on complex polynomials*. Proc. of the Symposium on Dynamical Systems, Mexico 1986. Lecture Notes in Math. **1345**, (1986), 11-72.

[BF] B. Branner y N. Fagella, *Homeomorphisms between limbs of the Mandelbrot set*. J. Geom. Anal. **9**, no. 3, (1999), 327-390.

[Br] A. D. Brjuno, *Convergence of transformations of differential equations to normal forms*. Dokl. Akad. Nauk USSR. **165**, (1965), 987-989.

[BC] X. Buff y A. Cheritat, *Ensembles de Julia quadratiques de mesure de Lebesgue strictement positive*. C. R. Math. Acad. Sci. Paris, **341,** (2005), 669–674.

[CG] L. Carleson y T. Gamelin, Complex Dynamics. Springer Verlag, New York, 1993.

[D1] A. Douady, *Systèmes Dynamiques Holomorphes*, Séminaire Bourbaki, 35é année. **599**, Astérisque 105–106 (1982), 39–63.

[D2] A. Douady, *Disques de Siegel et anneaux de Herman*, Séminaire Bourbaki, 1986-87. **677**, Astérisque 152–153, (1987), 151–172.

[DH1] A. Douady y J. Hubbard, *Étude dynamique des polynômes complexes*, Publications Mathématiques d'Orsay **84-02**, (1984), **85-04**, (1985).

[DH2] A. DOUADY Y J. HUBBARD, *On the dynamics of polynomial-like mappings*, Ann. Sci. Éc. Norm. Sup. **18**, (1985), 287–344.

[F] P. FATOU, *Mémoire sur les équations fonctionnelles*, Bull. S.M.F. **47** (1919) 161-271; **48** (1920), 33–94; 208–314.

[GS] J. GRACZYK Y G. ŚWIĄTEK, The real Fatou conjecture, Annals of Math. Studies **144**, Princeton University Press, Princeton, NJ, 1998.

[H1] P. HAÏSSINSKY, *Modulation dans l'ensemble de Mandelbrot*, The Mandelbrot Set, theme and variations, London Mathematical Society, Lecture Note Series **274**, (2000), 37–65.

[H2] P. HAÏSSINSKY, *Applications de la chirurgie holomorphe, notamment aux points paraboliques*, Ph. D. Thesis, Université de Paris-Sud, 1998.

[HW] G.H. HARDY Y E.M. WRIGHT, An Introduction to the Theory of Numbers, Oxford University Press, New York, 1979.

[He] M. HERMAN, *Are there critical points on the boundaries of a Siegel disk?* Comm. Math. Phys. **99**, (1985), 593–612.

[H] J. H. HUBBARD, *Local connectivity of Julia sets and bifurcation loci: three theorems of J.-C. Yoccoz*, Topological Methods in Modern Mathematics. A Symposium in Honor of John Milnor's 60 th. Birthday. Publish or Perish, Houston, TX, 467–511, 1993.

[J] G. JULIA, *Mémoire sur l'itération des functions rationnelles*, J. Math. Pures Appl. (7th series) **4**, (1918), 47–245.

[L] S. LANG, Differentiable manifolds, Addison-Wesley Publishing Co. Inc, Reading, MA, 1972.

[LV] O. LETHO Y K.J. VIRTANEN, Quasiconformal Mappings in the Plane. Springer-Verlag, Berlin, New York, 1973.

[LvS] G. LEVIN Y S. VAN STRIEN, *Local connectivity of the Julia set of real polynomials*, Ann. of Math. **147**, (1998), 471–541.

[L1] M. LYUBICH, *The dynamics of rational transforms: the topological picture*, Russian Math. Surveys **41:4**, (1986), 43–117.

[L2] M. LYUBICH, *Dynamics of quadratic polynomials*, I-II, Acta Mathematica **178**, (1997), 185–297.

[L3] M. LYUBICH, *Feigenbaum-Coullet-Tresser universality and Milnor's Hairness conjecture*, Ann. of Math. **149**, (1999), 319–420.

[L4] M. LYUBICH, *Almost every real quadratic map is either regular or stochastic*, Ann. of Math. **156**, (2002), 1–78.

[L5] M. LYUBICH, *Renormalization ideas in conformal dynamics*, Current Developments in Mathematics. (1995) International Press, Cambridge, MA, (1994), 155–190.

[LY] M. LYUBICH Y M. YAMPOLSKY, *Dynamics of quadratic polynomials: Complex bounds for real maps*, Ann. Inst. Fourier (Grenoble). **47** (1997), 1219–1255.

[MSS] R. MAÑÉ, P. SAD Y D. SULLIVAN, *On the dynamics of rational maps*, Ann. Sci. Éc. Norm. Sup. **16**, (1983), 193–217.

[M1] J. MILNOR, Dynamics in one complex variable. Third edition, Annals of Math. Studies. **160**, Princeton University Press, Princeton, NJ, 2006.

[M2] J. MILNOR, *Self-similarity and hairness in the Mandelbrot set*, Computers in Geometry and Topology, Lect. Notes in Pure and Appl. Math. **114**, (1989), 211–257.

[M3] J. MILNOR, *Local connectivity of Julia sets: expository lectures*, The Mandelbrot Set, theme and variations, London Math. Soc. Lecture Note Series **274**, (2000), 67–116.

[McM1] C. MCMULLEN, Complex Dynamics and Renormalization. Annals of Math. Studies. **135**, Princeton University Press, Princeton, NJ, 1994.

[McM2] C. MCMULLEN Renormalization and 3-manifolds which Fiber over the Circle. Annals of Math. Studies **142**, Princeton University Press, Princeton, NJ, 1996.

[McS] C. MCMULLEN Y D. SULLIVAN, *Quasiconformal homeomorphism and dynamics III: The Teichmuller space of a holomorphic dynamical system*, Advances in Math. **135**, (1998), 243–260.

[Pe] C. L. PETERSEN *Local connectivity of some Julia sets containing a circle with an irrational rotation*, Acta Mathematica **177**, (1996), 163–224.

[PeZa] C. L. PETERSEN Y S. ZAKERI, *On the Julia set of a typical quadratic polynomial with a Siegel disk*, Annals of Mathematics (2) **159**, (2004), no. 1, 1–52.

[Sh] M. Shishikura, *The boundary of the Mandelbrot set has Hausdorff dimension two*, Complex analytic methods in dynamical systems (Rio de Janeiro, 1992). Astérisque **222**, (1994), 7, 389–405.

[Sh1] M. Shishikura, *The Hausdorff dimension of the boundary of the Mandelbrot set and Julia sets*, Annals of Mathematics. **147**, (1998), 225–267.

[Sh2] M. Shishikura, *On the quasiconformal surgery of rational functions*, Ann. Sci. École Norm. Sup. **20**, (1987), 1–29.

[Si] C. L. Siegel, *Iteration of analytic functions*, Annals of Mathematics **43**, (1942), 607–612.

[S1] D. Sullivan, *Quasiconformal homeomorphism and dynamics* I*, Solution of the Fatou-Julia problem on wandering domains*, Annals of Mathematics **122**, (1985), 401–418.

[S2] D. Sullivan, *Conformal dynamical systems*, Geometric dynamics, Springer-Verlag Lecture Notes **1007**, (1983), 725–752.

[Y] J.-C. Yoccoz, *Linéarisation des germes de difféomorphismes holomorphes de* (**C**, 0). C.R. Acad. Sci. Paris **306**, (1988), 55–58.

Bol. Soc. Mat. Mexicana (3) Vol. 12, 2006

INVARIANTS ASSOCIATED TO THE PURE BRAID GROUP OF THE SPHERE

DANIEL JUAN-PINEDA AND SILVIA MILLAN-LÓPEZ

ABSTRACT. We prove that the Fibered Isomorphism Conjecture of Farrell and Jones holds for the Pure Braid Groups of the 2-sphere.

1. Introduction

The Fibered Isomorphism Conjecture, FIC, of T. Farrell and L. Jones [7], states that for any group G, the algebraic K-theory groups of the group ring $\mathbb{Z}G$ are completely determined by the corresponding algebraic K-theory groups of the group rings of virtually cyclic subgroups of G. FIC has been verified for several groups. For instance, for discrete cocompact subgroups of virtually connected Lie groups by Farrell and Jones in [7] and for a certain family of Mapping class groups by Berkove, Juan-Pineda, and Lu in [3]. In this paper we prove that the Pure Braid Group of the sphere $\mathbb{S}^2$ satisfies FIC, see section 3 for the exact statement.

The pure braid group of the sphere $\mathbb{S}^2$ is defined as the fundamental group of the configuration space $F(\mathbb{S}^2, n)$. A method to study these groups is to associate with them a fibration, see section 2. We use this fibration and the fact that the second homotopy group of some configuration spaces of the sphere are trivial to use an induction argument to prove the result. Our main result is the following:

MAIN THEOREM. *The Pure Braid Group of n strands of the sphere $\mathbb{S}^2$, $B_n(\mathbb{S}^2)$, satisfies FIC for all $n > 0$.*

This paper is divided in three sections: we recall background material in section 2, in section 3 we state our main result, supply its proof and a corollary, and in the last section we recall the general setup for FIC.

Before we start section 1 we recall two important results used in the proof of the main result.

Let $\phi\colon F \longrightarrow F$ be a self homeomorphism. The mapping torus of ϕ, F_ϕ, is by definition the quotient space of $F \times [0, 1]$ where $(x, 1)$ is identified with $(\phi(x), 0)$ and the projection onto the second factor induces a fiber bundle projection $q\colon F_\phi \longrightarrow \mathbb{S}^1$ with fiber F.

2000 *Mathematics Subject Classification*: 19B28, 20F36.

Keywords and phrases: K-theory, braid groups.

Research partially supported by CONACyT and DGAPA research grants. The second author was supported by the Clark Fellowship program from Binghamton University.

LEMMA (1.1). ([1], Lemma 2.4) *Let $p\colon E \longrightarrow M$ be a fiber bundle with $\pi_2(M) = 1$, arc connected fiber F, and $\phi\colon F \longrightarrow F$ be a self homeomorphism. Let $g\colon F_\phi \longrightarrow E$ be a bundle map covering $\alpha\colon \mathbb{S}^1 \longrightarrow M$ and $\hat{\alpha}\colon \mathbb{S}^1 \longrightarrow E$ be a lift of α. Let $[\hat{\alpha}] \in \pi_1(E)$ denote the homotopy class of $\hat{\alpha}$. Then $\phi_\sharp([\beta]) = [\hat{\alpha}][\beta][\hat{\alpha}]^{-1}$ in $Out(\pi_1(F))$, where the fiber containing the base point $e \in E$ is identified with F via g.*

THE ISOTOPY EXTENSION THEOREM. ([9], Ch. 8, Thm. 1.4) *Let $V \subset M$ be a compact submanifold and $F\colon V \times [0, 1] \longrightarrow M$ an isotopy of V. If either $F(V \times I) \subset \partial M$ or $F(V \times I) \subset M\text{-}\partial M$, then F extends to a diffeotopy of M having compact support.*

We thank Tom Farrell for his interest and advice and the referee for his careful reading and suggestions.

2. The Pure Braid Group of the Sphere $\mathbb{S}^2$

Let $Q_m = \{q_1, \ldots, q_m\}$ be a fixed set of m distinct points of a connected space X. Define

$$F(X - Q_m, n) = \{(x_1, \ldots, x_n) \mid x_i \in X - Q_m,\ x_i \neq x_j,\ \text{if } i \neq j\}.$$

Observe that when $m = 0$ in the above definition we obtain

$$F(X, n) = \{(x_1, \cdots, x_n) \mid x_i \in X,\ x_i \neq x_j,\ \text{if } i \neq j\}.$$

Definition (2.1). The Pure Braid Group, $B_n(X)$, of n strands of X is the fundamental group of $F(X, n)$.

The following definition was introduced by Aravinda, Farrell and Roushon to study the Whitehead groups of pure braid groups:

Definition (2.2). ([1], 1.1) A discrete group Γ is called *strongly poly-free* if there exists a finite filtration by subgroups $1 = \Gamma_0 \subseteq \Gamma_1 \subseteq \cdots \subseteq \Gamma_n = \Gamma$ such that the following conditions are satisfied:

(1). Γ_i is normal in Γ for each i.

(2). Γ_{i+1}/Γ_i is a finitely generated free group for all i.

(3). For each $\gamma \in \Gamma$ and each i there is a compact surface F and a diffeomorphism $f\colon F \to F$ such that the induced homomorphism $f_\#$ on $\pi_1(F)$ is equal to c_γ in $Out(\pi_1(F))$, where c_γ is the action of γ on Γ_{i+1}/Γ_i by conjugation and $\pi_1(F)$ is identified with Γ_{i+1}/Γ_i via a suitable isomorphism.

The third condition says that the algebraic action of γ on Γ_{i+1}/Γ_i can be *geometrically realized.*

Now we recall a theorem about fibrations of configuration spaces and two theorems which state the conditions to determine when a group satisfies FIC.

THEOREM (2.3). [5] *Let M denote a manifold of dimension ≥ 2 and $Q_m \subset M$ a finite set of m distinct points. Then the projection to the first r coordinates is a locally trivial fiber bundle,*

$$F(M - Q_m, n) \xrightarrow{p} F(M - Q_m, r)$$

with fiber $F(M - Q_{m+r}, n - r)$.

THEOREM (2.4). ([8], Thm. 1.1). *Let Γ be an extension of a finite group by a strongly poly-free group, i.e.* $1 \longrightarrow H \longrightarrow \Gamma \longrightarrow F \longrightarrow 1$ *is a short exact sequence where H is a strongly poly-free group and F is a finite group. Then Γ satisfies FIC.*

THEOREM (2.5). ([3], Thm. 2.7). *Let Γ be a group such that* $1 \longrightarrow K \longrightarrow \Gamma \longrightarrow G \longrightarrow 1$ *is a short exact sequence where G satisfies FIC, K is a finitely generated free group, and for any $t \in G$ of infinite order, the action of the lift $\hat{t}$ on K can be geometrically realized. Then Γ satisfies FIC.*

3. The Main Result

MAIN THEOREM. *The Pure Braid Group of n strands on the sphere $\mathbb{S}^2$, $B_n(\mathbb{S}^2)$, satisfies FIC for all $n > 0$.*

Proof. We prove the theorem by induction on the number of strands of the pure braid group $B_n(\mathbb{S}^2)$. First we look at the cases for $n = 1, 2, 3, 4$.

For $n = 1$, we have that $B_1(\mathbb{S}^2) = \pi_1(F(\mathbb{S}^2, 1) = \pi_1(\mathbb{S}^2) = 1$ which satisfies FIC.

For $n = 2$, $B_2(\mathbb{S}^2) = \pi_1(F(\mathbb{S}^2, 2))$. To find this group we look at the fiber bundle associated to $F(\mathbb{S}^2, 2)$.

The projection on the first coordinate $F(\mathbb{S}^2, 2) \xrightarrow{p} F(\mathbb{S}^2, 1)$ is a fiber bundle with fiber $F(\mathbb{S}^2 - Q_1, 1) = \mathbb{S}^2 - *$ by Theorem (2.3). Now we apply the homotopy long exact sequence to this fibration to obtain a short exact sequence

$$\cdots \longrightarrow \pi_2(\mathbb{S}^2) \longrightarrow \pi_1(\mathbb{S}^2 - *) \longrightarrow \pi_1(F(\mathbb{S}^2, 2)) \longrightarrow \pi_1(\mathbb{S}^2) \longrightarrow 1$$

which implies that $\pi_1(F(\mathbb{S}^2, 2) = 1$. Thus $B_2(\mathbb{S}^2)$ satisfies FIC.

It is known that $B_3(\mathbb{S}^2) = \pi_1(F(\mathbb{S}^2, 3) = \mathbb{Z}/2$, [6]. Since $\mathbb{Z}/2$ is virtually cyclic, then $B_3(\mathbb{S}^2)$ satisfies FIC.

Now let $n = 4$, $B_4(\mathbb{S}^2) = \pi_1(F(\mathbb{S}^2, 4))$. As before, the projection on the first three coordinates $F(\mathbb{S}^2, 4) \xrightarrow{p} F(\mathbb{S}^2, 3)$ is a fiber bundle with fiber $F(\mathbb{S}^2 - Q_3, 1)$. Using the homotopy long exact sequence and the following facts: $\pi_0(\mathbb{S}^2 - Q_{n-1}) = 1$, for $n > 0$ and $\pi_2(F(\mathbb{S}^2, n-1)) = 1$, for $n > 2$ [6], we have the following short exact sequence

$$1 \longrightarrow \pi_1(\mathbb{S}^2 - Q_3) \longrightarrow \pi_1(F(\mathbb{S}^2, 4)) \longrightarrow \pi_1(F(\mathbb{S}^2, 3)) \longrightarrow 1.$$

As $\pi_1(\mathbb{S}^2 - Q_3)$ is strongly poly-free [1], and $\pi_1(F(\mathbb{S}^2, 3))$ is finite [6], we can apply Theorem (2.4) to conclude that $\pi_1(F(\mathbb{S}^2, 4))$ satisfies FIC.

For the induction step assume $\pi_1(F(\mathbb{S}^2, n-1))$ satisfies FIC for $n > 4$. The projection on the first $n-1$ coordinates $p\colon F(S^2, n) \longrightarrow F(\mathbb{S}^2, n-1)$ is a fiber bundle with fiber $F(\mathbb{S}^2 - Q_{n-1}, 1) = \mathbb{S}^2 - Q_{n-1}$. From the homotopy long exact sequence and the facts used for the case $n = 4$ we have the following short exact sequence

$$1 \longrightarrow \pi_1(\mathbb{S}^2 - Q_{n-1}) \longrightarrow \pi_1(F(\mathbb{S}^2, n)) \xrightarrow{p_\#} \pi_1(F(\mathbb{S}^2, n-1)) \longrightarrow 1$$

where $\pi_1(\mathbb{S}^2 - Q_{n-1})$ is a finitely generated free group. To apply Theorem (2.5) it is left to check the following condition: for any $[t] \in \pi_1(F(\mathbb{S}^2, n-1))$ of infinite order, the action of the lift $[\hat{t}]$ on $\pi_1(\mathbb{S}^2 - Q_{n-1})$ can be geometrically realized, *i.e.*, there exists a compact surface F and a diffeomorphism $f \colon F \longrightarrow F$ such that *i)* $\pi_1(F) = \pi_1(\mathbb{S}^2 - Q_{n-1})$ and *ii)* $f_\sharp$ is conjugation by $[\hat{t}]$. To verify this condition we need an alternative description of the pure braid group.

Let $E(Q_n, \mathbb{S}^2)$ be the space of all embeddings of Q_n in $\mathbb{S}^2$. Note that there is an identification of $E(Q_n,\mathbb{S}^2)$ with $F(\mathbb{S}^2,n)$. Hence $\pi_1(E(Q_n,\mathbb{S}^2)) = \pi_1(F(\mathbb{S}^2,n))$.

Now we proceed to verify the realizability condition. First let $\mathbf{F} = \mathbb{S}^2 - (n-1)\overset{o}{D}$, were $(n-1)\overset{o}{D}$ is the union of $(n-1)$ open discs. Note that $\mathbf{F}$ is compact and that $\pi_1(\mathbb{S}^2 - Q_{n-1}) = \pi_1(\mathbf{F})$.

Next let $[t] \in \pi_1(F(\mathbb{S}^2, n-1)$ be of infinite order; then there exists $[\gamma] \in \pi_1(F(\mathbb{S}^2,n))$ such that $p_\sharp([\gamma]) = [t]$. Fix γ, a representative of the element $[\gamma] \in B_n(\mathbb{S}^2)$. Note that we can regard the loop γ as an isotopy γ_t, $t \in [0,1]$ of the inclusion $Q_n \subset \mathbb{S}^2$ back to itself. Since Q_n is a compact submanifold of $\mathbb{S}^2$, we can apply the Isotopy Extension Theorem to obtain a diffeotopy $\hat{\gamma}_t \colon \mathbb{S}^2 \longrightarrow \mathbb{S}^2$, $t \in [0,1]$ extending γ_t and starting at $id_{\mathbb{S}^2}$, *i.e.*, $\hat{\gamma}_0 = id_{\mathbb{S}^2}$, $\hat{\gamma}_t(j) = \gamma_t(j)$ for all $t \in [0,1]$ and for all $j \in Q_n$.

Observe that $\hat{\gamma}_1 \colon \mathbb{S}^2 \longrightarrow \mathbb{S}^2$ is a diffeomorphism such that $\hat{\gamma}_1(j) = \gamma_1(j)$ for all $j \in Q_n$. Hence $\hat{\gamma}_1$ leaves $\mathbb{S}^2 - Q_n$ invariant. We now use Lemma (1.1) for the fibration $p \colon F(\mathbb{S}^2, n) \to F(\mathbb{S}^2, n-1)$ by letting $f \colon \mathbb{S}^2 - Q_n \longrightarrow \mathbb{S}^2 - Q_n$ be the diffeomorphism $\hat{\gamma}_1$ restricted to $\mathbb{S}^2 - Q_n$, $\hat{\alpha} = \gamma, \alpha = p \circ \gamma,$ and

$$g((z,t)) = \gamma_t(z).$$

Lastly, we apply Theorem (2.5) to conclude that $\pi_1(F(\mathbb{S}^2, n))$ satisfies FIC. $\square$

The fact that FIC holds for $\Gamma = B_n(\mathbb{S}^2)$ tells us that the groups $K_i(\mathbb{Z}\Gamma)$ for $i \le 1$ are determined by the groups $K_i(\mathbb{Z}V)$, where V varies over the virtually cyclic subgroups of Γ. On the other hand it is known, see [4], Corollary 6.3, that $B_n(\mathbb{S}^2) \cong \mathbb{Z}/2 \times B_{n-3}(\mathbb{R}^2 - Q_2)$ for $n > 3$ and that $B_{n-3}(\mathbb{R}^2 - Q_2)$ is torsion-free. Hence the virtually cyclic subgroups of Γ are isomorphic to one of the following groups: $\{1\}$, $\mathbb{Z}/2$, $\mathbb{Z}$ or $\mathbb{Z}/2 \times \mathbb{Z}$. The lower algebraic K theory of the group rings of any of the groups in this list has been calculated, see [2], Ch. VII and Ch. XI. From these calculations we get that the Whitehead group $Wh(V)$, the class group $\tilde{K}_0(\mathbb{Z}V)$, and $K_i(\mathbb{Z}V)$, $i < 0$, vanish for any of the groups in the above list. From this we have

COROLLARY (3.1). *Let Γ be the pure braid group of the sphere. Then $Wh(\Gamma)$, $\tilde{K}_0(\mathbb{Z}\Gamma)$ and $K_i(\mathbb{Z}\Gamma)$ vanish for all $i < 0$.*

4. The Fibered Isomorphism Conjecture

We recall the Fibered Isomorphism Conjecture formulated in [7], 1.6. Let $\mathcal{S} \colon TOP \to SPECTRA$ be a covariant homotopy functor. Let $\mathbf{B}$ be the category of continuous surjective maps: objects in $\mathbf{B}$ are continuous maps $p \colon E \to B$, where E, B are objects in TOP, and morphisms between $p_1 \colon E_1 \to B_1$ and

$p_2 \colon E_2 \to B_2$ consist of continuous maps $f \colon E_1 \to E_2$ and $g \colon B_1 \to B_2$ making the following diagram commute

$$\begin{array}{ccc} E_1 & \xrightarrow{f} & E_2 \\ {\scriptstyle p_1}\Big\downarrow & & {\scriptstyle p_2}\Big\downarrow \\ B_1 & \xrightarrow{g} & B_2. \end{array}$$

In this setup, Quinn [10] constructs a functor from $\mathbf{B}$ to $\Omega - \mathit{SPECTRA}$. The value of this spectrum at $p \colon E \to B$ is denoted by

$$\mathbb{H}(B; \mathcal{S}(p)),$$

and has the property that its value at the object $E \to *$ is $\mathcal{S}(E)$. The map of spectra $\mathbb{A}$ associated to

$$\begin{array}{ccc} E & \xrightarrow{id} & E \\ {\scriptstyle p}\Big\downarrow & & {\scriptstyle p}\Big\downarrow \\ B & \xrightarrow{g} & *. \end{array}$$

is known as the *Quinn* assembly map.

Given a discrete group Γ, let $\mathcal{E}$ be a universal Γ-space for the family of virtually cyclic subgroups of Γ [7], Appendix, and denote by $\mathcal{B}$ the orbit space $\mathcal{E}/\Gamma$. Let X be any free and properly discontinuous Γ- space, and $p \colon X \times_\Gamma \mathcal{E} \to B$ be the map determined by the projection onto $\mathcal{B}$. The *Fibered Isomorphism Conjecture* (FIC) for $\mathcal{S}$ and Γ is the assertion that

$$\mathbb{A} \colon \mathbb{H}(\mathcal{B}; \mathcal{S}(p)) \to \mathcal{S}(X/\Gamma)$$

is a weak equivalence of spectra. This conjecture was made in [7], 1.7, for the functors $\mathcal{S} = \mathcal{P}()$, $\mathcal{K}()$, and $\mathcal{L}^{-\infty}$, the pseudoisotopy, algebraic K-theory and $\mathcal{L}^{-\infty}$-theory functors. In this paper we mean FIC as FIC for the functor $\mathcal{S} = \mathcal{P}()$.

Received October 21, 2004

Final version received October 03, 2005

DANIEL JUAN-PINEDA
INSTITUTO DE MATEMÁTICAS
UNAM CAMPUS MORELIA
APARTADO POSTAL 61-3 (XANGARI)
58089 MORELIA, MICHOACÁN
MÉXICO
daniel@matmor.unam.mx

SILVIA MILLAN-LÓPEZ
DEPARTMENT OF MATHEMATICS
BINGHAMTON UNIVERSITY
BINGHAMTON, N. Y. 13902
USA
millan@math.binghamton.edu

References

[1] C. S. Aravinda, F.T. Farrell, and S.K. Roushon, *Algebraic K-Theory of Pure Braid groups*, Asian J. Math. **4** (2), (2000), 337–344.

[2] H. Bass, Algebraic K-theory, Benjamin, 1968.

[3] E. Berkove, D. Juan-Pineda, Q. Lu, *Algebraic K-theory of mapping class groups*, K-Theory **32** (1), (2004), 83–100.

[4] F. Cohen and J. Pakianathan, *Configuration Spaces and Braid Groups*, available at `www.math.rochester.edu/people/faculty/jonpak/`.

[5] E. Fadell and L. Neuwirth, *Configuration Spaces*, Math. Scand. **10** (1962), 111–118.

[6] E. Fadell and J. Van Buskirk, *The Braid Groups of E^2 and $\mathbb{S}^2$*, Duke Math. J. **29** (1962), 243–258.

[7] F. T. Farrell and L. Jones, *Isomorphism Conjecture in algebraic K-theory*, J. Amer. Math. Soc. **6** (1993), 249–297.

[8] F. T. Farrell and Sayed K. Roushon, The Whitehead Groups of Braid Groups Vanish, International Mathematics Research Notice 2000, No.10.

[9] M. W. Hirsch, Differential Topology, Springer-Verlag, N.Y, 1994.

[10] F. Quinn, *Ends of maps II*, Invent. Math. **68** (1982) (3), 353–424.

Bol. Soc. Mat. Mexicana (3) Vol. 12, 2006

DUALITY AND DOMINATING EXTENSION THEOREMS IN NONCANCELLATIVE NORMED CONES

S. ROMAGUERA, E. A. SÁNCHEZ-PÉREZ, AND O. VALERO

ABSTRACT. Let C be a cone. We study the relation between the structure of the space of linear functions from C to R and the cancellativity of the additive operation in C. In particular, we prove a separation result by defining an equivalence relation Q that is given by the structure of the cone, and we show that a linear function on a subcone S can be extended to the whole cone C whenever its associated linear function on the quotient cone S/Q can be extended to the cancellative cone C/Q. This provides a technique for the generalization of the Hahn-Banach type theorems on extensions of real functionals on noncancellative normed cones.

1. Introduction

Let $\mathbb{R}$ be the set of real numbers and $\mathbb{R}^+$ the set of non-negative real numbers. We say that an algebraic structure $(C, +, \cdot)$ is a *cone* if $(C, +)$ is an abelian monoid and the product $\cdot : \mathbb{R}^+ \times C \to C$ satisfies the usual axioms of a linear space when restricted to nonnegative scalars. We denote by 0 the neutral element of C. If C is a cone, we say that a map $f : C \to R$ is a *functional* if it is a linear map, i.e., if for every pair of elements $x, y \in C$,

$$f(x + y) = f(x) + f(y),$$

and for every $r \in R^+$, $f(rx) = rf(x)$. Note that $f(0) = 0$ for every functional f. We call the set of all these functionals the *algebraic dual* of C, and we denote it by C'.

In this paper we study the relations between the structure of C' and the cancellativity of $(C, +)$. As we will show, this structure depends only on a particular quotient of C which is a cancellative cone that will be called the cancellative quotient of C. Although in recent years there have been obtained several Hahn-Banach extensions theorems for functionals on cones, in the context where these results have been developed it is implicitly assumed that the cone is cancellative –when the result has been found in a Functional Analysis framework–, or the extension result concerns other properties assumed also for the cone –for instance, an order relation– (see for example [8, 1, 2, 9, 5]). In other cases, the dominated extension theorems are given for extended functionals, i.e., functionals for which the value ∞ can be attained (see [6] and [2]); however, these results can also be applied in order to obtain extension and separation theorems for finite valued functionals (see Remark (3.8)). Here our

2000 *Mathematics Subject Classification*: 22A30, 20K99, 46A20, 46A22.

Keywords and phrases: cone, normed cone, duality.

The authors acknowledge the support of the Spanish Ministry of Science and Technology and FEDER under grant BFM2003-02302.

aim is to show that dominated extension theorems for functionals on noncancellative cones can be formulated in terms of cancellative cones. In fact, the extension of a functional defined on a noncancellative cone is given by the extension of the associated functional defined on the corresponding cancellative quotient cone. Thus, the concept of a cancellative quotient is the main key of the technique that we propose here. It is based on the definition of an equivalence relation that we call the Q-relation, which is related to the canonical construction of the locally convex cones in the sense of W. Roth (see [6], [7]).

Let $p\colon C \to \mathbb{R}^+$ be an homogeneous function, i.e., $p(rx) = rp(x)$ for every $r \in \mathbb{R}^+$ and $x \in C$. We say that a functional f is *dominated* by p if $f(x) \leq p(x)$ for all x in C. An homogeneous function p on a cone C is a seminorm if $p(x+y) \leq p(x) + p(y)$ for every $x, y \in C$ and we say that the pair (C, p) is a seminormed cone. The function p induces a topology on C by means of the following family of sets. If $x \in C$, we define for every $\epsilon > 0$ the ball

$$B_p(x,\epsilon) = \{y \in C\colon \text{there is } z \in C \text{ such that } z + x = y,\, p(z) < \epsilon\}.$$

Thus the family $\{B_p(x,\epsilon) : \epsilon > 0\}$ is a basis of neighbourhoods of x. A functional that is continuous with respect to this topology will be called a continuous functional. Although it is not used later on, recall that a seminorm is said to be a quasi-norm if it defines a quasi-metric d_p on C by the formula $d_p(x,y) = \inf\{p(z)\colon x + z = y\}$ if there exists such an element z, and $d_p(x,y) = \infty$ otherwise (see [4]). In this case we say that (C,p) is a quasi-normed cone, and the topology that is considered for this space is the one induced by d_p, which is the same as the one associated to p as a seminorm.

We use standard notation. Let (C,p) and (D,q) be seminormed cones. We say that a function $\phi\colon (C,p) \to (D,q)$ is an isometric isomorphism if it is a linear bijection and $q(\phi(x)) = p(x)$ for every $x \in C$ (see [4]). We say that a cone C is 1+-dimensional if it can be written as $\{rx : r \in \mathbb{R}^+\}$ for an element $x \in C$. A pair of 1+-dimensional subcones C_0 and C_1 of a cone C are linearly independent if there do not exist $\lambda \in \mathbb{R}^+$, $x \in C_0 - \{0\}$ and $y \in C_1 - \{0\}$ such that $\lambda x = y$ or $\lambda x + y = 0$. It is assumed that a cancellative cone C can be identified with an algebraically closed subset of a linear space (see [2]), i.e., a subset of a linear space that satisfies that for $x, y \in C$ and $r \in \mathbb{R}^+$, $x + y \in C$ and $rx \in C$. Thus, for $B \subset C$ we can consider the linear subspace linspan $\{B\}$ of X defined as the (non necessarily positive) linear combinations of elements of B.

2. The dual cone and the cancellative quotient of a cone

In this section we define the cancellative quotient of a cone C and establish the relation between this cone and the algebraic dual of C. Consider a seminorm p defined on C. We define a function $p^*\colon C' \to R^+ \cup \{\infty\}$ by

$$p^*(f) = \sup_{p(x)\leq 1} f(x), \qquad x \in C.$$

Obviously p^* is well-defined, since $f(0) = 0$ for every $f \in C'$ and then $p^*(f) \geq 0$.

Remark (2.1). If C is a cone, the algebraic dual C' endowed with the pointwise sum and product by nonnegative real numbers is also a cone. Moreover, if $r \in R^+$ and $f, g \in C'$, it is clear that

$$p^*(rf) = rp^*(f),$$

and

$$\begin{aligned} p^*(f+g) &= \sup_{p(x)\leq 1} (f(x)+g(x)) \\ &\leq \sup_{p(x)\leq 1} f(x) + \sup_{p(x)\leq 1} g(x) \leq p^*(f)+p^*(g). \end{aligned}$$

Definition (2.2). We define the dual cone of a seminormed cone as the pair (C^*, p^*), where

$$C^* = \{f \in C' \colon p^*(f) < \infty\}.$$

It is clear by Remark (2.1) that (C^*, p^*) is a seminormed cone.

Note that if the functional $f \colon C \to R$ is continuous (where C is endowed with the topology induced by the seminorm p), then $p^*(f) < \infty$.

Let C be a cone and define the following relation Q on $C \times C$. A pair of elements $x, y \in C$ are *Q-related* if there is an element $v \in C$ such that $x+v = y+v$. In this case we write xQy.

This relation is obviously reflexive and symmetric. To show that it is transitive, consider the elements $x, y, z \in C$ and suppose that xQy and yQz. Then there are $v, w \in C$ such that $x+v = y+v$ and $y+w = z+w$. Since the sum is commutative and associative, this implies that $x+v+w = y+v+w = z+v+w$. Thus xQz, and hence Q is an equivalence relation.

If $x \in C$, let us denote by $[x]$ the equivalence class of x with respect to Q, and

$$C/Q = \{[x] \colon x \in C\}.$$

We will denote by P_Q the corresponding quotient map $P_Q \colon C \to C/Q$ given by $P_Q(x) = [x]$, $x \in C$.

If (C, p) is a seminormed cone, let us also define

$$\overline{p}([x]) = \inf\{p(x') \colon x' \in [x]\}, \qquad \text{for all } x \in C.$$

LEMMA (2.3). *The set C/Q endowed with the usual sum and product of equivalence classes $[x]+[y] := [x+y]$ and $r[x] := [rx]$, where $x, y \in C$ and $r \in R^+$, is a cancellative cone. Moreover, if (C, p) is a seminormed cone, $\overline{p}$ is a seminorm on C/Q.*

Proof. It is immediate to show that both the sum of equivalence classes and the natural product are well-defined.

To show that the cone is cancellative, take two elements $[x], [y] \in C/Q$ and suppose that there is an element $[v] \in C/Q$ such that $[x]+[v] = [y]+[v]$. Thus, $[x+v] = [y+v]$, and we can find an element $w \in C$ such that $x+v+w = y+v+w$. Therefore $[x] = [y]$.

Consider the seminorm p on C, and take two elements $[x], [y] \in C/Q$. Let $\epsilon > 0$. Then there are $x' \in [x]$ and $y' \in [y]$ that satisfy $p(x') < \overline{p}([x]) + \epsilon$ and $p(y') < \overline{p}([y]) + \epsilon$. Since $x'+y' \in [x+y]$, we obtain $\overline{p}([x+y]) \leq p(x'+y') < \overline{p}([x]) + \overline{p}([y]) + 2\epsilon$. The same kind of argument proves that it is homogeneous. $\square$

In the following we show that the structure of the dual cone (C^*, p^*) depends only on the cancellative quotient of the seminormed cone (C, p).

LEMMA (2.4). *The quotient map* $P_Q\colon (C,p) \to (C/Q, \overline{p})$ *is a continuous linear surjection.*

Proof. The map P_Q is obviously linear and surjective. Let $[x] \in C/Q$ and $\epsilon > 0$. Then if $y \in B_p(x,\epsilon)$ there is $z \in C$ such that $y = x + z$ and $p(z) < \epsilon$. Therefore, $[y] = [x+z] = [x] + [z]$ and $\overline{p}([z]) \leq p(z) < \epsilon$, which implies $P_Q(B_p(x,\epsilon)) \subset B_{\overline{p}}([x],\epsilon)$. Thus, P_Q is also continuous. □

Definition (2.5). Let C be a cone. We say that C' separates points if for every pair of different elements $x, y \in C$ there is a functional $f \in C'$ such that $f(x) \neq f(y)$.

PROPOSITION (2.6). *Let C be a cone. Then the following are equivalent.*
1) C is cancellative.
2) The quotient map P_Q between C and C/Q is a bijection.
3) The algebraic dual C' separates points.

Consequently, if (C,p) is a cancellative seminormed cone, then P_Q defines an isometric isomorphism between (C,p) and $(C/Q,\overline{p})$.

Proof. To prove 1) $\Rightarrow$ 2) we just need to show that P_Q is injective. Consider $x, y \in C$ and suppose that $P_Q(x) = [x] = [y] = P_Q(y)$. Then there is $v \in C$ such that $x + v = y + v$. Since C is cancellative, we directly obtain $x = y$. To prove 2) $\Rightarrow$ 1), suppose that C is not cancellative. Then there are $x, y, v \in C$ such that $x + v = y + v$ but $x \neq y$. Then $P_Q(x) = P_Q(y)$ and $x \neq y$, and so P_Q is not a bijection. To see 1) $\Rightarrow$ 3), it is enough to use the following construction. Let $x, y \in C$, $x \neq y$. We can suppose without loss of generality that they are linearly independent, since otherwise we can construct a linear function that separates x and y using an obvious modification of the following argument. Since C is cancellative, we can consider it as an algebraically closed subset of a linear space X. In each linear space Y, if S is a linearly independent subset of vectors, we can find another subset D such that $S \cup D$ is a basis of Y (see Th. 2.4, Ch. IV in [3]). Thus, consider the subset $S = \{x, y\}$ and find a subset D that completes the basis $B = S \cup D$. Let us define the linear function $f\colon X \to R$ by $f(x) = 1$, $f(y) = 0$ and $f(z) = 0$ for every $z \in D$. It is clear that the restriction of f to C gives a map in C' and obviously $f(x) \neq f(y)$.

Finally, let us show 3) $\Rightarrow$ 1). Suppose that C is not cancellative. Then there are elements $x, y, v \in C$ such that $x + v = y + v$ and $x \neq y$. Then if $f \in C'$,

$$f(x) + f(v) = f(x+v) = f(y+v) = f(y) + f(v),$$

and so $f(x) = f(y)$. Therefore there is no function $f \in C'$ satisfying $f(x) \neq f(y)$.

The fact that P_Q defines also an isometry is a direct consequence of Lemma (2.4) and the definition of $\overline{p}$. This finishes the proof. □

Remark (2.7). Note that for every $f \in C'$, $f(x) = f(y)$ if $x \in [y]$.

THEOREM (2.8). *Let C be a cone. Then the formula $\Psi(f)([x]) := f(x)$, $f \in C'$, $x \in C$, defines a function*

$$\Psi\colon C' \to (C/Q)'$$

that is an isomorphism. Moreover, if (C,p) is a seminormed cone, then Ψ defines an isometric isomorphism between (C^*, p^*) and $((C/Q)^*, \overline{p}^*)$.

Proof. We first note that Ψ is well-defined by Remark (2.7). Next, we show that $\Psi(f) \in (C/Q)'$. If $r, s \in R^+$ and $[x], [y] \in C/Q$,

$$\begin{aligned}\Psi(f)(r[x] + s[y]) &= \Psi(f)([rx + sy]) = f(rx + sy) \\ &= rf(x) + sf(y) = r\Psi(f)([x]) + s\Psi(f)([y]).\end{aligned}$$

On the other hand, it is routine to prove that Ψ is an isomorphism.

Now suppose that (C, p) is a seminormed cone. Then $(C/Q, \overline{p})$ is also a seminormed cone, as a consequence of Lemma (2.3). Let $f \in C^*$. The inequality

$$p^*(f) \leq \sup\{\Psi(f)([x]) \colon \overline{p}([x]) \leq 1\} = \overline{p}^*(\Psi(f))$$

is a direct consequence of the definition of $\overline{p}$. For the other inequality, let $\epsilon > 0$. Then there is an element $[x_0] \in C/Q$ such that $\overline{p}([x_0]) \leq 1$ and

$$(1 - \epsilon)\overline{p}^*(\Psi(f)) = (1 - \epsilon) \sup\{\Psi(f)([x]) \colon \overline{p}([x]) \leq 1\} \leq \Psi(f)([x_0]).$$

Now take an element $z_0 \in C$ such that $z_0 \in [x_0]$ and

$$p(z_0) \leq \overline{p}([x_0]) + \epsilon \leq 1 + \epsilon.$$

Then

$$p^*(f) = \sup\{f(x) \colon p(x) \leq 1\} \geq f(\frac{z_0}{1 + \epsilon}) = \Psi(f)(\frac{[x_0]}{1 + \epsilon}) \geq \frac{1 - \epsilon}{1 + \epsilon}\overline{p}^*(\Psi(f)).$$

Since these inequalities hold for every $\epsilon > 0$, we obtain the result. $\square$

The cancellativity of the additive operation can also be formulated in terms of the extension properties of functionals defined on subcones. Let us illustrate the situation with an example. Consider the cone C_0 defined by the elements of R^2

$$C_0 = \{(x, 0) \colon x \in \mathbb{R}^+\} \cup \{(0, y) \colon y \in \mathbb{R}^+\},$$

with the usual product by nonnegative scalars of $\mathbb{R}^2$ and the sum

$$\begin{aligned}(x_1, 0) + (x_2, 0) &:= (x_1 + x_2, 0), && x_1, x_2 \in \mathbb{R}^+, \\ (x, 0) + (0, y) &:= (x, 0), && x \neq 0, \quad x, y \in \mathbb{R}^+, \\ (0, y_1) + (0, y_2) &:= (0, y_1 + y_2)), && y_1, y_2 \in \mathbb{R}^+,\end{aligned}$$

and assuming that $(0, 0)$ is the neutral element. A direct calculation shows that for $x \neq 0$, $[(x, 0)] = \{(x, 0)\}$, but for every $y \in \mathbb{R}^+$,

$$[(0, y)] = \{(0, y') \colon y' \in \mathbb{R}^+\} = [(0, 0)].$$

Thus, $C_0/Q = \{[(x, 0)] \colon x \in \mathbb{R}^+\} = \mathbb{R}^+$. Therefore, $(C_0/Q)'$ is isomorphic to C_0' by means of the linear map $\Psi_0 \colon C_0' \to (C_0/Q)'$ given by

$$\Psi_0(f)([(x, 0)]) = f((x, 0)).$$

Moreover, since $[(0, y)] = [(0, 0)]$, $f((0, y)) = f((0, 0)) = 0$ for every $y \in \mathbb{R}^+$ and every functional $f \in C'$. As a consequence, if we consider the subcone $C_1 = \{(0, y) \colon y \in \mathbb{R}^+\}$ of C and a nontrivial linear map $g \colon C_1 \to \mathbb{R}$, it is not possible to find an extension $\overline{g}$ of g to the whole C_0, since $\overline{g}((0, y)) = g((0, y)) = 0$ for every $y \in \mathbb{R}^+$. This motivates the following result.

PROPOSITION (2.9). *Let C be a cone. Then the following are equivalent.*

1) *The neutral element* 0 *is the unique element of* $[0]$.

2) *For every* $1+$*-dimensional subcone* C_0 *of* C *and each linear function* $f_0\colon C_0 \to \mathbb{R}$ *there exists an extension to a functional* $f \in C'$.

Proof. Suppose that 0 is the unique element in $[0]$, and consider a $1+$-dimensional subcone $C_0 := \{rx\colon r \in \mathbb{R}^+\}$, $x \neq 0$, and a linear function $f_0\colon C_0 \to \mathbb{R}$. The cancellative quotient C/Q can be identified with a subcone of a linear space V, and then we can find a basis B containing $[x]$, since $x \neq 0$ and then $[x] \neq [0]$. We can define a linear function $\overline{f}\colon V \to \mathbb{R}$ by $\overline{f}([x]) = f_0(x)$ and $\overline{f}(v) = 0$ for every $v \in B - \{[x]\}$. Since every element $[z] \in C/Q$ can be identified with an element of V, the formula

$$f(z) = \overline{f}([z]), \qquad z \in C,$$

defines a (linear) function $f\colon C \to \mathbb{R}$. Obviously, f is an extension of f_0.

Conversely, suppose that there is an element $x \neq 0$ such that $x \in [0]$. Then there is an element $z \in C$ such that $x + z = z$. Consider the subcone $C_0 := \{rx\colon r \in \mathbb{R}^+\}$ and the function $f_0\colon C_0 \to \mathbb{R}$ given by $f_0(rx) = r$. Then there is no extension of f_0 to C, since every linear function $f\colon C \to R$ satisfies $f(x) = f(0) = 0$, as a consequence of Remark (2.7). □

However, there are cones that are not cancellative but satisfy the condition $[0] = \{0\}$. The following modification of the example above shows this. Let us define the cone C_1 by the same elements that the cone C_0, also with the usual product by non negative scalars and the sum

$$\begin{aligned}
(x_1, 0) + (x_2, 0) &:= (x_1 + x_2, 0), && x_1, x_2 \in \mathbb{R}^+,\\
(x, 0) + (0, y) &:= (x + y, 0), && x > 0, \quad x, y \in \mathbb{R}^+,\\
(0, y_1) + (0, y_2) &:= (0, y_1 + y_2)), && y_1, y_2 \in \mathbb{R}^+.
\end{aligned}$$

In this case, $[(0, 0)] = (0, 0)$, since $(0, 0)$ is the only element $z \in C_1$ that satisfies $v + z = v$ for some $v \in C_1$. However, note that for every $y > 0$, $[(0, y)] = [(y, 0)]$, and then C_1 is not cancellative.

PROPOSITION (2.10). *Let C be a cone. Then the following are equivalent.*

1) C is cancellative.

2) For every pair of linearly independent $1+$*-dimensional subcones* C_0 *and* C_1 *of* C *and each pair of linear functions* $f_0\colon C_0 \to \mathbb{R}$ *and* $f_1\colon C_1 \to \mathbb{R}$ *there exists an extension to a function* $f \in C'$.

Proof. If C is cancellative, the proof of 1) $\Rightarrow$ 3) of Proposition (2.6) gives the key to produce a simultaneous extension f of f_0 and f_1 to C. If $0 \neq x \in C_0$ and $0 \neq y \in C_1$, since C_0 and C_1 are linearly independent it is possible to construct a basis B of a vector space V including C that contains x and y. Then it is enough to define an extension $\overline{f}\colon V \to R$ as $\overline{f}(x) := f_1(x)$, $\overline{f}(y) := f_2(y)$ and $\overline{f}(z) := 0$ for every $z \in B - \{x, y\}$, and by linearity for the rest of the elements of V. The desired extension to C is the restriction $f := \overline{f}|_C$.

To prove that 2) implies 1), suppose that C is not cancellative. Then there are elements $x, y, v \in C$ such that $x + v = y + v$ but $x \neq y$. We can suppose without loss of generality that $x \neq 0$. Consider the subcones $C_0 := \{rx\colon r \in R^+\}$

and $C_1 = \{ry : r \in R^+\}$, and the functionals $f_0 : C_0 \to R^+$ and $f_1 : C_1 \to R^+$ given by $f_0(rx) = r$ and $f_1(ry) = 0$ for every $r \in R^+$. It is easy to see that there is no simultaneous extension of f_0 and f_1, since as a consequence of Remark (2.7) such an extension f must satisfy $1 = f_0(x) = f(x) = f(y) = 0$, a contradiction. □

Obviously, if C is cancellative then 0 is the unique element of [0], and then the statements of Proposition (2.10) imply the existence of extensions of functionals defined on 1+-dimensional subcones, as a consequence of Propositions (2.9) and (2.10).

3. Applications: Hahn-Banach type theorems for noncancellative cones

In the previous section we have found the relations between the algebraic dual of a cone and the cancellativity of the cone. Moreover, we have characterized this property in terms of the existence of (algebraic) extensions of functionals defined on 1+-dimensional subcones. In this section we present several results for dominated extensions of functionals defined on noncancellative cones, which can be applied in the context of seminormed cones. The technique that we use is based on the construction of the cancellative quotient associated to a cone C.

From now on we will deal with extensions of functionals defined on a subcone C_0 of a cone C. We will always consider the equivalence relation Q defined by the cone C, i.e., the equivalence class $[x]$ of an element $x \in C_0$ will be defined taking elements v of the whole cone C. Therefore if $x, y \in C_0$, xQy if and only if there is an element $v \in C$ such that $x + v = y + v$.

LEMMA (3.1). *Consider a subcone C_0 of a cone C and the set*

$$C_0/Q := \{[x] \in C/Q : x \in C_0\}.$$

Then C_0/Q is a subcone of C/Q.

Proof. If $x, y \in C_0$, then $x+y \in C_0$ and $[x]+[y] = P_Q(x)+P_Q(y) = P_Q(x+y) = [x + y]$. This implies that $[x + y] \in C_0/Q$. A similar argument gives that for every $r \in R^+$ and $x \in C_0$, $[rx] \in C_0/Q$. Therefore C_0/Q is a subcone of C. □

Let C_0 be a subcone of a cone C, and consider a functional $f_0 : C_0 \to R$. A direct consequence of Remark (2.7) is that if there is a linear extension of f_0 to C and $x, y \in C_0$, then $f_0(x) = f_0(y)$ if $y \in [x]$. This motivates Definition (3.2) and Lemma (3.3) below.

Definition (3.2). Let C_0 be a subcone of a cone C. A function $f_0 : C_0 \to R$ is called Q-compatible if for each pair $x, y \in C_0$ such that xQy, we have $f_0(x) = f_0(y)$.

The following result is essentially known, so we omit its (easy) proof.

LEMMA (3.3). *Consider a subcone C_0 of a cone C and a Q-compatible linear function $f_0 : C_0 \to R$. Then f_0 defines a linear function $f_0/Q : C_0/Q \to R$ by means of the formula $f_0/Q([x]) := f_0(x)$, $[x] \in C_0/Q$.*

Conversely, if $f/Q\colon C_0/Q \to R$ *is a linear function, then the formula* $f(x) := f/Q([x])$, $x \in C_0$, *provides a functional* $f\colon C_0 \to R$ *that is* Q*-compatible.*

It follows from the above lemma that if $f \in C'$ then f/Q is exactly $\Psi(f)$. This fact will be used in the proof of our next result.

PROPOSITION (3.4). *Consider a subcone* C_0 *of a cone* C. *Let* $f_0\colon C_0 \to R$ *be a linear function. Then the following hold.*

1) *If there exists a linear extension* $f\colon C \to R$ *of* f_0, *then there is a linear extension* $f_Q\colon C/Q \to R$ *of the linear function* $f_0/Q\colon C_0/Q \to R$.

2) *If there exists a linear extension* $f_Q\colon C/Q \to R$ *of the linear function* $f_0/Q\colon C_0/Q \to R$, *then there is a linear extension* $f\colon C \to R$ *of* f_0.

In this case the corresponding extensions can be defined to satisfy $\Psi(f) = f/Q = f_Q$. *Moreover, if* (C, p) *is a seminormed cone, then* $f \in C^*$ *if and only if* $f_Q \in (C/Q)^*$ *and* $p^*(f) = \overline{p}^*(f_Q)$.

Proof. 1) If there is a linear extension of $f_0\colon C_0 \to R$, then f_0 is Q-compatible by Remark (2.7). Then f_0/Q is a well-defined linear function by Lemma (3.3). Moreover, if f extends f_0, we can define $f_Q = f/Q = \Psi(f)$, and then $f_Q([x]) = f(x) = f_0(x) = f_0/Q([x])$ for every $[x] \in C_0/Q$. Thus f_Q provides an extension of f_0/Q.

2) Suppose that we have an extension f_Q of a linear function f_0/Q defined by a functional $f_0\colon C_0 \to R$. Note that we assume that f_0 is Q-compatible, since f_0/Q is defined. Moreover, Lemma (3.3) gives $f_0(x) = f_0/Q([x])$ for every $x \in C_0$. Let us define the function $f\colon C \to R$ by $f(x) := f_Q([x])$ for every $x \in C$. It is obviously a well-defined linear function that extends f_0. Note that $f_Q(f) = \Psi(f)$. The statement for the case of seminormed cones is a direct consequence of Theorem (2.8). □

Theorem (2.8) and Proposition (3.4) establish an adequate framework for the automatic generalization of extension theorems to the context of noncancellative cones. In the following we provide an application of these results. In particular, we generalize to the case of noncancellative cones a Hahn-Banach type theorem that is known for cancellative cones and has been recently published in [5].

Definition (3.5). ([5]) We say that a cancellative cone C is well-generated by an algebraic basis $V = \{v_i\colon i \in I\}$ of linspan$\{C\}$ if for every subset J of I and each $i_0 \in I - J$, we can find for every element

$$y \in \operatorname{linspan}\{\{v_i\colon i \in J\},\ v_{i_0}\} \cap C,$$

a representation $y = x + \lambda v_{i_0}$, where $x \in C \cap \operatorname{linspan}\{v_i\colon i \in J\}$ and $\lambda \in R$.

Definition (3.6). ([5]) Let C be a cancellative cone and let C_0 be a subcone of C. We say that C_0 is V-compatible if for the algebraic basis $V = \{v_i\colon i \in I\}$ of linspan $\{C\}$ there is a subset J of I such that $V_0 = \{v_i\colon i \in J\}$ is an algebraic basis of linspan $\{C_0\}$.

COROLLARY (3.7). *Let* (C, p) *be a seminormed cone such that* C/Q *is well-generated by an algebraic basis* V *and let* C_0 *be a subcone of* C *such that* C_0/Q

is V-compatible. Suppose that $f_0 \in C_0'$ is Q-compatible, and that there is a constant $K > 0$ satisfying

$$f_0(x+z) - f_0(x) \leq K\overline{p}([z])$$

for each $x \in C_0$ and $z \in C$ such that $x + z \in C_0$. Then there exists an extension $f \in C'$ such that

$$f(z) \leq Kp(z), \qquad z \in C.$$

Moreover, $f_0 \in C_0^*$, $f \in C^*$ and $p^*(f_0) = p^*(f)$.

Proof. Since $f_0 \colon C_0 \to R$ is a Q-compatible linear map, Lemma (3.3) makes clear that the functions $f_0/Q([x]) := f_0(x)$ and $\phi([x]) := K\overline{p}([x])$, $[x] \in C_0/Q$, satisfies the inequality

$$f_0/Q([x]+[z]) - f_0/Q([x]) \leq \phi([z])$$

for each $[x] \in C_0/Q$ and $[z] \in C/Q$ such that $[x+z] \in C_0/Q$. Therefore f_0/Q satisfies the conditions of Theorem 8 of [5]. An application of this result gives an extension $\overline{f}_Q$ to the whole cone C/Q such that

$$\overline{f}_Q([z]) \leq \phi([z]), \qquad [z] \in C/Q,$$

and then the function $f \colon C \to R$ given by

$$f(x) := \overline{f}_Q([x]) \qquad x \in C,$$

satisfies

$$f(x) \leq \phi([x]) \leq Kp(x), \qquad x \in C.$$

It is clear that f extends f_0. Moreover, $q^*(f) = q^*(f_0)$ as a consequence of Theorem (2.8) and Theorem 8 of [5]. This finishes the proof. □

Obviously, the results of this paper can be directly applied to obtain dominated extensions for functionals on normed cones, since each normed cone is in particular a seminormed cone.

Remark (3.8). One of the referees has pointed out to the authors the following facts:

(1) The seminormed case discussed here may be modelled as a locally convex cone in the sense of [6], [7].

(2) Theorem 3.1 of [7] provides a separation result related to Proposition (2.6) above.

(3) From Theorem 4.1 of [6] one can deduce the following result on extension of functionals: Let D be a subcone of C. A real-valued functional f on D can be extended to a real-valued functional on D if and only if there is a convex and absorbing subset A such that $0 \in A$ and such that for all $x, y \in X$ we have $|f(x) - f(y)| \leq 1$ whenever $x + u = y + v$ for some $u, v \in A$.

Acknowledgement

The authors are very grateful to the referees for their valuable suggestions, in particular, for calling our attention to reference [7].

Received June 16, 2005

Final version received February 06, 2006

S. Romaguera,
E. A. Sánchez-Pérez
Escuela de Caminos
Instituto de Matemática Pura y Aplicada
Universidad Politécnica de Valencia
46071 Valencia
Spain
sromague@mat.upv.es
easancpe@mat.upv.es

O. Valero
Departamento de Ciencias Matématicas e Informática
Universidad de las Islas Baleares
Edificio Anselm Turmeda
07122 Palma, Baleares
Spain
o.valero@uib.es

References

[1] C. Alegre, J. Ferrer, V. Gregori, *On the Hahn-Banach theorem in certain linear quasi-uniform structures*, Acta Math. Hungar. **82** (1999), 315–320.

[2] B. Fuchssteiner, W. Lusky, Convex Cones, North Holland Math. Stud. **56**, Amsterdam. 1981.

[3] T. W. Hungerford, Algebra, Springer-Verlag, Berlin. 1974.

[4] S. Oltra, O. Valero, *Isometries on quasi-normed cones and bicompletion*, New Zealand J. Math. **33** (2004), 83–90.

[5] S. Romaguera, E. A. Sánchez Pérez, O. Valero, *Dominated extensions of functionals and V-convex functions on cancellative cones*, Bull. Austral. Math. Soc. **67** (2003), 87–94.

[6] W. Roth, *Hahn-Banach type theorems for locally convex cones*, J. Aust. Math. Soc. Ser. A, **68** (2000), 104–125.

[7] W. Roth, *Separation properties for locally convex cones*, J. Convex Anal. **9** (2002), 301–307.

[8] W. Rudin, Functional Analysis, McGraw-Hill. New York. 1973.

[9] R. Tix, *Some results on Hahn-Banach type theorems for continuous d-cones*, Theoret. Comput. Sci. **264** (2001), 205–218.

Bol. Soc. Mat. Mexicana (3) Vol. 12, 2006

INEXISTENCE OF INVARIANT MEASURES FOR GENERIC RATIONAL DIFFERENTIAL EQUATIONS IN THE COMPLEX DOMAIN

MARCO BRUNELLA

ABSTRACT. We prove that a generic holomorphic foliation by curves on the complex projective space has no invariant measures. The main ingredient is an index formula for invariant measures.

A *foliation by curves* of degree d (≥ 0) on the complex projective space $\mathbb{C}P^n$ is generated by a nontrivial holomorphic section (a rational vector field)

$$s \in H^0(\mathbb{C}P^n, \Theta_{\mathbb{C}P^n} \otimes \mathcal{O}(d-1))$$

where $\Theta_{\mathbb{C}P^n}$ is the tangent sheaf and $\mathcal{O}(d)$ is the line bundle of degree d [G-O]. One usually assumes that the zero set of s, i.e., the singular set of the foliation, has codimension at least 2, so that s is uniquely determined up to a multiplicative constant. In this way, the space $\mathrm{Fol}\,(d, n)$ of degree d foliations on $\mathbb{C}P^n$ can be identified with a Zariski–open subset of the projective space $\mathrm{Proj}\,(H^0(\mathbb{C}P^n, \Theta_{\mathbb{C}P^n} \otimes \mathcal{O}(d-1)))$ (whose dimension is $\binom{n+d-1}{d}(n+d+1)-1$), and as such it has a natural topology.

Given a foliation $\mathcal{F}$, with singular set $\mathrm{Sing}\,(\mathcal{F})$, there is a classical notion of *invariant measure* for the nonsingular foliation $\mathcal{F}|_{\mathbb{C}P^n \setminus \mathrm{Sing}\,(\mathcal{F})}$, see for instance [Sul]. We shall prove that, generically, invariant measures do not exist, when $d \geq 2$:

THEOREM. *Given $n \geq 2$ and $d \geq 2$, there exists an open and dense subset $\mathcal{U} \subset \mathrm{Fol}\,(d, n)$ such that any $\mathcal{F} \in \mathcal{U}$ has no invariant measure.*

This result must be compared with (and is based on) a result by Lins Neto and Soares [L-S], asserting that there exists an open and dense $\mathcal{U} \subset \mathrm{Fol}\,(d, n)$ such that every $\mathcal{F} \in \mathcal{U}$ has the following two properties:

i) all the singularities of $\mathcal{F}$ are hyperbolic (see below, Section 2, for the definition);

ii) no algebraic curve is invariant by $\mathcal{F}$.

In fact, our $\mathcal{U}$ is the same as in [L-S], and the proof of the Theorem above consists in showing that every foliation on $\mathbb{C}P^n$ which satisfies i) and ii) has no invariant measure.

When $n = 2$, this is extremely simple. Indeed, by Bott's vanishing principle (see [B-B] and Section 1 below) and positivity properties of $\mathbb{C}P^n$, the support of an invariant measure is forced to meet one or more singularities of the

2000 *Mathematics Subject Classification*: 32Uxx, 34Mxx, 37Fxx.

Keywords and phrases: holomorphic foliations, invariant measures.

foliation [CLS]. If $n = 2$, on a neighbourhood of a hyperbolic singularity there are very few invariant measures: only the ones concentrated on the two local separatrices. This forces the invariant measure to be, globally, concentrated on an invariant algebraic curve, which however does not exist by assumption.

When $n \geq 3$, however, the proof is slightly more elaborated, because a hyperbolic singularity can be in the *Siegel domain* [CKP], and in that case there are a lot of invariant measures on a neighbourhood of the singularity. Thus, we need to replace Bott's vanishing principle by a more quantitative formula, given in Section 1 below, and then we prove that those invariant measures around the hyperbolic singularities give no contribution. This formula is a higher dimensional generalisation of a result of [Bru].

Let us observe the following consequence of our Theorem:

COROLLARY. *For every $\mathcal{F} \in \mathcal{U}$ all the leaves are hyperbolic. More precisely, there exists no nonconstant $f\colon \mathbb{C} \to \mathbb{C}P^n$ tangent to $\mathcal{F}$ (and possibly passing through* $\mathrm{Sing}\,(\mathcal{F})$*).*

Indeed, by Ahlfors' lemma [Bru], such an f could be used to construct an invariant measure for $\mathcal{F}$. When $n = 2$, this Corollary was already observed in [C-G].

1. Baum–Bott formula for invariant measures

Let $\mathcal{F}$ be a holomorphic foliation by curves on a complex manifold X, of dimension n, and assume that $\mathrm{Sing}\,(\mathcal{F})$ is a discrete subset. Let $X^0 = X \setminus \mathrm{Sing}\,(\mathcal{F})$.

Recall, after [Sul], that there exists a natural bijective correspondence between invariant measures for $\mathcal{F}|_{X^0}$ and invariant closed positive currents of bidimension $(1, 1)$ (we refer e.g. to [Har] for the basic theory of positive currents). Here, a closed positive current T on X^0 is said to be *invariant* by $\mathcal{F}$ if $T(\Theta) = 0$ for every 2-form Θ which vanishes on the leaves of $\mathcal{F}$. In other words, the value $T(\Theta)$ depends only on the restriction of Θ to the leaves of $\mathcal{F}$. Locally, on X^0, we may choose coordinates $(z_1, ..., z_n)$ so that $\mathcal{F}$ is generated by $\frac{\partial}{\partial z_1}$. Setting

$$\alpha_j = dz_1 \wedge \cdots \wedge dz_{j-1} \wedge dz_{j+1} \wedge \cdots \wedge dz_n\,,$$

the positive current T can be locally written as

$$T = \sum_{j,k=1}^{n} f_{jk}\sqrt{-1}\alpha_j \wedge \overline{\alpha}_k$$

for suitable complex valued measures f_{jk}. The $\mathcal{F}$-invariance of T means that $T \wedge dz_j \equiv 0 \equiv T \wedge d\overline{z}_j$ for every $j \neq 1$, whence $f_{jk} \equiv 0$ for $(j, k) \neq (1, 1)$ and so

$$T = f_{11}\sqrt{-1}\alpha_1 \wedge \overline{\alpha}_1\,.$$

The closedness of T means that f_{11} does not depend on z_1 (more precisely, the distributional derivatives of f_{11} along z_1 and $\overline{z}_1$ are zero). Thus T also does not depend on z_1 and can be projected to a positive measure on the local transversal $\{z_1 = 0\}$. This procedure, repeated on each foliated chart on X^0, associates to T a measure transverse to $\mathcal{F}$ and invariant by the holonomy.

Conversely, given such a measure we may construct on X^0 a closed positive current T, invariant by $\mathcal{F}$, by firstly integrating a 2-form along the leaves and secondly by integrating the result with respect to the transverse measure. The closedness of T follows from the holonomy invariance of the transverse measure. See [Sul] for more details.

The closed positive current T is at the beginning defined only on X^0, but if $X \setminus X^0 = \mathrm{Sing}\,(\mathcal{F})$ is discrete then T can be extended, in a unique way, to a closed positive current on X. In the following, we shall forget invariant measures and we will concentrate on invariant closed positive currents on X.

Let $p \in \mathrm{Sing}\,(\mathcal{F})$. We now introduce an index

$$\mathrm{Res}\,(\mathcal{F}, T, p) \in \mathbb{C}$$

which, roughly speaking, represents the twisting of the normal bundle of $\mathcal{F}$ along T and around p.

Choose, on a neighbourhood U of p, a holomorphic vector field v generating $\mathcal{F}$ and vanishing only at p, and choose also a holomorphic n-form Ω on U without zeroes. Set $\omega = i_v\Omega$: it is a holomorphic $(n-1)$-form which, outside p, spans the determinant of the conormal bundle of $\mathcal{F}$, $\det N^*_{\mathcal{F}}$.

On $U^* = U \setminus \{p\}$, take a smooth $(1,0)$-form β such that

$$d\omega = \beta \wedge \omega\,.$$

Such a β always exists: for instance, if $v = \sum_{j=1}^n F_j \frac{\partial}{\partial z_j}$ and $\Omega = dz_1 \wedge \cdots \wedge dz_n$, then we may take

$$\beta = \frac{\mathrm{div}\,(v)}{\sum_{j=1}^n |F_j|^2} \sum_{j=1}^n \overline{F}_j dz_j$$

where $\mathrm{div}\,(v) = \sum_{j=1}^n \frac{\partial F_j}{\partial z_j}$. The $(1,0)$-form β is not uniquely determined, but its restriction to the leaves is (once ω is fixed). Note that $d\omega = \beta \wedge \omega$ can be rewritten as $i_v\beta = \mathrm{div}_{\,\Omega}(v)$, where $\mathrm{div}_{\,\Omega}(v)$ is the divergence of v with respect to Ω. The restriction $\beta|_{\mathcal{F}}$ can be seen as a section of the canonical bundle $K_{\mathcal{F}} = T^*_{\mathcal{F}}$, and as such it is holomorphic and therefore holomorphically extensible at p. However, in general we cannot find a holomorphic β (on U^* or U, it is the same), unless $\mathrm{div}_{\,\Omega}(v)$ vanishes at p at a higher order (so that it belongs to the ideal generated by the components of v).

Take now a smooth function φ on U, equal to 0 on a neighbourhood of p and equal to 1 outside a compact subset of U. Hence $d\varphi \wedge \beta$ is a smooth 2-form on U, with compact support. We set:

$$\mathrm{Res}\,(\mathcal{F}, T, p) = \frac{1}{2\pi\sqrt{-1}} T(d\varphi \wedge \beta).$$

An alternative expression is the following, where χ_U denotes the characteristic function of U:

$$\mathrm{Res}\,(\mathcal{F}, T, p) = \frac{1}{2\pi\sqrt{-1}} T(\chi_U d(\varphi\beta)).$$

Indeed, $\chi_U d(\varphi\beta) = d\varphi \wedge \beta + \chi_U \varphi d\beta$, and $T(\chi_U \varphi d\beta) = 0$ because $d\beta|_{\mathcal{F}} \equiv 0$ ($\beta|_{\mathcal{F}}$ is holomorphic, hence closed). Thus, if we assume that β is defined on a neighbourhood of $\overline{U}$, we can also write:

$$\mathrm{Res}\,(\mathcal{F}, T, p) = \frac{1}{2\pi\sqrt{-1}} d(\chi_U T)(\beta)$$

where $d(\chi_U T)$ is the boundary current of $T|_U = \chi_U T$, supported on ∂U (around which $\varphi\beta = \beta$). Therefore, the index $\mathrm{Res}\,(\mathcal{F}, T, p)$ is a sort of residue of the foliated 1-form $\beta|_{\mathcal{F}}$ at p, measured with the current T.

It is easy to check that this definition is well posed, i.e. $\mathrm{Res}\,(\mathcal{F}, T, p)$ depends only on the germs of $\mathcal{F}$ and T at p, and not on the several choices done so far to define it. For instance, if we change v and/or Ω then ω changes by a multiplicative factor $f \in \mathcal{O}^*(U)$, $\tilde{\omega} = f\omega$, and so $\tilde{\beta} = \beta + \frac{df}{f}$ (up to an inessential term which vanishes on the leaves). Then $T(d\varphi \wedge \beta) = T(d\varphi \wedge \tilde{\beta})$ because $T(d\varphi \wedge \frac{df}{f}) = T(d((\varphi - 1)\frac{df}{f})) = dT((\varphi - 1)\frac{df}{f}) = 0$ (or, equivalently, because $d(\chi_U T)(\frac{df}{f}) = 0$).

We can now state the index formula for $c_1(\det N^*_{\mathcal{F}}) \cdot [T]$, where $c_1(\det N^*_{\mathcal{F}}) \in H^2(X, \mathbb{R})$ is the Chern class of $\det N^*_{\mathcal{F}} \in \mathrm{Pic}\,(X)$ and $[T] \in H_2(X, \mathbb{R})$ is the homology class of T. Recall that $\det N^*_{\mathcal{F}}$ is the line bundle on X locally generated by holomorphic $(n-1)$-forms having $\mathcal{F}$ as kernel (as the form ω above).

Proposition (1.1). *Let X be a complex manifold of dimension n, let $\mathcal{F}$ be a foliation by curves on X with discrete singular set, and let T be a closed positive current on X, of bidimension* $(1, 1)$*, with compact support, and invariant by $\mathcal{F}$. Then*

$$c_1(\det N^*_{\mathcal{F}}) \cdot [T] = \sum_{p \in \mathrm{Sing}\,(\mathcal{F}) \cap \mathrm{Supp}\,(T)} \mathrm{Res}\,(\mathcal{F}, T, p)\,.$$

Proof. It is a straightforward generalisation of [Bru, Lemme 4], which is an application of Baum–Bott localisation principle [B-B]. As in [Bru], we can construct a smooth closed 2-form Θ, representing $c_1(\det N^*_{\mathcal{F}})$ in De Rham's sense, such that:

i) if $p \in \mathrm{Sing}\,(\mathcal{F})$ then, on a neighbourhood U_p of p, Θ is equal to $\frac{1}{2\pi\sqrt{-1}} d(\varphi_p \beta_p)$, where φ_p and β_p are as in the definition of $\mathrm{Res}\,(\mathcal{F}, T, p)$;

ii) outside $\cup_p U_p$, $\Theta|_{\mathcal{F}} \equiv 0$.

Hence,

$$c_1(\det N^*_{\mathcal{F}}) \cdot [T] = T(\Theta) = \sum_p T(\chi_{U_p}\Theta) = \sum_p \mathrm{Res}\,(\mathcal{F}, T, p).$$

□

Remark (1.2). We shall not need such a generality, but it could be useful for other purposes to have a similar formula for foliations with nonisolated (compact) singularities. The problem is to define an index $\mathrm{Res}\,(\mathcal{F}, T, \Gamma)$ for every connected component Γ of $\mathrm{Sing}\,(\mathcal{F})$. This is easy when $\det N^*_{\mathcal{F}}$ is trivial around Γ, a little more problematic otherwise.

2. Hyperbolic singularities and their index

A singular point p of $\mathcal{F}$ is said to be *hyperbolic* if around p the foliation is generated by a vector field v whose linear part at p has eigenvalues $\lambda_1, \ldots, \lambda_n$

satisfying

$$\lambda_j \notin \mathbb{R}\cdot \lambda_k \qquad \forall\, j \neq k\,.$$

The structure of hyperbolic singularities has been elucidated by [CKP], in the linear case, and by [Cha], in the nonlinear one. In fact, by [Cha] a foliation on a neighbourhood of a hyperbolic singularity is always topologically linearizable. Let us resume some known results.

There is a basic dichotomy for hyperbolic singularities: $p \in \operatorname{Sing}(\mathcal{F})$, hyperbolic, is in the *Siegel domain* if the convex hull (in $\mathbb{C}$) of the eigenvalues $\{\lambda_j\}_{j=1}^n$ contains the origin 0, in the *Poincaré domain* otherwise. Of course, the first possibility can occur only if $n \geq 3$.

Let

$$v = \sum_{j=1}^n F_j \frac{\partial}{\partial z_j}$$

be a vector field generating $\mathcal{F}$ around $p = 0$. The real analytic variety

$$M = \{z = (z_1, \dots, z_n) \mid \sum_{j=1}^n \overline{z_j} F_j(z) = 0\},$$

which is the subset where $\mathcal{F}$ is tangent to the spheres $S_r = \{\sum_{j=1}^n |z_j|^2 = r^2\}$, is smooth outside 0 and on a sufficiently small neighbourhood of 0, and there it is transverse to $\mathcal{F}$.

In the Poincaré domain we have $M = \{0\}$, i.e., small spheres around 0 are everywhere transverse to $\mathcal{F}$. The real, one-dimensional, nonsingular foliations $\mathcal{L}_r = \mathcal{F} \cap S_r$ are all isomorphic, and they have a "Morse–Smale" structure: there are n periodic trajectories (arising from the n invariant complex curves of $\mathcal{F}$ at p), with hyperbolic holonomy, and any other trajectory tends to (different) periodic trajectories in the past or in the future. The foliation $\mathcal{F}$ has, around p, a (topologically) conical structure over $\mathcal{L}_r$.

In the Siegel domain, on the contrary, M has (real) codimension 2. On a small ball $B_r = \{\sum_{j=1}^n |z_j|^2 < r^2\}$ the leaves of $\mathcal{F}_r = \mathcal{F} \cap B_r$ cutting M are discs, properly embedded in B_r and cutting M only one time. These discs are called *Siegel leaves*. Hence, the union of all Siegel leaves is an open subset of B_r which, differentiably, fibers over $M^* = M \setminus \{0\}$ with fiber $\mathbf{D}$. The other leaves of $\mathcal{F}_r$ are called *Poincaré leaves*. Their union (plus 0) is a union of complex submanifolds $\{L_I\}$ through 0, each L_I corresponding to a maximal subcollection of eigenvalues $\{\lambda_j\}_{j\in I}$ in the Poincaré domain; the dimension of L_I is the cardinality of I, and its tangent space at 0 is the direct sum of the eigenspaces associated to λ_j, $j \in I$. On each L_I the foliation $\mathcal{F}|_{L_I}$ has at 0 a singularity in the Poincaré domain, whose structure has been already described.

Suppose now that on B_r we have a closed positive current T invariant by $\mathcal{F}_r$. Through 0, there are n invariant curves $C_1, \dots, C_n$, tangent at 0 to the eigenspaces associated to $\lambda_1, \dots, \lambda_n$. For r sufficiently small, each C_j is a disc, properly embedded in B_r, and we can decompose

$$T = T_0 + \sum_{j=1}^n m_j \delta_{C_j}$$

where δ_{C_j} is the integration current over C_j and $m_j \geq 0$ is the mass attributed to C_j by the invariant measure (in terms of T, it is its Lelong number along C_j). Obviously, T_0 also is a closed positive current invariant by $\mathcal{F}_r$.

LEMMA (2.1).

$$\mathrm{Res}\,(\mathcal{F}_r, T_0, p) = 0.$$

Proof. If p belongs to the Poincaré domain, then $T_0 \equiv 0$. Indeed, T induces an invariant measure for $\mathcal{L}_r = \mathcal{F} \cap S_r$, and the Morse–Smale structure of $\mathcal{L}_r$ implies that such a measure is concentrated on the n periodic trajectories of $\mathcal{L}_r$. It follows that T is concentrated on the n invariant curves $C_1, \dots, C_n$, that is $T_0 \equiv 0$.

If p belongs to the Siegel domain, let $\{L_I\}$ be the collection of smooth submanifolds through 0 filled by the Poincaré leaves, and let L be their union. Write

$$T = \chi_L T + \chi_{B_r \setminus L} T.$$

This decomposition of T corresponds to the decomposition of the invariant measure into the part on L and the part on $B_r \setminus L$. For each L_I, $\chi_{L_I} T$ is a closed positive current invariant by $\mathcal{F}_r|_{L_I}$, which has at 0 a singularity in the Poincaré domain. By the previous discussion, we obtain $\chi_L T = \sum_{j=1}^n m_j \delta_{C_j}$ and therefore

$$T_0 = \chi_{B_r \setminus L} T.$$

On $B_r \setminus L$, the foliation has a global cross section M^*, and it is defined by a submersion

$$\pi \colon B_r \setminus L \to M^*$$

whose fibers are discs, with boundary on S_r. The current T_0 naturally induces a measure μ on M^*, so that on any test 2-form η we have, as in Section 1,

$$T_0(\eta) = \int_{M^*} \left(\int_{\pi^{-1}(q)} \eta \right) d\mu(q).$$

Let us use this expression to compute $\mathrm{Res}\,(\mathcal{F}_r, T_0, p) = \frac{1}{2\pi\sqrt{-1}} T_0(d\varphi \wedge \beta)$. Because $\beta|_{\pi^{-1}(q)}$ is holomorphic, we have

$$\int_{\pi^{-1}(q)} d\varphi \wedge \beta = \int_{\pi^{-1}(q)} d(\varphi\beta) = \int_{\partial(\pi^{-1}(q))} \beta = 0$$

for every $q \in M^*$. A fortiori,

$$\mathrm{Res}\,(\mathcal{F}_r, T_0, p) = \int_{M^*} 0 \, d\mu = 0.$$

□

We can now complete the proof of the Theorem.

Let $\mathcal{F}$ be a foliation in $\mathbb{C}P^n$ without invariant algebraic curves and all of whose singularities are hyperbolic. By [L-S], the set of these foliations in $\mathrm{Fol}\,(d, n)$, $d \geq 2$, is open (even in the real analytic Zariski topology) and dense. Suppose, by contradiction, that $\mathcal{F}$ admits an invariant closed positive current T. The absence of invariant algebraic curves implies that, on a neighbourhood of $p \in \mathrm{Sing}(\mathcal{F})$, the local invariant curves of $\mathcal{F}$ at p have no mass, i.e. $T = T_0$ in the notation above. Hence $\mathrm{Res}(\mathcal{F}, T, p) = 0$ by Lemma (2.1), and

$$c_1(\det N^*_{\mathcal{F}}) \cdot [T] = 0$$

by Proposition (1.1).

On the other hand, $\det N^*_{\mathcal{F}} = K_{\mathbb{C}P^n} \otimes T_{\mathcal{F}} = \mathcal{O}(-n-1) \otimes \mathcal{O}(1-d) = \mathcal{O}(-n-d)$ is a negative line bundle, which has negative degree on any positive homology class. This contradiction shows that $\mathcal{F}$ has no invariant measure.

Remark (2.2). By the same argument, we see that for any foliation $\mathcal{F}$ all of whose singularities are hyperbolic the only $\mathcal{F}$-invariant measures are those concentrated on $\mathcal{F}$-invariant algebraic curves.

Received November 30, 2004

Final version received January 26, 2006

MARCO BRUNELLA
INSTITUT DE MATHÉMATIQUES DE BOURGOGNE
9, AVENUE SAVARY
21078 DIJON
FRANCE
brunella@u-bourgogne.fr

REFERENCES

[B-B] P. BAUM, R. BOTT, *On the zeroes of meromorphic vector fields*, Essais en l'honneur de De Rham (1970), 29-74.

[Bru] M. BRUNELLA, *Courbes entières et feuilletages holomorphes*, Enseign. Math. **45** (1999), 195–216.

[CKP] C. CAMACHO, N. KUIPER, J. PALIS, *The topology of holomorphic flows with singularity*, Inst. Hautes Etudes Sci. Publ. Math. **48** (1978), 5–38.

[CLS] C. CAMACHO, A. LINS NETO, P. SAD, *Minimal sets of foliations on complex projective spaces*, Inst. Hautes Etudes Sci. Publ. Math. **68** (1988), 187–203.

[C-G] A. CANDEL, X. GOMEZ–MONT, *Uniformization of the leaves of a rational vector field*, Ann. Inst. Fourier **45** (1995), 1123–1133.

[Cha] M. CHAPERON, *C^k-conjugacy of holomorphic flows near a singularity*, Inst. Hautes Etudes Sci. Publ. Math. **64** (1986), 143–183.

[G-O] X. GOMEZ–MONT, L. ORTIZ–BOBADILLA, *Sistemas dinámicos holomorfos en superficies*, Aportaciones Mat. Investig. **3** (1989).

[Har] R. HARVEY, *Holomorphic chains and their boundaries*, Proc. Sympos. Pure Math. **30** (1977), 309–382.

[L-S] A. LINS NETO, M. SOARES, *Algebraic solutions of one-dimensional foliations*, J. Differential Geom. **43** (1996), 652–673.

[Sul] D. SULLIVAN, *Cycles for the dynamical study of foliated manifolds and complex manifolds*, Invent. Math. **36** (1976), 225–255.

Bol. Soc. Mat. Mexicana (3) Vol. 12, 2006

A FIXED POINT APPROACH TO THE STABILITY OF THE CUBIC FUNCTIONAL EQUATION

SOON-MO JUNG AND TAE-SOO KIM

ABSTRACT. Cădariu and Radu applied the fixed point method to the investigation of the Cauchy and Jensen functional equations. In this paper, we will adapt the idea of Cădariu and Radu to prove the Hyers-Ulam-Rassias stability of the cubic functional equation for a large class of functions from a vector space into a complete β-normed space.

1. Introduction

In 1940, S. M. Ulam [15] gave a wide ranging talk before the mathematics club of the University of Wisconsin in which he discussed a number of important unsolved problems. Among those was the question concerning the stability of group homomorphisms:

Let G_1 be a group and let G_2 be a metric group with the metric $d(\cdot, \cdot)$. Given $\varepsilon > 0$, does there exist a $\delta > 0$ such that if a function $h\colon G_1 \to G_2$ satisfies the inequality $d(h(xy), h(x)h(y)) < \delta$ for all $x, y \in G_1$, then there exists a homomorphism $H\colon G_1 \to G_2$ with $d(h(x), H(x)) < \varepsilon$ for all $x \in G_1$?

The case of approximately additive functions was solved by D. H. Hyers [5] under the assumption that G_1 and G_2 are Banach spaces. Indeed, he proved that each solution of the inequality $\|f(x+y) - f(x) - f(y)\| \leq \varepsilon$, for all x and y, can be approximated by an exact solution, say an additive function. Ten years after the publication of Hyers's theorem, D. G. Bourgin extended the theorem of Hyers and stated it in his paper [1] without proof.

Th. M. Rassias [13] attempted to weaken the condition for the bound of the norm of the Cauchy difference as follows,

$$\|f(x+y) - f(x) - f(y)\| \leq \varepsilon(\|x\|^p + \|y\|^p),$$

and generalized the result of Hyers. Since then, the stability of several functional equations has been extensively investigated.

The term Hyers-Ulam-Rassias stability originates from this historical background. The terminology can also be applied to the case of other functional equations. For more detailed definitions of such terminologies, we can refer to [4], [6], [7], [9], [10], [14].

Let E_1 and E_2 be real vector spaces. A function $f : E_1 \to E_2$ is called a *cubic function* if and only if f is a solution function of the cubic functional equation

$$f(2x+y) + f(2x-y) = 2f(x+y) + 2f(x-y) + 12f(x). \tag{1.1}$$

2000 *Mathematics Subject Classification*: Primary 39B82; secondary 47H10.
Keywords and phrases: Hyers-Ulam-Rassias stability, cubic functional equation, fixed point method.

It is known that a function $f: E_1 \to E_2$ is a cubic function if and only if there exists a function $B: E_1 \times E_1 \times E_1 \to E_2$ such that $f(x) = B(x, x, x)$ for all $x \in E_1$, and B is symmetric for any one fixed variable and is additive for two fixed variables. (See [8], Theorem 2.1).

Recently, L. Cădariu and V. Radu [3] applied the fixed point method to the investigation of the Cauchy additive functional equation (ref. [2], [12]). Using this clever idea, they present a new proof.

In this paper, we will adopt the idea of Cădariu and Radu to prove the Hyers-Ulam-Rassias stability of the cubic functional equation for a large class of functions between a vector space and a complete β-normed space.

2. Preliminaries

Let X be a set. A function $d: X \times X \to [0, \infty]$ is called a *generalized metric* on X if and only if d satisfies

(M_1) $d(x, y) = 0$ if and only if $x = y$;

(M_2) $d(x, y) = d(y, x)$ for all $x, y \in X$;

(M_3) $d(x, z) \le d(x, y) + d(y, z)$ for all $x, y, z \in X$.

Note that the only substantial difference of the generalized metric from a metric is that the range of a generalized metric includes infinity.

We now introduce one of fundamental results of fixed point theory. For the proof, refer to [11].

THEOREM (2.1). *Let (X, d) be a generalized complete metric space. Assume that $\Lambda: X \to X$ is a strictly contractive operator with the Lipschitz constant $L < 1$. If there exists a nonnegative integer k such that $d(\Lambda^{k+1}x, \Lambda^k x) < \infty$ for some $x \in X$, then the following are true:*

(a) *The sequence $\{\Lambda^n x\}$ converges to a fixed point x^* of Λ;*

(b) *x^* is the unique fixed point of Λ in*

$$X^* = \{y \in X \mid d(\Lambda^k x, y) < \infty\};$$

(c) *If $y \in X^*$, then*

$$d(y, x^*) \le \frac{1}{1-L} d(\Lambda y, y).$$

Throughout this paper, we fix a real number β with $0 < \beta \le 1$ and let $\mathbb{K}$ denote either $\mathbb{R}$ or $\mathbb{C}$. Suppose E is a vector space over $\mathbb{K}$. A function $\|\cdot\|_\beta: E \to [0, \infty)$ is called a β-norm if and only if it satisfies

(N_1) $\|x\|_\beta = 0$ if and only if $x = 0$;

(N_2) $\|\lambda x\|_\beta = |\lambda|^\beta \|x\|_\beta$ for all $\lambda \in \mathbb{K}$ and all $x \in E$;

(N_3) $\|x + y\|_\beta \le \|x\|_\beta + \|y\|_\beta$ for all $x, y \in E$.

3. Main results

In the following theorems, by using the idea of Cădariu and Radu (see [2], [3]), we will prove the Hyers-Ulam-Rassias stability of the cubic functional equation in a more general setting.

THEOREM (3.1). *Let E_1 and E_2 be vector spaces over $\mathbb{K}$ and let E_2 be a complete β-normed space, where $0 < \beta \leq 1$. Suppose $\varphi\colon E_1 \times E_1 \to [0, \infty)$ is a given function and there exists a constant L, $0 < L < 1$, such that*

$$\varphi(2x, 0) \leq 8^{\beta} L\varphi(x, 0) \tag{3.2}$$

for all $x \in E_1$. Furthermore, let $f : E_1 \to E_2$ be a function which satisfies

$$\|f(2x + y) + f(2x - y) - 2f(x + y) - 2f(x - y) - 12f(x)\|_{\beta} \leq \varphi(x, y) \tag{3.3}$$

for all $x, y \in E_1$. If φ satisfies

$$\lim_{n\to\infty} \frac{\varphi(2^n x, 2^n y)}{8^{n\beta}} = 0 \tag{3.4}$$

for every $x, y \in E_1$, then there exists a unique cubic function $T\colon E_1 \to E_2$ such that

$$\|f(x) - T(x)\|_{\beta} \leq \frac{1}{16^{\beta}} \frac{1}{1 - L} \varphi(x, 0) \tag{3.5}$$

for all $x \in E_1$.

Proof. First, we set $X = \{h \mid h : E_1 \to E_2\}$ and introduce a generalized metric on X as follows,

$$d(g, h) = \inf\{C \in [0, \infty] \mid \|g(x) - h(x)\|_{\beta} \leq C\varphi(x, 0) \text{ for all } x \in E_1\}.$$

Let $\{f_n\}$ be a Cauchy sequence in (X, d). According to the definition of Cauchy sequences, there exists, for any given $\varepsilon > 0$, a positive integer N_ε such that $d(f_m, f_n) \leq \varepsilon$ for all $m, n \geq N_\varepsilon$. By considering the definition of the generalized metric d, we see that

$$\forall\, \varepsilon > 0\ \exists\, N_\varepsilon \in \mathbb{N}\ \forall\, m, n \geq N_\varepsilon\ \forall\, x \in E_1 : \|f_m(x) - f_n(x)\|_{\beta} \leq \varepsilon\varphi(x, 0) \tag{3.6}$$

If x is any fixed point of E_1, (3.6) implies that $\{f_n(x)\}$ is a Cauchy sequence in E_2. Since E_2 is complete, $\{f_n(x)\}$ converges in E_2 for each $x \in E_1$. Hence we can define a function $f : E_1 \to E_2$ by

$$f(x) = \lim_{n\to\infty} f_n(x)$$

for any $x \in E_1$.

If we let m increase to infinity, it follows from (3.6) that for any $\varepsilon > 0$, there exists a positive integer N_ε with $\|f_n(x) - f(x)\|_{\beta} \leq \varepsilon\varphi(x, 0)$ for all $n \geq N_\varepsilon$ and for all $x \in E_1$; i.e., for any $\varepsilon > 0$, there exists a positive integer N_ε such that $d(f_n, f) \leq \varepsilon$ for any $n \geq N_\varepsilon$. This fact leads us to the conclusion that $\{f_n\}$ converges in (X, d). Hence (X, d) is a complete space (*cf*. the proof of [3], Theorem 2.5.)

We now define an operator $\Lambda\colon X \to X$ by

$$(\Lambda h)(x) = \frac{1}{8} h(2x)$$

for all $x \in E_1$.

First, we assert that Λ is strictly contractive on X. Given $g, h \in X$, let $C \in [0, \infty]$ be an arbitrary constant with $d(g, h) \leq C$; i.e.,

$$\|g(x) - h(x)\|_{\beta} \leq C\varphi(x, 0)$$

for all $x \in E_1$. If we replace x in the last inequality by $2x$ and make use of (3.2), then we have

$$\|(\Lambda g)(x) - (\Lambda h)(x)\|_\beta \leq LC\varphi(x, 0)$$

for every $x \in E_1$, i.e., $d(\Lambda g, \Lambda h) \leq LC$. Hence, we conclude that $d(\Lambda g, \Lambda h) \leq Ld(g, h)$ for any $g, h \in X$.

Next, we assert that $d(\Lambda f, f) < \infty$. If we substitute $y = 0$ in (3.3) and we divide both sides by 16^β, then (3.2) establishes

$$\|(\Lambda f)(x) - f(x)\|_\beta \leq \frac{1}{16^\beta}\varphi(x, 0)$$

for any $x \in E_1$, i.e.,

$$d(\Lambda f, f) \leq \frac{1}{16^\beta} < \infty. \tag{3.7}$$

Then it follows from Theorem (2.1) (a) that there exists a function $T: E_1 \to E_2$ which is a fixed point of Λ, such that $d(\Lambda^n f, T) \to 0$ as $n \to \infty$, i.e.,

$$\lim_{n\to\infty} \frac{1}{8^n} f(2^n x) = T(x) \tag{3.8}$$

for all $x \in E_1$.

By Theorem (2.1) (c) and (3.7), we obtain

$$d(f, T) \leq \frac{1}{1-L} d(\Lambda f, f) \leq \frac{1}{16^\beta}\frac{1}{1-L}; \tag{3.9}$$

i.e., the inequality (3.5) is true for all $x \in E_1$.

Now substitute $2^n x$ and $2^n y$ for x and y in (3.3), respectively. If we divide both sides of the resulting inequality by $8^{n\beta}$, and let n go to infinity, it follows from (3.4) and (3.8) that T is a cubic function.

Assume that inequality (3.5) is also satisfied with another cubic function $T_1: E_1 \to E_2$ besides T. (We know that T_1 is a fixed point of Λ.) In view of (3.5) and the definition of d, we conclude that inequality (3.9) is true with T_1 in place of T. Due to Theorem (2.1) (*b*), we get $T = T_1$. This proves the uniqueness of T. □

The above result can be compared with [8], Theorem 3.1. In our theorem, we dealt with the Hyers-Ulam-Rassias stability of the cubic functional equation (1.1) for a complete β-normed space, where $0 < \beta \leq 1$, while the authors of [8] proved the Hyers-Ulam-Rassias stability of (1.1) for a Banach space, which is the case of a complete 1-normed space.

In a similar way as in the proof of Theorem (3.1), we also apply Theorem (2.1) and prove the following theorem.

THEOREM (3.10). *Let E_1 and E_2 be a vector space over $\mathbb{K}$ and a complete β-normed space over $\mathbb{K}$, respectively. Assume that $\varphi: E_1 \times E_1 \to [0, \infty)$ is a given function and there exists a constant L, $0 < L < 1$, such that*

$$\varphi(x, 0) \leq \frac{1}{8^\beta} L\varphi(2x, 0)$$

for all $x \in E_1$. Furthermore, assume that $f: E_1 \to E_2$ is a given function and satisfies the inequality (3.3) for all $x, y \in E_1$. If φ satisfies

$$\lim_{n\to\infty} 8^{n\beta}\varphi\left(\frac{x}{2^n}, \frac{y}{2^n}\right) = 0$$

for every $x, y \in E_1$, then there exists a unique cubic function $T\colon E_1 \to E_2$ such that

$$\|f(x) - T(x)\|_\beta \leq \frac{1}{16^\beta}\frac{L}{1-L}\varphi(x,0) \tag{3.11}$$

for any $x \in E_1$.

Proof. We use the definitions for X and d, the generalized metric on X, as in the proof of Theorem (3.1). Then (X, d) is complete.

We define an operator $\Lambda\colon X \to X$ by

$$(\Lambda h)(x) = 8h\left(\frac{x}{2}\right)$$

for all $x \in E_1$. We apply the same argument as in the proof of Theorem (3.1) and prove that Λ is a strictly contractive operator. Moreover, we prove

$$d(\Lambda f, f) \leq \frac{1}{16^\beta}L \tag{3.12}$$

instead of (3.7).

According to (a) of Theorem (2.1), there exists a function $T\colon E_1 \to E_2$ which is a fixed point of Λ such that

$$\lim_{n\to\infty} 8^n f\left(\frac{x}{2^n}\right) = T(x)$$

for each $x \in E_1$.

Using Theorem (2.1) (c) and (3.12), we get

$$d(f, T) \leq \frac{1}{1-L}d(\Lambda f, f) \leq \frac{1}{16^\beta}\frac{L}{1-L},$$

which implies the validity of inequality (3.11).

If we replace x and y in (3.3) by $\frac{x}{2^n}$ and $\frac{y}{2^n}$, respectively, and then multiply both sides of the resulting inequality by $8^{n\beta}$, then we can prove that T is a cubic function. For the uniquness of T we can apply the same argument as in the last part of the proof of Theorem (3.1). □

In the following corollaries, we will give some special cases for Theorems (3.1) and (3.10).

COROLLARY (3.13). *Fix a positive number p less than 3 and choose a constant β with $\frac{p}{3} < \beta \leq 1$. Let E_1 and E_2 be a normed space over $\mathbb{K}$ and a complete β-normed space over $\mathbb{K}$, respectively. If a function $f : E_1 \to E_2$ satisfies*

$$\|f(2x+y)+f(2x-y)-2f(x+y)-2f(x-y)-12f(x)\|_\beta \leq \varepsilon\left(\|x\|^p + \|y\|^p\right) \tag{3.14}$$

for all $x, y \in E_1$ and for some $\varepsilon > 0$, then there exists a unique cubic function $T\colon E_1 \to E_2$ such that

$$\|f(x) - T(x)\|_\beta \leq \frac{\varepsilon}{2^\beta|8^\beta - 2^p|}\|x\|^p \tag{3.15}$$

for any $x \in E_1$.

Proof. If we set $\varphi(x, y) = \varepsilon(\|x\|^p + \|y\|^p)$ for all $x, y \in E_1$ and if we set $L = \frac{2^p}{8^\beta}$, then we have $0 < L < 1$ and

$$\varphi(2x, 0) = 2^p \varepsilon \|x\|^p = 2^p \varphi(x, 0) = 8^\beta L \varphi(x, 0)$$

for all $x \in E_1$.

Furthermore, we get

$$\frac{\varphi(2^n x, 2^n y)}{8^{n\beta}} = L^n \varepsilon(\|x\|^p + \|y\|^p) \to 0, \text{ as } n \to \infty,$$

for any $x, y \in E_1$.

According to Theorem (3.1), there exists a unique cubic function $T : E_1 \to E_2$ such that the inequality (3.15) holds for every $x \in E_1$. □

COROLLARY (3.16). *Assume that p is a real constant larger than 3. Let E_1 and E_2 be a normed space over $\mathbb{K}$ and a complete β-normed space over $\mathbb{K}$, respectively. If $f : E_1 \to E_2$ is a function which satisfies the inequality (3.14) for all $x, y \in E_1$ and for some $\varepsilon > 0$, then there exists a unique cubic function $T : E_1 \to E_2$ such that the inequality (3.15) holds for all $x \in E_1$.*

Proof. If we set $\varphi(x, y) = \varepsilon(\|x\|^p + \|y\|^p)$ for any $x, y \in E_1$, then we obtain

$$\varphi(x, 0) = \varepsilon\|x\|^p = \frac{1}{8^\beta} L \varphi(2x, 0)$$

for each $x \in E_1$, where $L = \frac{8^\beta}{2^p}$ is less than 1 because $0 < \beta \le 1$ and $p > 3$.

Moreover, we have

$$8^{n\beta} \varphi\Big(\frac{x}{2^n}, \frac{y}{2^n}\Big) = L^n \varepsilon(\|x\|^p + \|y\|^p) \to 0, \text{ as } n \to \infty,$$

for all $x, y \in E_1$.

In view of Theorem (3.10), there exists a unique cubic function $T : E_1 \to E_2$ for which the inequality (3.15) is true for any $x \in E_1$. □

Can we also expect the Hyers-Ulam-Rassias stability of the cubic functional equation for the special case of $p = 3$? It is still open whether the cubic functional equation (1.1) is stable when $p = 3$.

Received July 04, 2005

Final version received December 15, 2005

SOON-MO JUNG
MATHEMATICS SECTION
COLLEGE OF SCIENCE AND TECHNOLOGY
HONG-IK UNIVERSITY
339-701 CHOCHIWON
KOREA
smjung@hongik.ac.kr

TAE-SOO KIM
DEPARTMENT OF MATHEMATICS
CHUNGBUK NATIONAL UNIVERSITY
361-763 CHEONGJU, KOREA
kimtaesoo@chungbuk.ac.kr

References

[1] D. G. Bourgin, *Classes of transformations and bordering transformations*, Bull. Amer. Math. Soc. **57** (1951), 223–237.

[2] L. Cădariu and V. Radu, *Fixed points and the stability of Jensen's functional equation*, JIPAM. J. Inequal. Pure and Appl. Math. **4** (1) (2003), Art. 4 (`http://jipam.vu.edu.au`)

[3] L. Cădariu and V. Radu, *On the stability of the Cauchy functional equation: a fixed point approach*, Grazer Math. Ber. **346** (2004), 43–52.

[4] G. L. Forti, *Hyers-Ulam stability of functional equations in several variables*, Aequationes Math. **50** (1995), 143–190.

[5] D. H. Hyers, *On the stability of the linear functional equation*, Proc. Nat. Acad. Sci. USA. **27** (1941), 222–224.

[6] D. H. Hyers, G. Isac and Th. M. Rassias, Stability of Functional Equations of Several Variables, Birkhäuser, Basel/Boston, 1998.

[7] D. H. Hyers and Th. M. Rassias, *Approximate homomorphisms*, Aequationes Math. **44** (1992), 125–153.

[8] K.-W. Jun and H.-M. Kim, *The generalized Hyers-Ulam-Rassias stability of a cubic functional equation*, J. Math. Anal. Appl. **274** (2002), 867–878.

[9] S.-M. Jung, *Hyers-Ulam-Rassias stability of functional equations*, Dynam. Systems Appl. **6** (1997), 541–566.

[10] S.-M. Jung, Hyers-Ulam-Rassias Stability of Functional Equations in Mathematical Analysis, Hadronic Press, Palm Harbor, 2001.

[11] B. Margolis and J. Diaz, *A fixed point theorem of the alternative for contractions on a generalized complete metric space*, Bull. Amer. Math. Soc. **74** (1968), 305–309.

[12] V. Radu, *The fixed point alternative and the stability of functional equations*, Fixed Point Theory – An International Journal on Fixed Point Theory Computation and Applications **4** (2003), 91–96.

[13] Th. M. Rassias, *On the stability of the linear mapping in Banach spaces*, Proc. Amer. Math. Soc. **72** (1978), 297–300.

[14] Th. M. Rassias, *On the stability of functional equations and a problem of Ulam*, Acta Appl. Math. **62** (2000), 23–130.

[15] S. M. Ulam, A Collection of Mathematical Problems, Interscience Publ., New York, 1960.

Bol. Soc. Mat. Mexicana (3) Vol. 12, 2005

A TRANSVERSAL PROPERTY OF FAMILIES OF EIGHT OR NINE UNIT DISKS

T. BISZTRICZKY, F. FODOR AND D. OLIVEROS

ABSTRACT. For a family $\mathcal{F}$ of n disjoint unit disks in the plane with the property $T(4)$, we show that if there is an $(n-2)$-transversal that strictly separates two elements of $\mathcal{F}$ then $\mathcal{F}$ has the property $T-1$; that is, it has an $(n-1)$-transversal. We apply this generic result to verify that $T(4)$ implies $T-1$ for families $\mathcal{F}$ of eight or nine disks.

1. Introduction

Let $\mathcal{F}$ denote a family of mutually disjoint translates of a compact convex set (with non-empty interior) in the Euclidean plane. Then $\mathcal{F}$ has a *transversal* and the property T if there is a line that intersects all members of $\mathcal{F}$. If there is a line that meets all but at most k members of $\mathcal{F}$ then $\mathcal{F}$ has property $T-k$. Next, if each k-element subfamily of $\mathcal{F}$ has the property T then we say that $\mathcal{F}$ has the property $T(k)$. Finally, an *m-transversal* of $\mathcal{F}$ is a line that meets m elements of $\mathcal{F}$. An m-transversal is *separating* if there are elements of $\mathcal{F}$ in each of the open half-planes determined by it.

We recall that Transversal Theory has its origin in Helly's Theorem [5] and [7], and that major results about the property $T(k)$ are due to Tverberg for $k=5$, and Katchalski and Lewis for $k=3$. With regard to $T(3)$, A. Heppes ([2], [3]) has recently shown $T-2$ for disjoint unit disks. With regard to $T(4)$, it is known that for families of disjoint unit disks: $T(4)$ does not imply $T[1]$, and $T(4)$ implies $T-5$ [3]. Since $T(5)$ implies T, and $T(3)$ implies $T-2$, it is reasonable to conjecture that $T(4)$ implies $T-1$. This has been verified for six and seven disks in [4].

Henceforth, we assume that the members of $\mathcal{F}$ are unit disks. We denote points of the plane by $a, b, c, \ldots,$ and all other sets by $A, B, C, \ldots$. The convex hull of $A \cup B$ is denoted by $[A, B]$.

In the following, we assume also that an n-element $\mathcal{F}$ has the property $T(4)$. In Section 2, we present general results about $\mathcal{F}$. The aim is the Main Lemma: the existence of an $(n-1)$-transversal follows from the existence of a separating $(n-2)$-transversal, and we show the existence of such a line in the Reduction Lemma. The gain is in Section 3, where we show in a straightforward manner that if $|n| \leq 9$ then $\mathcal{F}$ has property $T-1$.

2000 *Mathematics Subject Classification*: 52A35.

Keywords and phrases: Helly's theorem, Katchalski-Lewis conjecture, transversal line, unit disk.

First named author partially supported by an NSERC Discovery Grant.

Second named author partially supported by the Békésy György Postdoctoral Fellowship of the Hungarian Ministry of Education, and by Hungarian OTKA Grant F042959.

2. Separating $(n-2)$-Transversals

In this section, we assume that $|\mathcal{F}| = n \geq 8$ and that H is a separating $(n-2)$-transversal of $\mathcal{F}$. For disjoint A and B in $\mathcal{F}$, the *tangential separators* are the two common tangents that separate them.

LEMMA (2.1). *Let $\{X, Y\} \subset \mathcal{F}$, H strictly separate X and Y, and L and L' be the tangential separators for X and Y. Let $\mathcal{H} = \{B_i \in \mathcal{F} \mid B_i \cap H \neq \emptyset$ and $B_i \cap \left[X, Y\right] = \emptyset\}$. Then either each $B_i \in \mathcal{H}$ meets L or each $B_i \in \mathcal{H}$ meets L'.*

Proof. Let $B_j \in \mathcal{H}$ and H be horizontal. We note that some vertical line that meets $[X, Y]$ is disjoint from B_j. Hence, $\{X, Y, B_j\}$ has either positive-sloped or negative-sloped transversals. Assume the former. Then by $T(4)$, $\{X, Y, B_j, B_i\}$ has only positive-sloped transversals for each $B_i \in \mathcal{H}$. Thus, each $B_j \in \mathcal{H}$ meets either L or L', depending upon which has the smaller positive slope. □

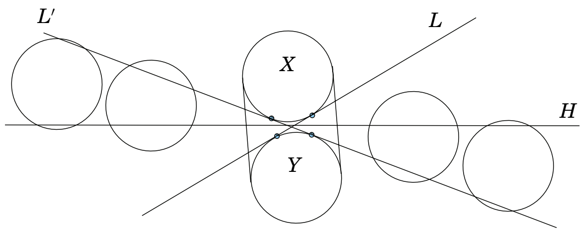


Figure 1.

REDUCTION LEMMA. *Let $\{X, Y\} \subset \mathcal{F}$ such that H strictly separates X and Y, and there is a line orthogonal to H that intersects X and Y. Then there is an $(n-1)$-transversal of $\mathcal{F}$ or an $(n-2)$-transversal H' and $\{X', Y'\} \subset \mathcal{F}$ such that H' strictly separates X' and Y', and at most two elements of $\mathcal{F} \setminus \{X', Y'\}$ meet the line segment $H' \cap \left[X', Y'\right]$.*

Sketch of proof. Let L and L' be the tangential separators of X and Y, and (cf. Lemma (2.1)) $L \cap B_i \neq \emptyset$ for each $B_i \in \mathcal{H}$. We assume that H is horizontal and L has positive slope. Then L' is vertical or has negative slope.

We show that $|\mathcal{F}| \leq 7$ or

(RL1) *there is at most one element of $\mathcal{F}$ in each open half-plane determined by L, and*

(RL2) *no element of $\mathcal{F} \setminus \{X, Y\}$ meets $\left[L \cap X, L \cap Y\right]$; that is, L is an $(n-1)$-transversal or L is an $(n-2)$-transversal that strictly separates, say, A and C in $\mathcal{F}$ and no element of $\mathcal{F} \setminus \{X, Y, A, C\}$ meets $L \cap \left[A, C\right]$.*

Proof of (RL1). Let L^+ and L^- be the open half-planes determined by L, and let A and B be disks such that $A \cup B \subset L^-$. Then A and B meet $\left[X, Y\right]$ by Lemma (2.1). We assume that L' is vertical and show that $|\mathcal{F}| \leq 7$. The advantage of the orthogonality of H and L' is the case calculations (cf. Figure 2). As will be readily seen from the observations, if $|\mathcal{F}| \leq 7$ with X and Y in the assumed limit positions then $|\mathcal{F}| \leq 7$ with X and Y in arbitrary positions.

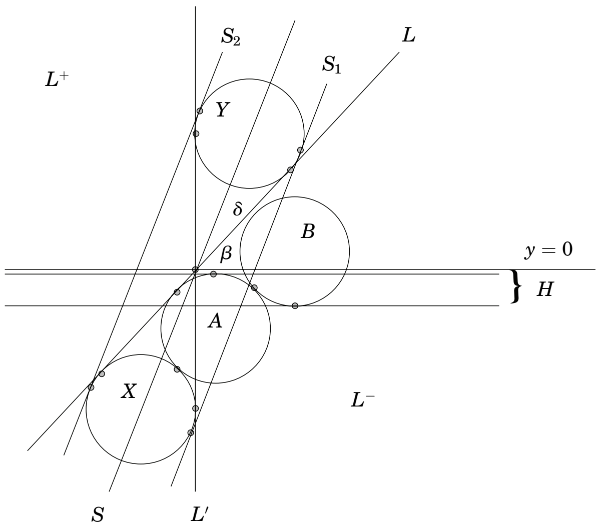


Figure 2.

With $L \cap L'$ as the origin, β as the acute angle between H and L, $\alpha = \beta + \delta$ as the acute angle between H and the line S through the centres $(-1, -\tan\alpha)$ and $(1, \tan\alpha)$ of X and Y, and as needed, A and B in limit positions (tangential to each other and possibly to L, S_1 and H), we have the following observations:

i) *A and B are the only elements of $\mathcal{F}$ contained in L^-.*

Consider A with centre $(2\csc\alpha - 1, -\tan\alpha)$ and B with centre $(2(\csc\alpha + \cot\alpha) - 1, 2 - \tan\alpha)$. Then S_1 is tangent to A and B, the line $y = 1 - \tan\alpha$ is tangent to A, B and X. Then, any disk tangent to L from the right and to X from above meets $A \cup B$. Hence, any $U \in \mathcal{F}$ such that $U \subset L^-$ meets A or B.

ii) *A and B exist only if $\beta > 23°$.*

Consider A with centre $(1, -\tan\alpha)$ and β with centre $(2\cos 2\beta + 1, 2\sin 2\beta - \tan\alpha)$. Then A meets X and B tangentially, L supports B and if $\beta \leq 23°$, $B \cap [X, Y] = \emptyset$. Specifically, if $\beta \leq 23°$ and $A \subset L^-$ then no element of $\mathcal{F} \setminus \{X, Y\}$ meets both $[L \cap X, L \cap Y]$ and $S_1 \cap [X, Y]$.

iii) *There is no $W \in \mathcal{F}$, disjoint from $[X, Y]$, meeting the upper component of $L \setminus [X, Y]$.*

Consider A and B such that H supports A from above, and B meets L and S_1 tangentially. Let W be a unit disk that is disjoint from $[X, Y]$, meets B tangentially and is supported by H from below. Then W may be an element of $\mathcal{F}$ only if $L \cap (\mathrm{int} W) \neq \emptyset$.

If $\beta = 30°$ then L meets A, B and W tangentially. If $\beta > 30°$ then L meets A and B tangentially but is disjoint from W. If $\beta < 30°$ and decreasing then A loses contact with L, H approaches $y = 1 - \tan\alpha$ and L is again disjoint from W.

iv) *There is a $W \in \mathcal{F}$, disjoint from $[X, Y]$, meeting the lower component of $L \setminus [X, Y]$ only if $\beta < 50°$.*

Consider the disk W meeting L and S_2 tangentially. Then W may be a member of $\mathcal{F}$ only if it meets the H with minimum height; that is, $H : y = c$ (minimum). If L supports A and B then $c = 4\sin\beta - \cot\delta - 1$, and $H \cap W = \emptyset$ for $\beta \geq 50^\circ$.

Let $\beta \geq 50^\circ$ and suppose that L does not support A, say. Then H may be lower than above, but W may be a member of $\mathcal{F}$ only if it meets L^* (the tangential separator of Y and A, that meets X) by $T(4)$ for $\{X, Y, A, W\}$. Let β^* be the acute angle determined by L^* and H. Then $\beta^* > \beta$ and the argument that $H \cap W = \emptyset$, when L^* supports W, is the same as above with β^* replacing β and a suitable X^* replacing X.

v) *Let $W_0 \in \mathcal{F}$ be disjoint from $[X, Y]$ and meet the lower component of $L \setminus [X, Y]$. Then $|\mathcal{F} \setminus \{X, Y, A, B, W_0\}| \leq 2$.*

In order to maximize the number of W in $\mathcal{F} \setminus \{X, Y, A, B, W_0\}$, we may assume that W_0 is as far as possible from $[X, Y]$; that is, W_0 meets L and H tangentially. Then L supports A and B, $\beta \geq 30^\circ$ and the centre of each W is in $L^+ \cup L$. Since $\beta < 50^\circ$; A or B (say B) is disjoint from L', and so, $T(4)$ for $\{X, Y, B, W\}$ yields that $L \cap W \neq \emptyset$.

Consider the disk W_1 with centre w on L and supported by H from below. Then $\operatorname{dist}(L \cap W_0, w) = 2\cot(\beta/2) - \cot\beta < 3\sqrt{3}$ for $\beta \geq 33^\circ$, and there is at most one W such that $L \cap W \subset [L \cap W_0, L \cap W_1]$.

Let $\beta < 33^\circ$. Then the reflections of A and B about L meet Y, and it follows from $L \cap A \subset [L \cap W_0, L \cap B]$ that $L \cap W \subset [L \cap W_0, L \cap A]$. We note that a limit H supports B from below or A from above. In either case, $\operatorname{dist}(L \cap W_0, L \cap A) < 2 + 2\sqrt{3}$ and hence, there are at most two W such that $L \cap W \subset [L \cap W_0, L \cap A]$.

Since each $W \in \mathcal{F} \setminus \{X, Y\}$ meets H and a vertical line meets X and Y, it is a straightforward consequence of $T(4)$ that $|\{W \in \mathcal{F} \setminus \{X, Y\} \mid W \cap [X, Y] \neq \emptyset\}| \leq 5$. From this and v), we obtain that L is an $(n-1)$-transversal or L is an $(n-2)$-transversal that strictly separates two elements of $\mathcal{F}$. In the latter case, we remark that because L is a tangential separator of X and Y, it is a limit $(n-2)$-transversal.

Proof of (RL2). In view of the above, we may assume that, say, A and C are the only elements of $\mathcal{F}$ that are disjoint from L, $A \subset L^-$, $C \subset L^+$ and H is a tangential separator of A and C. Then H is a unique horizontal $(n-2)$-transversal, and it supports A from above and C from below.

Let $B \in \mathcal{F} \setminus \{X, Y\}$ meet both $S_1 \cap [X, Y]$ and $[L \cap X, L \cap Y]$, and $W \in \mathcal{F}$. We need to show that $|\mathcal{F}| \leq 7$.

Let $S_1 \cap B \subset [S_1 \cap A, S_1 \cap Y]$. We argue as in iii) that no W, disjoint from $[X, Y]$, meets the upper component of $L \setminus [X, Y]$, and note that there are W disjoint from $[X, Y]$ only if $\beta < 40^\circ$. (Consider W meeting L and S_1 tangentially, and L supporting a limit A. Then W is supported from above by $y = c_1 = \sin\beta(\sec\beta - 2\csc\delta) + \sec\beta + 1$, $H : y = c_2 = 2\sin\beta - \cot\delta + 1$ (recall H is unique) and $c_1 \geq c_2$ only if $\beta < 40^\circ$).

Let $W_0 \in \mathcal{F}$ meet the lower component of $L \setminus [X, Y]$. As in v); in order to maximize the number of W in $\mathcal{F} \setminus \{X, Y, A, C, B, W_0\}$ that are disjoint from

$\left[X,Y\right]$ or meet $S_2 \cap \left[X,Y\right]$ and $\left[L \cap X, L \cap Y\right]$, we may assume that W_0 meets L and H tangentially, and L supports A. If $\beta \leq 36°$ then $\text{dist}(L \cap X, L \cap Y) = 2\cot\delta < 4$, $\text{dist}\left(L \cap A, L \cap Y\right) < 2$ and it follows from $S_1 \cap B \subset \left[S_1 \cap A, S_1 \cap Y\right]$ and $B \cap \left(A \cup Y\right) = \emptyset$ that $B \cap \left[L \cap X, L \cap Y\right] = \emptyset$; a contradiction. Thus $36° < \beta < 40°$, $H : y = c_2 > 0$ and $\text{dist}(L \cap W_0, L \cap Y) < \cot\delta + \cot\left(\frac{\beta}{2}\right) < 2 + 2\sqrt{3}$. It now follows that only B meets both $S_1 \cap \left[X,Y\right]$ and $\left[L \cap X, L \cap Y\right]$, and there is at most one W that meets both $S_2 \cap \left[X,Y\right]$ and $\left[L \cap W_0, L \cap Y\right]$; that is, $|\mathcal{F} \setminus \{X,Y,A,C,B,W_0\}| \leq 1$.

Let $S_1 \cap A \subset [S_1 \cap B, S_1 \cap Y]$. If $\beta < 50°$ then $L' \cap A = \emptyset$ and we obtain (as in v) that $\{X,Y,A,C\}$ satisfies $T(4)$ only if $L \cap C \neq \emptyset$; a contradiction. Let $\beta \geq 50°$. Then we argue as in iv) that there is no W, disjoint from $\left[X,Y\right]$, meeting the lower component of $L \setminus \left[X,Y\right]$, and as in v) that there is at most one W, disjoint from $\left[X,Y\right]$, meeting the upper component of $L \setminus [X,Y]$. Let W_0 be such a disk. Then no $W \in \mathcal{F} \setminus \{X,Y\}$ meets both $\left[L \cap X, L \cap Y\right]$ and $\left[S_1 \cap A, S_1 \cap Y\right]$. From this it follows that no $W \neq B$ meets both $\left[L \cap X, L \cap Y\right]$ and $S_1 \cap \left[X,Y\right]$, and thus, either $\mathcal{F} = \{X,Y,A,C,B,W_0\}$ or there is a $B' \in \mathcal{F}$ meeting both $\left[L \cap X, L \cap Y\right]$ and $S_2 \cap \left[X,Y\right]$. The existence of W_0 and iv) yield that $S_2 \cap B' \subset \left[S_2 \cap C, S_2 \cap X\right]$, and we obtain as above that $|\mathcal{F}| \leq 7$; that is, $\mathcal{F} = \{X,Y,A,C,B,B',W_0\}$. □

Main Lemma. *Let H be an $(n-2)$-transversal of $\mathcal{F}$ and $\{X,Y\} \subset \mathcal{F}$ such that H strictly separates X and Y. Then there is an $(n-1)$-transversal of $\mathcal{F}$.*

In view of the Reduction Lemma, it is sufficient to prove that $\mathcal{F}$ has a $(n-1)$-transversal if

(ML1) *no line orthogonal to H intersects both X and Y, or*

(ML2) *at most two elements of $\mathcal{F} \setminus \{X,Y\}$ meet the line segment $H \cap [X,Y]$.*

Proof of (ML1). Since no line orthogonal to H intersects both X and Y and with H horizontal, no vertical line intersects both X and Y, and, say, both tangential separators L and L' have positive slope. Let L be the one with the smaller slope. Then by Lemma (2.1), $B_i \cap L \neq \emptyset$ for each $B_i \in \mathcal{H}$. We may assume that there exists a $B_k \in \mathcal{F}$ such that $B_k \cap L = \emptyset$. Then $B_k \cap [X,Y] \neq \emptyset$. Let L^+ and L^- be the open half-planes determined by L with $Y \cap L^- = \emptyset$ and, say, $B_k \subset L^+$. We further assume that the tangential separator, denoted by M for X and B_k, that intersects Y has maximum slope for all such B_k; that is, if $B_i \cap [X,Y] \neq \emptyset$ and $B_i \subset L^+$ then $M \cap B_i \neq \emptyset$.

We note (cf. Figure 3) that the existence of H yields that $Y \cap [X, B_k] = \emptyset$, and that the existence of L yields that $X \cap [B_k, Y] = \emptyset$. Thus,

(a) $N \cap B_k \subset [N \cap X, N \cap Y]$ for any transversal N of $\{X,Y,B_k\}$.

We claim that M or L intersect all but at most one member of $\mathcal{F}$. Let M^+ and M^- be the open half-planes determined by M with $X \cap M^+ = \emptyset$. Let $B_i \in \mathcal{F} \setminus \{X,Y,B_k\}$ and N be a transversal of $\{X,Y,B_k,B_i\}$. We note that $H \cap B_i \neq \emptyset$. It is clear that

(b) if $H \cap B_i$ is left of $H \cap B_k$ then $Y \cap [X, B_i] = \emptyset$, and thus by (a) either $N \cap B_i \subset [N \cap X, N \cap Y]$ or

$$N \cap X \subset [N \cap B_i, N \cap B_k] \subset [N \cap B_i, N \cap Y].$$

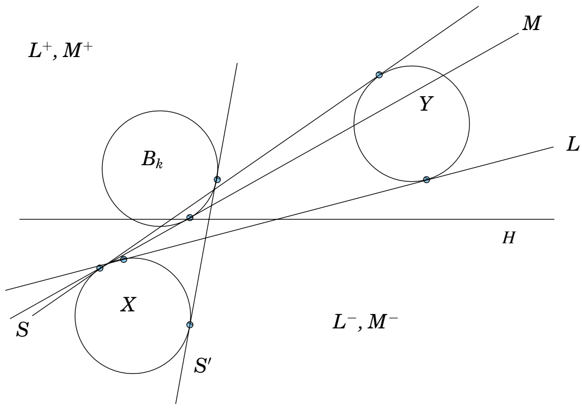


Figure 3. Configuration in Case I

Similarly, we obtain that

(c) if $H \cap B_i$ is right of $H \cap B_k$ then $X \cap [Y, B_i] = \emptyset$ and either $N \cap B_i \subset [N \cap X, N \cap Y]$ or

$$N \cap Y \subset [N \cap B_i, N \cap B_k] \subset [N \cap B_i, N \cap X].$$

CASE (1). $B_i \cap [X, Y] = \emptyset$.

If $H \cap B_i$ is left (right) of $H \cap B_k$ then (b) ((c)) implies that $X \cap [B_i, B_k]$ ($B_k \cap [X, B_i]$) is not empty, and that is possible only if $M \cap B_i \neq \emptyset$.

CASE (2). $B_i \cap [X, Y] \neq \emptyset$.

We assume that B_0 and B_1 are such B_i, and that $M \cap (B_0 \cup B_1) = \emptyset$.

If $H \cap B_0$ is left of $H \cap B_k$ then $B_0 \cap L^+ \neq \emptyset$, and it follows from the maximum property of M that $L \cap B_0 \neq \emptyset$; cf. Figure 4. We note that B_0 meets L and S_0 (the upper supporting line of X, Y and $[X, Y]$) from an obtuse angled side, and hence it is unique among B_i with $H \cap B_i$ left of $H \cap B_k$. Then $H \cap B_1$ is right of $H \cap B_k$, $B_1 \subset M^-$ and $B_k \cap [X, B_1]$ is empty. It now follows from (a) and (c) that $B_1 \cap [X, B_k] \neq \emptyset$. Since B_1 meets H and S_1 (the lower supporting line of X, B_k and $[X, B_k]$) from an obtuse angled side, it follows that it is unique among B_i with $H \cap B_i$ right of $H \cap B_k$.

In summary: M is a transversal for $\mathcal{F} \setminus \{B_0, B_1\}$ and, as a consequence, L intersects X, Y, B_0, B_1 and each B_i with $H \cap B_i \not\subset [H \cap B_0, H \cap B_1]$. Since $B_k \subset L^+$ and $B_1 \cap [X, B_k] \neq \emptyset$, it is easy to check that there is no B_i such that $H \cap B_i \subset [H \cap B_k, H \cap B_1]$.

Suppose that $B_2 \in \mathcal{F}$ is such that $H \cap B_2 \subset [H \cap B_0, H \cap B_k]$ and $L \cap B_2 = \emptyset$; cf. Figure 5. Then $M \cap B_2 \neq \emptyset$ and we observe that the slope of any transveral of $\{X, B_0, B_k\}$ is positive and larger than the slope of M. Since B_1 meets such a transversal by $T(4)$, it follows that $M \cap B_1 \neq \emptyset$; a contradiction. Thus, L is a transversal of $\mathcal{F} \setminus \{B_k\}$. $\square$

Now, assume that H is an $(n-2)$-transversal that meets $B_1, B_2, \ldots, B_{n-2} \in \mathcal{F}$. Let λ be the angle of the tangential separators L and L' for X and Y as indicated in Figure 6.

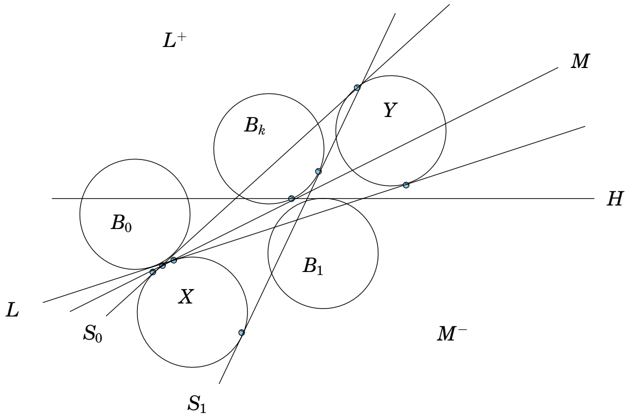


Figure 4. Configuration in Case 2

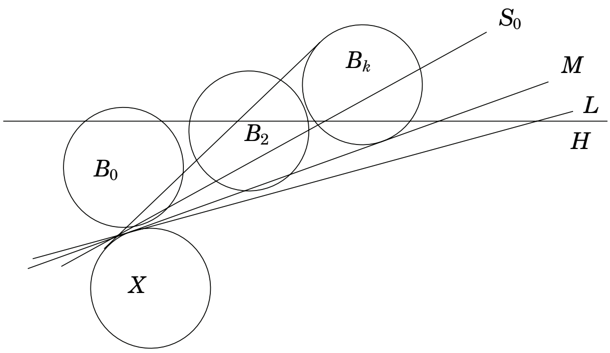


Figure 5.

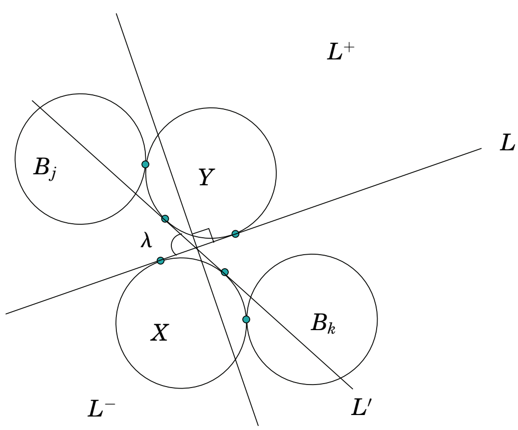


Figure 6. Extremal cases

We shall use the following lemma in the proof of (ML2).

LEMMA (2.2). *Let $\lambda \leq \frac{\pi}{3}$, L meet each $B_i \in \mathcal{H}$, and B_j and B_k be two members of $\mathcal{F}$ that are disjoint from L. Then L is an $(n-2)$-transversal, L separates B_j and B_k, and there is a line L^* that is orthogonal to L and strictly separates B_j and B_k.*

Proof. We note that both B_j and B_k meet $[X, Y]$, and hence, it follows from $\lambda \leq \frac{\pi}{3}$ that they meet also L'.

Let $B_k \subset L^-$. It is clear that B_k meets $[X, Y]$ from the right, and we claim that it is the only member of $\mathcal{F} \backslash \mathcal{H}$ that meets $[X, Y]$ from the right. It is sufficient to consider the case where $\lambda = \frac{\pi}{3}$ and B_k is as far as possible from Y (B_k meets X tangentially and its centre is on L'), and then direct computation of the distance between Y and B_k yields the claim. Thus, $B_j \subset L^+$ and L is an $(n-2)$-transversal.

In order to verify the existence of L^*, it is again sufficient to consider the case where $\lambda = \frac{\pi}{3}$ and B_k (B_j) meets $X(Y)$ tangentially and its centre is on L'. Then direct computation shows that the line, through $L \cap L'$ and orthogonal to L, strictly separates B_j and B_k. $\square$

Proof of (ML2). We assume that H is the line $y = 0$ and that the x-coordinate of the center of Y is no smaller than that of X. Furthermore, there is a vertical line that meets both X and Y. Let L be the tangential separator of X and Y that determines the smaller acute angle with H. We assume that the slope of L is positive. Then L meets each $B_i \in \mathcal{H}$ by Lemma (2.1). Since there are only two sets, A and C of $\mathcal{F}$, say that meet $H \cap [X, Y]$, we may assume from the proof of the Reduction Lemma that L is an $(n-2)$-transversal of $\mathcal{F}$ that strictly separates A and C, and no element of $\mathcal{F} \backslash \{X, Y\}$ meets $[L \cap X, L \cap Y]$. Notice that one can move L tangentially on the boundary of X until it becomes an $(n-3)$-transversal, and that L can be moved on the boundary of Y in a similar manner. We denote the first line by L_X, and the second by L_Y.

We assume that (cf. Figure 7) that $Y \cap L^- = \emptyset$, $A \subset L^+$ and that $H \cap C$ is to the right of $H \cap A$.

In view of (ML1) and Lemma (2.2), we may assume that $\lambda > \frac{\pi}{3}$ and that the acute angle determined by L and H is not arbitrarily small. The former and $L \cap B_i \neq \emptyset$ for each $B_i \in \mathcal{H}$ yield that $L' \cap B_i = \emptyset$ for each $B_i \in \mathcal{H}$. The latter means that we may assume that H is in a limit position; that is, no line H' with $0 <$ slope $H' <$ slope L is a transversal of $\mathcal{F} \backslash \{X, Y\} = \mathcal{H} \cup \{A, C\}$. Thus, H is a tangential separator of A and C that separates $A \cup Y$ and $X \cup C$. Now if $C \subset L^+$ then $C \cap [X, Y] \neq \emptyset$ and $C \cap L = \emptyset$ clearly yield that there is no $B_i \in \mathcal{H}$ such that $B_i \cap [X, Y] = \emptyset$ and B_i meets both H and L. Thus, $\mathcal{H} = \emptyset$ and $|\mathcal{F}| = 4$; a contradiction.

Finally, let M and M' be the tangential separators of A and X. Since A meets $[X, Y]$ and is disjoint from L, it follows that, say, M meets Y and $M' \cap Y = \emptyset$. We note that $C \subset L^-$ yields that $M \cap C = \emptyset$.

We consider now the relative positions of A, C, X and Y.

CASE (1). $(Y \cup C) \cap [X, A] = \emptyset$.

Note that $M \cap Y \neq \emptyset$, $M' \cap Y = \emptyset$ and L_Y strictly separates A, X and meets Y. By Lemma (2.1), it follows that all B_i to the right of $[Y, C]$ meet M. Next, $C \cap [A, X] = \emptyset$ and $M \cap C = \emptyset$ imply that $M' \cap C \neq \emptyset$. Let H_C be an $(n-3)$-transversal that we obtain by moving H tangentially on the boundary of C until it loses contact with A. Now H_C strictly separates A and X and meets C. Since $C \cap M' \neq \emptyset$, Lemma (2.1) implies that all B_i to the right of $[Y, C]$ meet M'. This means that the acute angle μ determined by M and M' is less than $\pi/3$. Similarly, each B_j to the left of $[A, X]$ and disjoint from $[A, X]$ meets both M and M'. As in the proof of Lemma (2.2), $\mu < \frac{\pi}{3}$ implies that there is at most one B_i that meets $[A, X]$ from the left and this B_i meets M or M'. Thus, M is a transversal of $\mathcal{F}\backslash\{C\}$ or M' is a transversal of $\mathcal{F}\backslash\{Y\}$.

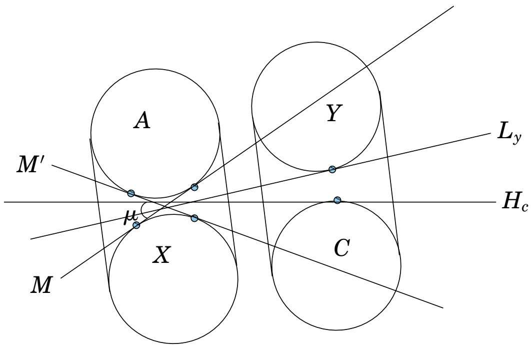


Figure 7. Configuration in Case 1

Due to the symmetry between the relative positions of A, X and C, Y, we may now assume that both $(Y \cup C) \cap [A, X]$ and $(X \cup A) \cup [Y, C]$ are non-empty.

Next, we recall from [8] that there is no directed transversal of $\{A, C, X, Y\}$ that meets the disks in both of the following orders: (a, y, c, x) and (y, a, x, c).

CASE (2). *Y and C meet $[A, X]$ and A meets $[Y, C]$.*

Then $X \cap [Y, C] = \emptyset$. Let N and N' be the tangential separators of Y and C. As with M and M', we assume that $N \cap X \neq \emptyset$ and $N \cap A = \emptyset = N' \cap X$. Recall that $C \cap L = \emptyset$ and that L is an $(n-2)$-transversal of $\mathcal{F}$. Since L_X strictly separates Y and C, all disks disjoint from $[Y, C]$ meet N by Lemma (2.1). It is easy to check that no $B_i \neq A$ meets $[Y, C]$ from the left. There may be some B_i meeting $[Y, C]$ from the right.

Suppose that there exists a B_i that meets $[Y, C]$ from the right and does not meet N. First recall that $B_i \cap (L\backslash[X, Y]) = \emptyset$. But applying $C \cap L = \emptyset$ and $T(4)$ for $\{X, Y, C, B_i\}$, it is easy to check that $B_i \cap L' \neq \emptyset$. This implies that $\lambda < \frac{\pi}{3}$; a contradiction. Thus, N is a transversal of $\mathcal{F}\backslash\{A\}$.

CASE (3). *Y and C meet $[A, X]$ and X meets $[Y, C]$.*

This case can be treated the same way as the previous one by virtue of symmetry. Interchange A by X, Y by C, N by N', L by L', and H by H' where H' is the other tangential separator of A and C. Now, $A \cap [Y, C] = \emptyset$ and we obtain that N' is a transversal of $\mathcal{F}\backslash\{X\}$.

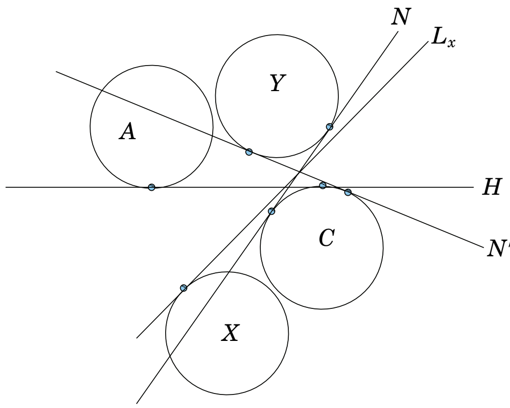


Figure 8. Configuration in Case 2

CASE (4). $C \cap [A, X] = \emptyset$ *and* $Y \cap [A, X] \neq \emptyset$.

Note that H_C strictly separates A and X, and so by Lemma (2.1), all B_i disjoint from $[A, X]$ meet M or M'. Since $M \cap C = \emptyset$, all such B_i meet M'.

We note that any transversal for $\{A, X, C\}$ intersects the sets in the order (a, x, c). Furthermore, $Y \cap M' = \emptyset$ implies that any transversal for $\{A, X, Y\}$ meets the sets in one of the orders (y, a, x) and (a, y, x). Any transversal for $\{A, X, Y, C\}$ thus meets the sets in one of the orders (y, a, x, c) and (a, y, x, c). In case of the former, X and A meet $[C, Y]$ and (interchanging A by C, X by Y, etc.) we argue as in Cases 2 and 3. In case of the latter, $X \cap [Y, C] \neq \emptyset$ and hence, we may assume that $A \cap [C, Y] = \emptyset$, and this implies that $A \cap N' \neq \emptyset$; cf. Figure 9. Since H_A strictly separates C and Y, it follows that all B_i, disjoint from $[C, Y]$, meet N'. In summary:

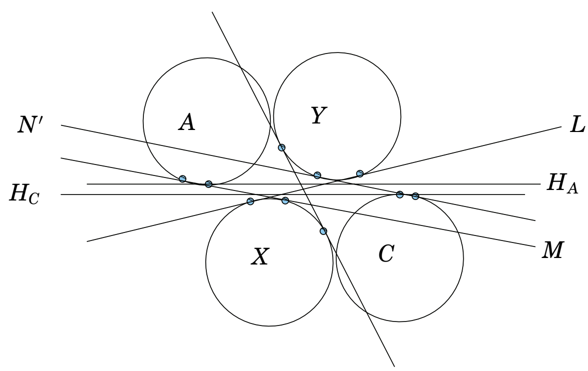


Figure 9. Configuration in Case 4

All B_i, disjoint from $[X, A]$, meet M'. In particular, all B_i to the right of $[Y, C]$ are disjoint from $[A, X]$ and hence meet M'.

All B_i, disjoint from $[C, Y]$ meet N'. In particular, all B_i to the left of $[A, X]$ are disjoint from $[C, Y]$ and hence meet N'.

If $M' \| N'$ then $M'(N')$ is a transversal of $\mathcal{F}\backslash\{Y\}$ ($\mathcal{F}\backslash\{X\}$). Let $M' \not\| N'$. Then $M' \cap N'$ is either to the left of $[A, X]$ or to the right of $[Y, C]$. In case of the former, any B_i to the right of $[Y, C]$ meets M', and hence N'. Thus, N' is a transversal of $\mathcal{F}\backslash\{X\}$. In case of the latter, any B_i to the left of $[A, X]$ meets N', and hence M'. Thus, M' is a transversal of $\mathcal{F}\backslash\{Y\}$.

The remaining case of $C \cap [A, X] \neq \emptyset$ and $Y \cap [A, X] \neq \emptyset$ is handled as above with the interchange of A by C, and X by Y. □

3. The Properties $T(4)$ and $T - 1$

Let $|\mathcal{F}| \leq 9$, and recall ([6] and [7]) that there exist $Z \neq W$ in $\mathcal{F}$ such that $X \cap \left[Z, W\right] \neq \emptyset$ for any $X \in \mathcal{F}' = \mathcal{F}\backslash\{Z, W\}$. Let S_1 and S_2 be the two supporting lines of $\left[Z, W\right]$ that meet both Z and W. Since we wish to verify that $\mathcal{F}$ has the property $T - 1$, we may assume that each S_i meets at least two elements of $\mathcal{F}'$ and that no $X \in \mathcal{F}'$ meets both Z and W.

THEOREM (3.1). ([4]) *If $\mathcal{F}$ has property $T(4)$ and $|\mathcal{F}| \leq 7$ then $\mathcal{F}$ has the property $T - 1$.*

We observe that the proof of Theorem (3.1), in the case $|\mathcal{F}| = 7$, is dependent upon the fact that one of the S_i's (say S_2) meets exactly two elements of $\mathcal{F}'$, and it is easy to check that the arguments are still valid when S_1 meets more than three elements of $\mathcal{F}'$. Accordingly, we assume that $8 \leq |\mathcal{F}| \leq 9$ and that each S_i meets at least three elements of $\mathcal{F}'$. We refer to Figure 10 for the relative positions of the elements of $\mathcal{F}$ with the understanding that G is missing in the case $|\mathcal{F}| = 8$.

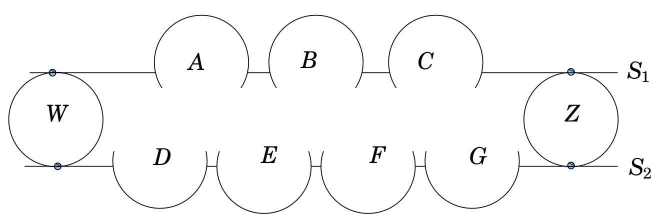


Figure 10.

In the following, we assume that the S_i are horizontal and we let $X_1 X_2 \ldots X_m$ (order unimportant) denote a line meeting the disks $X_1, X_2, \ldots, X_m$. Next, we observe that since S_i meets at least three disks of $\mathcal{F}'$, it is an immediate consequence of $T(4)$ that there is no U, V in $\mathcal{F}'$ such that S_1 meets U, S_2 meets V, one tangential separator of U and V meets Z and not W and the other meets W and not Z; cf. Figure 11. We call this a *forbidden configuration*.

Finally, let K denote the convex hull of the union of the disks in $\mathcal{F}$. We say that $X \in \mathcal{F}$ is a *boundary disk* (of K) if there is a supporting line L of K such that $L \cap K \subset X$. Clearly, W and Z are always boundary disks.

LEMMA (3.2). *Let $U \in \{A, B, C\}$ and $(X, Y) \in \left\{(D, F), \left(E, G\right) \left(D, G\right)\right\}$. Then $UXYWZ$ exists.*

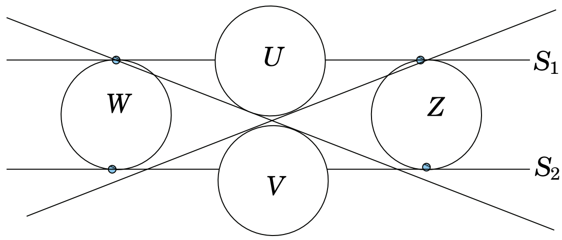


Figure 11.

Proof. By $T(4)$ there exist $L = UXYW$ and $M = UXYZ$. We assume that $L \cap Z = M \cap W = \emptyset$.

If L and M have slopes of the same sign then it is clear that $L \cap M \notin [W, Z]$, and it is easy to determine a line that meets each of U, X, Y, W and Z.

We suppose that there is no $UXYWZ$, and seek a contradiction. By the preceding, all $UXYW$ have slopes of one sign and all $UXYZ$ have slopes of the other sign; furthermore, there exist $L' = UXYW$ and $M' = UXYZ$ that determine a minimum acute angle ψ.

The minimum property of ψ and the fact that there is an element of $\mathcal{F}$ meeting S_2 between X and Y readily yield that $L' \cap M' \in [W, Z]$; cf. Figure 12. It is now easy to check that ψ is minimal only if both L' and M' support both U and a $V \in \{X, Y\}$; that is, U, V, W and Z are in a forbidden configuration. $\square$

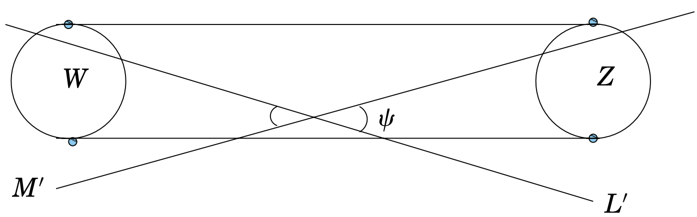


Figure 12.

We note that arguing as in Lemma (3.2) yields also that $ACVWZ$ exists for $V \in \{D, E, F, G\}$.

THEOREM (3.3). *Let $\mathcal{F} = \{A, B, C, D, E, F, W, Z\}$ be a family of eight mutually disjoint unit disks in the plane with the property $T(4)$. Then $\mathcal{F}$ has a 7-transversal.*

Proof. We refer to Figure 10, assume first that B is a boundary disk, and consider $L = BDFWZ$ from Lemma (3.2). It is clear that L meets A or C (otherwise, B does not meet the boundary of K). Let $L = ABDFWZ$. If $L \cap (C \cup E) = \emptyset$ then L is a separating 6-transversal and we apply the Main Lemma.

We assume that neither B nor E is a boundary disk, and note that each of $L = ADFWZ$, $M = BDFWZ$ and $N = CDFWZ$ meets E (otherwise, E is a boundary disk). If L, M and N are not 7-transversals then let $a \in L \cap A$,

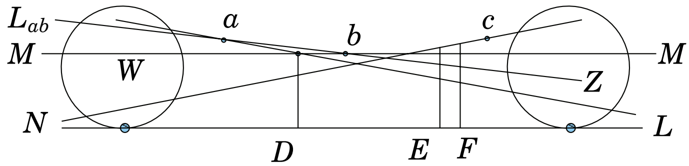


Figure 13.

$b \in M \cap B$, $c \in N \cap C$ and note that (cf. Figure 13) $M = BDEFW$ implies that, say, the line L_{ab}, through a and b, meets E and F. Now if $L_{ab} \cap (C \cup D) = \emptyset$ then L_{ab} is a separating 6-transversal and we apply the Main Lemma. □

THEOREM (3.4). *Let $\mathcal{F} = \{A, B, C, D, E, F, G, W, Z\}$ be a family of nine mutually disjoint unit disks in the plane with the property $T(4)$. Then $\mathcal{F}$ has an 8-transversal.*

Proof. With reference again to Figure 10, let $U \in \{A, B, C\}$. We apply Lemma (3.2) and note that, dependent upon which of D, E, F, G are boundary disks, a 6-transversal $UVXYWZ$ exists for some $V, X, Y \in \{D, E, F, G\}$. Examining such a 6-transversal, say, $L = ADEGWZ$, we observe that if L meets B or C then it is either an 8-transversal or a separating 7-transversal. In view of the Main Lemma, we may thus assume that

i) each $UVXYWZ$ meets only U among A, B and C.

From this and Figure 10, it follows that

ii) there is a $BVXYWZ$ that supports B.

From i) and ii), we obtain now the configuration depicted in Figure 14 with $L_b = BVXYWZ$ for some V, X, Y in $\{D, E, F, G\}$, and T_{ur} depicting the supporting line of $[U, R]$, U and R that meets both W and Z.

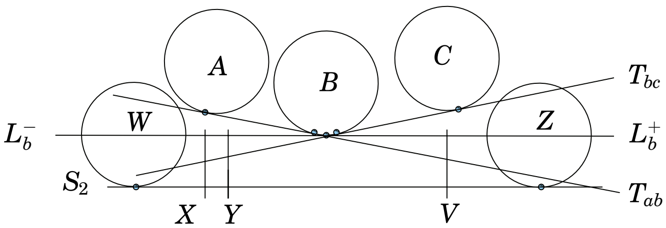


Figure 14.

We note that if T_{ab} or T_{bc} meets three of D, E, F, G then it is either an 8-transversal or a separating 7-transversal. Hence, we may assume that

iii) both T_{ab} and T_{bc} meet at most two of D, E, F, G.

Let L_b^- (L_b^+) denote the component of $L_b \backslash B$ that meets $W(Z)$. By (iii), each of L_b^- and L_b^+ meets at most two of D, E, F, G, and so, we may assume that L_b^- meets X and Y, and L_b^+ meets V. Then T_{bc} meets X and Y, T_{ab} meets V, T_{bc} is disjoint from V, and T_{ab} is disjoint from X or Y.

Let T_{xv} denote the supporting line of $[X, V]$, X and V with the property that $T_{xv} \cap \left[X, V\right] \subset \left[W, Z\right]$. Since $T_{bc} \cap V = \emptyset$, it follows that $V \cap \left[A, C\right] = \emptyset$ and

iv) if $T_{ab} \cap X = \emptyset$ then T_{xv} meets W and Z, and $X \cap \left[A, C\right] = \emptyset$.

Since $\{A, C, X, V\}$ has a transversal, it follows that

v) if $X \cap \left[A, C\right] = \emptyset$ then A and C (and hence B) meet $\left[X, V\right]$.

Since X meets L_b^-, it follows that

vi) if $X \cap \left[A, C\right] \neq \emptyset$ then T_{xv} meets B and C.

We note that iv) and v) imply that $T_{xv} = XVWZABC$. Similarly, we define T_{yv} and obtain that $T_{yv} = YVWZABC$ if Y satisfies iv).

It is clear that if $T_{ab} \cap (X \cup Y) = \emptyset$ then either T_{xv} meets Y or T_{yv} meets X, and there is an 8-transversal. It is also clear that if T_{ab} is disjoint from Y and meets X then T_{yv} meets X as well. Thus, we may assume that T_{ab} is disjoint from X and meets Y. Then $T_{xv} = XVWZABC$, and we may also assume that T_{xv} is disjoint from Y. It is clear that $T_{xv} \cap Y = \emptyset$ and $T_{ab} \cap Y \neq \emptyset$ imply that T_{yv} meets X, W and Z. Thus, either $T_{yv} = YVXWZABC$ by v) or $T_{yv} = YVXWZBC$ by vi). In the latter case, either T_{yv} separates A and the element of $\{D, E, F, G\} \setminus \{X, Y, V\}$ or it is an 8-transversal. $\square$

Received November 22, 2004

Final version received January 23, 2006

T. Bisztriczky
Department of Mathematics and Statistics
University of Calgary
Calgary
Canada
tbisztri@math.ucalgary.ca

F. Fodor
Department of Geometry
Bolyai Institute
University of Szeged
Szeged
fodorf@math.u-szeged.hu

D. Oliveros
Instituto de Matemáticas
Universidad Nacional Autónoma de México
Circuito Exterior
Ciudad Universitaria
04510 México, D.F.
México
deborah@math.ucalgary.ca

References

[1] B. Aronov, J.E. Goodman, R. Pollack, R. Wenger, *On the Helly number for hyperplane transversals to unit balls*, Discrete Comput. Geom. **24** (2–3), (2000), 171–176.

[2] A. HEPPES, *New upper bound on the narrowest transversal strip of $T(3)$-families of disks*, Discrete Comput. Geom., **34** (3), (2005), 463–474.

[3] A. HEPPES, *The width of the transversal strips of $T(3)$-families in the plane*, Discrete Comput. Geom. **34** (3), (2005), 455–461.

[4] T. BISZTRICZKY, F. FODOR AND D. OLIVEROS, *Large transversals to small families of unit disks*, Acta Math. Hungar. **106** (2005), 285–291.

[5] L. DANZER, B. GRÜNBAUM, AND V. KLEE, *Helly's theorem and its relatives*, in Convexity (V. Klee ed.), Proceedings of Symposia in Pure Mathematics, vol. 7, Amer. Math. Soc., Providence, RI, 1963, pp. 100–181.

[6] J. ECKHOFF, *Transversalenprobleme in der Ebene*, Arch. Math. (Basel) **24** (1973), 195–202.

[7] J. ECKHOFF. *Helly, Radon and Caratheodory type theorems*, in Handbook of Convex Geometry (P.M. Gruber and J.M. Wills, eds.), North Holland, Amsterdam, 1993.

[8] H. TVERBERG, *On geometric permutations and the Katchalski-Lewis conjecture on partial transversals for translates*, in Discrete and Computational Geometry (J. E. Goodman, R. Pollack, and W. Steiger, eds.), DIMACS Series in Discrete Mathematics and Theoretical Computer Science, vol. 6, Amer. Math. Soc., Providence, RI, 1991, pp. 351–361.

Bol. Soc. Mat. Mexicana (3) Vol. 12, 2006

ON THE STABLE NORM OF SURFACES

OSVALDO OSUNA

ABSTRACT. Given a Riemmanian metric on a closed surface, we consider the stable norm associated with the metric on the real homology group, and study some consequences when the operator of the Poincaré duality is an isometry respect to the stable norm. In analogy with the marked length spectrum problem we study, if two metrics with the same stable norm are isometric.

1. Introduction and results

The stable norm has attracted great attention in recent years, and has been extensively used in dynamics, geometry and analysis see [1], [2], [6]. Let (M, g) be a closed, oriented manifold equipped with a smooth Riemannian metric. We define a function called the stable norm on $H_1(M, \mathbb{R})$ by

$$|h|_s := \inf\{ \sum |r_i|\ l(\delta_i) \}, \quad \forall h \in H_1(M, \mathbb{R}),$$

where δ_i are simplexes, $r_i \in \mathbb{R}$, $\sum r_i\delta_i$ is a cycle representing h, and $l(\delta_i)$ is the length element induced by the metric g. This function is a norm on $H_1(M, \mathbb{R})$, and by duality it induces a stable norm on $H^1(M, \mathbb{R})$. It is clear that the stable norm should have some relevance in the study of the relationship between the dynamics of the geodesic flow and the geometry of g, see for instance [1] for some connections of the stable norm and minimizing measures of geodesic flows. An alternative definition of the stable norm was given by H. Federer in [4]. It is based on the notion of mass of a Lipschitz current. The stable norm of $h \in H_1(M, \mathbb{R})$ is then the minimal mass of a Lipschitz current in the homology class h. Let dx be the normalized Riemannian measure. From now on we suppose that $\mathrm{vol}(M) = 1$. If $\omega \in H^1(M, \mathbb{R})$, we denote by the same symbol the unique harmonic form in the cohomology class ω. It is well known that the integral $\{\int_M \omega \bigwedge *\omega\}^{\frac{1}{2}}$ (where $*$ denotes the Hodge operator) defines a L^2-norm on the space of differential one-forms on M, and an euclidean norm on $H^1(M, \mathbb{R})$. Recall that a norm on a vector space is called a *euclidean norm* if it is induced by an inner product.

In the case that M is a surface, the Poincaré duality defines an operator $P\colon H_1(M, \mathbb{R}) \to H^1(M, \mathbb{R})$, the so-called operator of Poincaré duality. The duality between homology and cohomology transfers the euclidean norm to $H_1(M, \mathbb{R})$ (we will use the notation $\|\cdot\|_2$ for all of them). We know that the operator of Poincaré duality is a linear isomorphism and an L^2-isometry. From now on we assume that M is a closed, oriented surface of genus ≥ 1 (note that if the genus is 0, then $H_1(M, \mathbb{R}) = H^1(M, \mathbb{R}) = \{0\}$, and $P \equiv 0$). In this work

2000 *Mathematics Subject Classification*: 53D25.

Keywords and phrases: Poincaré duality, stable norm, conjugate points, Lipschitz currents, marked length spectrum.

we relate some geometric aspects of certain operators, and the geometry of the stable norm with the underlying geometry of the manifolds. In particular, our goal is to study the situation when the operator of Poincaré duality is an isometry with respect to the stable norm. Our main result in this direction is the following.

THEOREM (1.1). *Let (M, g) be a closed, oriented, Riemannian surface of genus ≥ 1. If the operator of Poincaré duality is an isometry with respect to the stable norm, then (M, g) is a flat torus.*

The proof of Theorem (1.1) is based on some lemmas that are interesting on their own. In the last part of the paper we address the ridigity problem for the stable norm.

2. Preliminaries and Proofs

Let (N, g) be a closed, oriented manifold equipped with a smooth Riemannian metric. Given an one-form ω on N, its *comass* is defined as

$$\operatorname{comass}(\omega) = \sup_{x\in N} \sup_{v\in T_xN} \left\{ \frac{\omega(v)}{g(v, v)^{\frac{1}{2}}} \right\}.$$

The comass defines a norm on the on the space of one-forms on N. Thus we have induced a dual norm on the space of the one-currents, the so-called *mass*; more precisely, given a one-current γ, consider

$$\operatorname{mass}(\gamma) = \sup\{ \langle \omega, \gamma\rangle \mid \operatorname{comass}(\omega) \leq 1 \}.$$

Recall that the complex of Lipschitz currents on N is dual to the complex of smooth differential forms, and its homology is isomorphic to the real homology of N (see [4]). Now taking the infimum of mass(γ) over the currents in a homology class $[\gamma]$ we obtain a norm on $H_1(N, \mathbb{R})$ which is equal to the stable norm [4]. We will use the notation $|\cdot|_s$ for all of them (in homology and cohomology).

Before going to the proof of the Theorem (1.1), we will establish some results.

LEMMA (2.1). *If the operator of Poincaré duality is an $|\cdot|_s$-isometry, then*

$$\| h \|_2 = |h|_s \quad \textit{for all } h \in H_1(M, \mathbb{R}).$$

In particular the stable norm is euclidean.

Proof. Given $\omega \in H^1(M, \mathbb{R})$, it is well known that $\| \omega \|_2 \leq |\omega|_s$, for instance see [6], [11]. Let $P\colon H_1(M, \mathbb{R}) \to H^1(M, \mathbb{R})$ be the operator of Poincaré duality, thus $\| Ph \|_2 \leq |Ph|_s$ for all $h \in H_1(M, \mathbb{R})$. On the other hand, since P is an L^2-isometry, we have $\| h \|_2 = \| Ph \|_2$. Now we also suppose that P is an $|\cdot|_s$-isometry, therefore

$$\| h \|_2 = \| Ph \|_2 \leq |Ph|_s = |h|_s$$

for all $h \in H_1(M, \mathbb{R})$.

On the other hand, given ω a differential one-form, for $x \in M$ denote by $\| \omega_x \|$ the norm of the linear form on T_xM induced by g. Then we have

$\omega \wedge *\omega(x) = \| \omega_x \|^2 \, dx$ where dx is the normalized measure associated with g. By definition of comass we have

$$\| \omega \|_2^2 = \int_M \| \omega_x \|^2 \, dx \leq \int_M \sup_{x \in M} \| \omega_x \|^2 \, dx = (\text{comass}(\omega))^2$$

since the L^2-norm on the space of one-forms induces the dual L^2-norm on the space of Lipschitz one-currents. Dualizing the preceding inequality, we obtain

$$\text{mass}(\gamma) \leq \| \gamma \|_2 \, .$$

Minimizing both sides of above equation over the currents in a homology class h, we obtain

$$\| h \|_2 \geq |h|_s \quad \text{for all } h \in H_1(M, \mathbb{R}).$$

Now the result follows combining both inequalities. □

Mather in [8] introduces an interesting convex function $\alpha\colon H^1(M, \mathbb{R}) \to \mathbb{R}$. It is well-known that the function α and the stable norm in cohomology are related by $\frac{1}{2}|[\omega]|_s^2 = \alpha([\omega])$. The following results gives some information on α.

LEMMA (2.2). *Let (M, g) be a closed, oriented, Riemannian manifold. If the stable norm is euclidean, then the function α is differentiable.*

Proof. Suppose that $|[\omega]|_s = \{ \langle [\omega], [\omega] \rangle \}^{\frac{1}{2}}$, for some inner product $\langle \cdot, \cdot \rangle$, then $\alpha([\omega]) = \frac{1}{2} \langle [\omega], [\omega] \rangle$. Now the directional derivative of α at $[\omega]$ in the direction $[\eta]$ is given by

$$\begin{aligned} \tfrac{\partial \alpha}{\partial \eta}([\omega]) &= \lim_{t \to 0} \tfrac{\alpha([\omega]+t[\eta]) - \alpha([\omega])}{t} \\ &= \tfrac{1}{2} \lim_{t \to 0} \tfrac{\langle [\omega]+t[\eta], [\omega]+t[\eta] \rangle - \langle [\omega], [\omega] \rangle}{t} \\ &= \tfrac{1}{2} \lim_{t \to 0} 2 \langle [\omega], [\eta] \rangle = \langle [\omega], [\eta] \rangle, \end{aligned}$$

thus the directional derivative there exists in all directions and are continuous functions. Therefore α is differentiable at $[\omega]$. □

Note that if the stable norm $| \cdot |_s$ on cohomology is euclidean, then the dual norm, i.e., the stable norm $| \cdot |_s$ on $H_1(M, \mathbb{R})$, also is euclidean. Thus let us remark that the same argument in Lemma (2.2) gives that the function $| \cdot |_s^2$ on homology is differentiable.

We need the following result, that essentially was proved by G. Paternain, by combining Theorem C and corollary 2 in [11].

PROPOSITION (2.3). *Let (N, g) be a closed, Riemannian manifold whose dimension coincides with its first Betti number. If for every cohomology class the L^2 norm coincides with its stable norm, then N is a flat torus.*

Proof of the Theorem (1.1). From Lemma (2.1), if the operator of Poincaré duality is an $| \cdot |_s$-isometry, then the stable norm is a euclidean norm. Now, by Lemma (2.2), the square of the stable norm on $H_1(M, \mathbb{R})$ is a differentiable function. However, Massart in [9] showed that for a compact surface of genus ≥ 2, the unit sphere of the stable norm is not differentiable at each rational point. So the stable norm cannot be euclidean in this case, and in consequence the operator of Poincaré duality cannot be an $| \cdot |_s$-isometry. Now in the case that genus is 1, then $M = T^2$, therefore dim $H^1(M, \mathbb{R}) = 2$. Using Lemma (2.1) we have $\| \omega \|_2 = |\omega|_s$ for all cohomology classes, and proposition (2.3) implies that T^2 is free of conjugate points. □

3. Stable norm and Marked length spectrum

If M is a manifold and g_1, g_2 are two Riemannian metrics, we say that they have the same marked length spectrum if in each homotopy class of closed curves in M, the infimum of g_1-lengths of curves and the infimum of g_2-lengths of curves are the same.

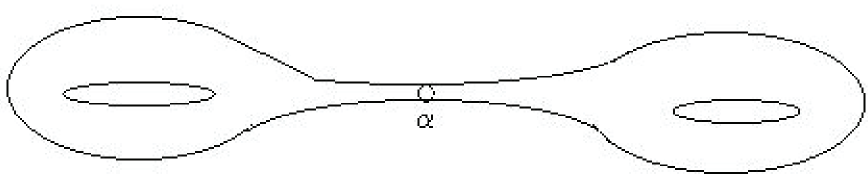
The marked length spectrum problem in general is, to show that two metrics with the same marked length spectrum are isometric. Of course, this cannot hold for arbitrary metrics (for example, if they are allowed to have conjugate points). If the metrics do not have conjugate points, then this problem has a solution as the rigidity theorem in [5], [10] shows. The strongest of these results (see [5]) states that

THEOREM (3.1). *Let M be a closed surface and g_1, g_2 two Riemannian metrics on M with g_1 having non-positive curvature and g_2 without conjugate points. If g_1 and g_2 have the same marked length spectrum then they are isometric by an isometry homotopic to the identity.*

Now inspired by the above results we can ask: given two metrics with the same stable norm, are they isometric? Unfortunately the next proposition show that this is not necesarily true. However, we obtain some affirmative results in the case of the n-torus.

PROPOSITION (3.2). *There exist two metrics g, g_1 on a surface of genus two with the same stable norm such that g has constant curvature -1, g_1 with negative curvature but g, g_1 are not isometric.*

Proof. We consider a surface of genus two, and take a metric g with constant curvature equal to -1 on it. We suppose that this surface has a thin neck between its two holes.



Figure

From the notion of stable norm for a homology class h given in the introduction, we need to consider cycles representing h, whose length is the least possible. If the length of neck of this surface (see figure) is sufficiently long with respect to its diameter, then any closed curve that crosses the neck will be longer than the closed curves around the holes of the surface. Thus for any curve crossing the neck and representing a homology class, we can always find other cycles which do not cross the neck with shorter length and representing this homology class. The idea is that the stable norm can be written as a sum of stable norms of the surfaces with boundary obtained by cutting the surface along α. Now, we consider a neighborhood of the closed curve α and on this neighborhood we make a smooth perturbation of the metric g, obtaining a new metric g_1 with negative curvature but not isometric to g. Again it is not difficult see that for any curve crossing the neck and representing a homology class, we can always find other cycles which do not cross the neck with shorter length respect to g_1 and representing this homology class. Thus g, g_1 have the same stable norm. $\square$

As was commented, there are some affirmative results for the question of rigidity of the stable norm in the case of the n-torus. First we consider the following lemma.

LEMMA (3.3). *Given two flat metrics g_1, g_2 on an n-torus T^n with the same stable norm, then (T^n, g_1) and (T^n, g_2) are isometric by an isometry homotopic to the identity.*

Proof. Denote by $\widetilde{g}_1, \widetilde{g}_2$ two metrics on $\mathbb{R}^n$ such that the proyection $p\colon \mathbb{R}^n \to T^n$ is a Riemannian covering of g_1, g_2 respectively. Given $v \in \mathbb{R}^n$, since $H_1(T^n, \mathbb{R}) = \mathbb{R}^n$ we can think of $v \in H_1(T^n, \mathbb{R})$. Using v we construct a vector field X_v on T^n and ω_v the dual 1 form of X_v. Since (T^n, g_1) is flat, then $|v|_s = ||v||_2$ and a calculation gives that $||v||_2^2 = g_1(X_v, X_v) = \widetilde{g}_1(v, v)$. The same is valid for g_2, but by hypothesis the stable norms are equal. Then $\widetilde{g}_1$ and $\widetilde{g}_2$ are isometrics, therefore g_1, g_2 are isometric as we wished to show. $\square$

The next result gives a positive answer to the rigidity problem in the case of T^2.

PROPOSITION (3.4). *Given two metrics g_1, g_2 on T^2 with the same stable norm, if (T^2, g_1) is flat then g_1, g_2 are isometric by an isometry homotopic to the identity.*

Proof. Since the metrics g_1, g_2 have the same stable norm, the operator of Poincaré duality has the same norm with respect to these metrics, but as there is an isometry respect to the stable norm of the flat metric g_1, this is an isometry respect to stable norm associated with g_2. Now, from theorem (1.1), (T^2, g_2) is flat, thus the result follows from the previous lemma. $\square$

Motivated by these results it seems interesting to solve the following question:

Question: Suppose that a Riemannian metric g on the n-torus T^n, $n \geq 3$ has the same stable norm as a flat metric g_0 on T^n. Then are (T^n, g) and (T^n, g_0) isometric by an isometry homotopic to the identity?

Acknowledgement

The author was partially supported by CIC of U.M.S.N.H., grant 4.18.

Received May 09, 2005

Final version received September 26, 2005

INSTITUTO DE FÍSICA Y MATEMÁTICAS
UNIVERSIDAD MICHOACANA
EDIF. C-3, CIUDAD UNIVERSITARIA
58040 MORELIA, MICH.
MÉXICO
osvaldo@ifm.umich.mx

REFERENCES

[1] V. BANGERT, *Minimal measures and minimizing closed normal one-currents*, *GAFA*, **Vol. 9**, (1999) 413–427.

[2] BURAGO D. YU., IVANOV S., *Riemannian tori without conjugate points are flat*, *GAFA*, **4**, (1994) 3, 259–269.

[3] BURAGO D. YU., IVANOV S. AND KLEINER B., *On the structure of the stable norm of periodic metrics*, Math. Res. Lett. **4**, (1997), 791–808.

[4] H. FEDERER, Geometric Measure Theory, Springer-Verlag, New York (1996). 676 p.

[5] C. CROKE, A. FATHI., J. FELDMAN, *The marked length-spectrum of a surface of nonpositive curvature*, *Topology*, **31** (1992), 847–855.

[6] E. GUTKIN AND D. MASSART, *Stable norm and L^2 norm on the homology of surfaces*, Comunicación técnica CIMAT, No. I-02-02, (2002).

[7] E. HOPF, *Closed surfaces without conjugate points*, Proc. Natl. Acad. Sci. USA **34**, (1948), 47–51.

[8] J. MATHER, *Action minimizing measures for positive definite Lagrangian Systems*, Math. Z., **201**, (1991), 169–207.

[9] D. MASSART, Stable norms of surfaces: local structure of the unit ball of rational directions, *Geom. Funct. Anal.* **7**, (1997), 996–1010.

[10] J. P. OTAL, *Le spectre marqué des longueurs des surfaces a courbure negative*, Ann. of Math, **131**, (1990), 151–162.

[11] G. PATERNAIN, *Schrodinger operators with magnetic fields and minimal action functional*, Israel J. Math., **123**, (2001), 1–2.

Bol. Soc. Mat. Mexicana (3) Vol. 12, 2006

A GENERALIZATION OF THE EQUALITY $p(X) = a(C_p(X))$

ÇETIN VURAL

ABSTRACT. Suppose that α is a family of subsets of a Tychonoff space X and m is an infinite cardinal. We define two new cardinal invariants $p_m^\alpha(X)$ and $a_m(X)$ which are generalizations of the point-finite cellularity $p(X)$ of the space X and the Alexandroff number $a(X)$ of the space X, respectively. Then we obtain the equality $m \cdot p_m^\alpha(X) = m \cdot a_m(C_\alpha(X))$. This result implies that $p^\alpha(X) = a(C_\alpha(X))$ and generalizes the equality $p(X) = a(C_p(X))$ established by Tkachuk in [1].

1. Introduction and Preliminaries

Throughout this paper, X is an infinite Tychonoff space, $\mathbb{R}$ is the real line with the usual topology and m is a fixed infinite cardinal.

Let $C(X)$ denote the set of all continuous real-valued functions on X. The notation $[A, B]$, where $A \subseteq X$ and $B \subseteq \mathbb{R}$, is defined by $[A, B] = \{f \in C(X) : f(A) \subseteq B\}$. Let α be a family of subsets of X. The family $\{[A, V] : A \in \alpha$ and V is open in $\mathbb{R}\}$ is a subbase for a topology on $C(X)$. The function space $C(X)$ having this topology is denoted by $C_\alpha(X)$.

The cardinality of a set A is denoted by card(A). For a real number r, $|r|$ denotes the absolute value of r. As usual $\aleph_0$ denotes the first infinite cardinal.

Let $T(X)$ be the set of all nonempty open subsets of a topological space X, and for an arbitrary cardinal τ let A_τ be the Alexandroff compactification of a discrete space of cardinality τ. Recall that *the point-finite cellularity* $p(X)$ of a space is defined as

$$p(X) = \sup\{\tau : \text{there is a point-finite family } \mathcal{V} \subseteq T(X) \text{ such that } \operatorname{card}(\mathcal{V}) = \tau\},$$

and recall that *the Alexandroff number* $a(X)$ of a space X is defined as

$$a(X) = \sup\{\tau : X \text{ has a subspace which is homeomorphic to } A_\tau\}.$$

Let $C_p(X)$ denote the space of continuous real-valued functions on X equipped with the topology of pointwise convergence. In [1, 2], Tkachuk proved that the point-finite cellularity of the space X is equal to the Alexandroff number of the space $C_p(X)$. One may ask whether the equality holds for the space $C_\alpha(X)$ instead of the space $C_p(X)$, as well. In this paper, we will show that $p^\alpha(X) = a(C_\alpha(X))$, where $p^\alpha(X)$ is the α-finite cellularity of the space X. (We define this in Section 2.)

2000 *Mathematics Subject Classification*: Primary 54A25; Secondary 54C35, 54B10.
Keywords and phrases: Cellularity, point-finite, Alexandroff number.

2. Two new cardinal invariants

Let α and β be two families of subsets of a set Y. We say that the family β is α-$< m$ if for every $A \in \alpha$, $\text{card}(\{B \in \beta : A \cap B \neq \emptyset\}) < m$. If $m = \aleph_0$, we say that the family β is *α-finite*.

Let α be a family of subsets of a space Y. For an arbitrary cardinal τ, we write $Y \in P_m^\alpha(\tau)$ if there is an α-$< m$ family $\mathcal{V} \subseteq T(Y)$ such that $\text{card}(\mathcal{V}) = \tau$.

Definition (2.1). Let Y be a topological space and let α be a family of subsets of Y. *The α-$< m$ cellularity* $p_m^\alpha(Y)$ of the space Y is defined to be the supremum of the cardinals τ such that $Y \in P_m^\alpha(\tau)$.

The expression $p_{\aleph_0}^\alpha(Y)$ is abbreviated as $p^\alpha(Y)$ and called *the α-finite cellularity* of the space Y.

Let A be a discrete space of cardinality τ, and let Ω be an object not in A. Let $A_m(\tau)$ be the set $A \cup \{\Omega\}$ with the following topology: open sets in $A_m(\tau)$ are sets of the form $\{\Omega\} \cup (A \backslash F)$, where $F \subseteq A$ and $|F| < m$, together with all subsets of A. It is clear that $A_{\aleph_0}(\tau) = A_\tau$.

We define a class $Q_m(\tau)$ of spaces as follows: $Y \in Q_m(\tau)$ if and only if there is a continuous one-to-one mapping $\varphi : A_m(\tau) \longrightarrow Y$.

Definition (2.2). Given a topological space Y, the generalized Alexandroff number $a_m(Y)$ of the space Y is defined to be the supremum of the cardinals τ such that $Y \in Q_m(\tau)$.

Observe that compactness of the space A_τ implies that $a_{\aleph_0}(Y) = a(Y)$.

THEOREM (2.3). *Let m and X be as in the introduction and let α be any cover of X. Then we have $m \cdot p_m^\alpha(X) = m \cdot a_m(C_\alpha(X))$.*

Proof. Since $\tau = \sup\{\sigma : \sigma \leq \tau \text{ and } \sigma \text{ is a regular cardinal}\}$ for any cardinal τ, it is sufficient to see that $X \in P_m^\alpha(\sigma) \Leftrightarrow C_\alpha(X) \in Q_m(\sigma)$ for each regular cardinal σ with $m < \sigma$.

Let σ be a regular cardinal with $m < \sigma$, and let $X \in P_m^\alpha(\sigma)$. From the definition of $P_m^\alpha(\sigma)$, there exists an α-$< m$ family $\mathcal{V} = \{V_i : i \in I\} \subseteq T(X)$ such that $\text{card}(I) = \sigma$. Choose an $x_i \in V_i$ for each $i \in I$. Since the space X is Tychonoff, there exists a continuous mapping f_i from X to $[0, 1]$ such that $f_i(x_i) = 1$ and $f_i(X \backslash V_i) \subseteq \{0\}$ for each $i \in I$. Let c_0 denote the constant zero function from X to $\mathbb{R}$. Define a function $\Psi : I \cup \{\Omega\} \longrightarrow C_\alpha(X)$ by $\Psi(\Omega) = c_0$ and $\Psi(i) = f_i$ for every $i \in I$. We claim that there exists a subset J of I such that $\text{card}(J) = \sigma$ and the restriction of the mapping Ψ to the set $J \cup \{\Omega\}$ is one-to-one. To prove this, we shall show first that $\text{card}\left(\Psi^{-1}(\Psi(i))\right) < \sigma$ for every $i \in I$. Since the family $\mathcal{V}$ is α-$< m$ and α is a cover of the space X, we have $\text{card}(\left\{j \in I : x_i \in V_j\right\}) < m$ for every $i \in I$. It is clear that $\left\{j \in I : f_j = f_i\right\} \subseteq \left\{j \in I : x_i \in V_j\right\}$ so $m < \sigma$ and the equality $\Psi^{-1}(\Psi(i)) = \left\{j \in I : f_j = f_i\right\}$ lead us to the fact that $\text{card}\left(\Psi^{-1}(\Psi(i))\right) < \sigma$ for every $i \in I$. Let $Z = \{\Psi(i) : i \in I\}$. $I = \cup\left\{\Psi^{-1}(g) : g \in Z\right\}$, $\text{card}\left(\Psi^{-1}(g)\right) < \sigma$ for each $g \in Z$, $\text{card}(I) = \sigma$ and regularity of σ lead us to the fact that $\text{card}(Z) = \sigma$. Let us choose for every $g \in Z$ a point $i_g \in \Psi^{-1}(g)$; obviously, the cardinality of the set $J = \{i_g : g \in Z\}$ is σ. It is clear that the restriction of the mapping Ψ to the set $J \cup \{\Omega\}$, say

φ, is one-to-one. Now, we shall prove that the mapping φ is continuous. Since J is a discrete space, it is sufficient to see that φ is continuous at the point Ω. Let $A \in \alpha$ and let W be an open subset of $\mathbb{R}$ such that $c_0 \in [A, W]$. Let $F = \{j \in J : A \cap V_j \neq \emptyset\}$. Since the family $\mathcal{V}$ is α-$< m$, we have $\mathrm{card}(F) < m$, and so the set $(J \backslash F) \cup \{\Omega\}$ is an open neighbourhood of Ω. It can be easily seen that $\varphi\left((J \backslash F) \cup \{\Omega\}\right) \subseteq [A, W]$. Therefore φ is continuous. Hence $C_\alpha(X) \in Q_m(\sigma)$.

Now, let $C_\alpha(X) \in Q_m(\sigma)$ for some regular cardinal σ with $m < \sigma$. By the definition of the class $Q_m(\sigma)$, there exists a continuous one-to-one mapping $\psi : A_m(\sigma) = I \cup \{\Omega\} \longrightarrow C_\alpha(X)$. Without loss of generality, we may assume that $\psi(\Omega) = c_0$. For each $i \in I$ and $n \in \omega$, define the set $U_{in} = \left\{x \in X : |\psi(i)(x)| > \frac{1}{n}\right\}$. It is clear that the set U_{in} is open in X for each $i \in I$ and $n \in \omega$. We claim that the family $\{U_{in} : i \in I\}$ is α-$< m$ for each $n \in \omega$. Let $A \in \alpha$. Since $\psi(\Omega) = c_0 \in \left[A, \left(-\frac{1}{n}, \frac{1}{n}\right)\right]$, $\left[A, \left(-\frac{1}{n}, \frac{1}{n}\right)\right]$ is open in $C_\alpha(X)$ and ψ is continuous, there is a $F_n \subseteq I$ such that $\mathrm{card}(F_n) < m$ and $\psi\left((I \backslash F_n) \cup \{\Omega\}\right) \subseteq \left[A, \left(-\frac{1}{n}, \frac{1}{n}\right)\right]$. Then, $A \cap U_{in} = \emptyset$ for every $i \in I \backslash F_n$. Therefore for each $n \in \omega$, the family $\{U_{in} : i \in I\}$ is α-$< m$. Let $S_n = \{i \in I : U_{in} \neq \emptyset\}$ for every $n \in \omega$. We claim that there is a $k \in \omega$ such that $\mathrm{card}(S_k) = \sigma$. To see this, suppose that $\mathrm{card}(S_n) < \sigma$ for every $n \in \omega$, and so $\mathrm{card}\left(\bigcup\{S_n : n \in \omega\}\right) < \sigma$. Since m is an infinite cardinal and $m < \sigma$, σ is an uncountable cardinal. As σ is a regular uncountable cardinal, there is an $i_0 \in I \backslash \bigcup\{S_n : n \in \omega\}$. This implies that $|\psi(i_0)(x)| \leq \frac{1}{n}$ for every $n \in \omega$ and $x \in X$. Therefore $\psi(i_0) = c_0$. But this contradicts the fact that ψ is one-to-one. Hence there is a $k \in \omega$ such that $\mathrm{card}(S_k) = \sigma$. Since the family $\{U_{ik} : i \in I\}$ is α-$< m$, we have $\mathrm{card}\left(\{U_{ik} : i \in S_k\}\right) = \sigma$, and so $X \in P_m^\alpha(\sigma)$. □

If $m = \aleph_0$ in the above theorem, the α-finite cellularity of the space X is equal to the Alexandroff number of the space $C_\alpha(X)$, i.e., we have the equality $p^\alpha(X) = a\left(C_\alpha(X)\right)$.

If $m = \aleph_0$ and α is the family of all nonempty finite subsets of X in the above theorem, we obtain the following equality which was established by Tkachuk in [1, 2].

Corollary (2.4). $p(X) = a\left(C_p(X)\right)$.

Observe that if X belongs to α in the above theorem then $p_m^\alpha(X) \leq m$ and $a_m(C_\alpha(X)) \leq m$. In particular, if α is the family of all non-empty closed subsets of X then $p_m^\alpha(X) \leq m$; if, in addition, X is compact and $C_k(X)$ is the space $C(X)$ with the compact-open topology, then $a_m(C_k(X)) \leq m$.

Received June 30, 2005

Final version received October 17, 2005

Gazi University
Department of Mathematics
06500 Ankara
Turkey
cvural@gazi.edu.tr

REFERENCES

[1] V. V. TKACHUK, *On cardinal invariants of Suslin number type*, Soviet Math. Dokl. **27**, (1983), 681–684.

[2] V. V. TKACHUK, *On a supertopological cardinal invariant*, Vestnik Mosk. Univ. Matematika **39** (4), (1984), 26–29.

Bol. Soc. Mat. Mexicana (3) Vol. 12, 2006

STABLE GEOMETRIC DIMENSION OF VECTOR BUNDLES OVER ODD-DIMENSIONAL REAL PROJECTIVE SPACES

MARTIN BENDERSKY AND DONALD M. DAVIS

ABSTRACT. In [6], the geometric dimension of all stable vector bundles over real projective space P^n was determined if n is even and sufficiently large with respect to the order 2^e of the bundle in $\widetilde{KO}(P^n)$. Here we perform a similar determination when n is odd and $e > 6$. The work is more delicate since P^n does not admit a v_1-map when n is odd. There are a few extreme cases which we are unable to settle precisely.

1. Statement of results

The geometric dimension $\mathrm{gd}(\theta)$ of a stable vector bundle θ over a space X is the smallest integer m such that θ is stably equivalent to an m-plane bundle. Equivalently, $\mathrm{gd}(\theta)$ is the smallest m such that the classifying map $X \xrightarrow{\theta} BO$ factors through $BO(m)$. The group $\widetilde{KO}(P^n)$ of equivalence classes of stable vector bundles over real projective space is a finite cyclic 2-group generated by the Hopf line bundle ξ_n.

In [6], it was shown that, for sufficiently large even n, the geometric dimension of a stable vector bundle over P^n depends only on its order in $\widetilde{KO}(P^n)$ and the mod 8 value of n. For bundles of order 2^e, this value, called $\mathrm{sgd}(n, e)$ or $\mathrm{sgd}(\overline{n}, e)$, where $\overline{n}$ is the mod 8 residue of n, was completely determined; its approximate value is $2e$. A key role in this analysis was played by KO-equivalences $P^{n+8}_{k+8} \to P^n_k$, defined if n is even, k is odd, and $n + 8 < 2k - 1$. Such maps do not exist when n is odd, and so the methods and results are somewhat more complicated. The term "stable" geometric dimension (sgd) refers to the fact that the geometric dimension achieves a stable value as n gets large within its congruence class.

An important role in [6] was played by the v_1-periodic spectrum functor Φ described in [7, 7.2]. We are interested in the stable portion of $[P^n, \Phi BSO(m)]$, i.e. the portion which persists under $j_m : BSO(m) \to BSO$. To achieve this, we define the **stable** portion

$$\mathbf{s}[P^n, \Phi BSO(m)] = [P^n, \Phi BSO(m)]/\ker(j_{m*}),$$

and similarly for spectral sequence groups that approximate these groups. The group $\mathbf{s}[P^n, \Phi BSO(m)]$ is cyclic since it maps injectively to the cyclic group $[P^n, \Phi BSO]$.

2000 *Mathematics Subject Classification*: 55S40, 55R50, 55T15.
Keywords and phrases: geometric dimension, vector bundles, homotopy groups.
We would like to thank Mark Mahowald for valuable conversations related to this work.

In [6], we proved that, if n is even,

$$\operatorname{sgd}(n,e) \le m \quad \text{iff} \quad \nu(\mathbf{s}[P^n, \Phi BSO(m)]) \ge e. \tag{1.1}$$

Here and throughout, $\nu(-)$ denotes the exponent of 2 in an integer, and if C is a cyclic group, then $\nu(C)$ denotes $\nu(|C|)$. The backwards implication has a simple and natural proof ([6, 1.5]), while the forward implication was proved by noting that all the requisite nonlifting results were already in the literature.

For odd n, we determine $\nu(\mathbf{s}[P^n, \Phi BSO(m)])$ completely in Theorem (1.2), provided $m \ge 12$. We prove in (2.1) that the backwards implication of (1.1) holds when n is odd, except that here this sgd refers to stable bundles of order 2^e over projective spaces of sufficiently large dimension $\equiv n \bmod 2^L$, with L usually, but perhaps not always, equal to 3. We will observe in Theorem (1.3) that, in almost all cases, known nonlifting results of Section 3 imply the converse; i.e. (1.1) holds in almost all cases when n is odd. However, there are some rare cases in which our computation of $\nu(\mathbf{s}[P^n, \Phi BSO(m)])$ suggests there should be an extra nonlifting result which we have been unable to establish.

Most of our work is devoted to proving the following theorem.

THEOREM (1.2). *If $m = 8i+d \ge 12$, then $\nu(\mathbf{s}[P^n, \Phi BSO(m)])) = 4i+t$, where t is given by the following table. The two entries indicated by asterisks must be decreased by 1 if $\nu(n+1-m) \ge \frac{1}{2}m - 2$.*

	d							
	0	1	2	3	4	5	6	7
n mod 8: 1	0	0	1^*	1	2	2	3	3
3	0	0	1	2	3	3	3	3
5	0	0	1	1	2	2	3^*	3
7	0	0	0	0	1	1	2	3

Combining this with (2.1) for liftings, and using (3.1) and (3.2) for nonliftings, yields the following result, which is our main theorem.

THEOREM (1.3). *Define $\delta(\overline{n}, e)$ by the table*

	e mod 4			
	0	1	2	3
$\overline{n}$: 1	0	0^*	0	0
3	0	0	-1	-2
5	0	0	0	0^*
7	0	2	2	1

Let $e \ge 7$. For sufficiently large $n \equiv \overline{n} \bmod 8$,[1] *the geometric dimension of stable vector bundles of order 2^e over P^n equals $2e + \delta(\overline{n}, e)$, except that entries indicated with an asterisk might be 1 greater than indicated if $\nu(n+1-2e) \ge e-2$.*

The idea of stable geometric dimension was first proposed in [10]. It was claimed there that if $e \ge 75$, then $\operatorname{sgd}(n,e) \le 2e + \delta(\overline{n}, e)$ with $\delta(\overline{n}, e)$ as in Theorem (1.3), ignoring the asterisks. We do not contradict those results here.

[1] If the asterisked entries are increased to 1, then $n \equiv \overline{n} \bmod 8$ must be modified to $n \equiv \overline{n} \bmod 2^{e-2}$ in these cases.

However, if the exotic nonlifting results mentioned above can be proved, they would contradict this lifting result of [10], for certain extreme cases with n odd. This does not seem to be out of the question, for the sentence near the bottom of [10, p.60] which includes a commutative diagram seems to lack justification, which could render that proof invalid.

For even-dimensional projective spaces, we also obtained, in [6], results about stable geometric dimension for bundles of order 2^e when $e < 7$. We could do that here for odd-dimensional projective spaces, but the arguments are extremely delicate. Consequently, we will defer these cases of small m and e to the future.

2. Proof of Theorem (1.2)

In this section, we prove Theorem (1.2). We begin with a general result similar to [6, 1.6].

PROPOSITION (2.1). *Let n be odd and e a fixed positive integer. For each m, there exists an integer L such that if $\nu(\mathbf{s}[P^n, \Phi BSO(m)]) \geq e$ then, for sufficiently large N satisfying $N \equiv n \mod 2^L$, the geometric dimension of any stable vector bundle of order 2^e over P^N is less than or equal to m.*

Proof. From the definition of ΦX in [12][2] as a periodic spectrum whose spaces are telescopes of

$$\Omega^{L_1} X \to \Omega^{L_1+2^L} X \to \cdots \to \Omega^{L_1+k2^L} X \to \cdots,$$

with $L_1 \equiv 0 \bmod 2^L$ for the 0^{th} space, it follows, using James periodicity, that

$$[P^n, \Phi BSO(m)] \approx \operatorname*{colim}_k [P^{n+k2^L}_{1+k2^L}, BSO(m)].$$

Thus the hypothesis implies that the stable bundle of order 2^e over P^{n+k2^L} lifts to $BSO(m)$ if k is sufficiently large. □

The informal claim that we made in Section 1 that L can usually be chosen to be 3 can be seen either from the fact that $\nu(\mathbf{s}[P^n, BSO(m)])$ determined in (1.2) usually only depends on n mod 8, or by restricting to P^{n-1} and using the result from [6] that geometric dimension over these even-dimensional projective spaces eventually only depends on the mod 8 value of $n-1$. The way in which Proposition (2.1) will be used in the proof of Theorem (1.2) is to use known nonlifting results ((3.1) and (3.2)) to assert that $\nu(\mathbf{s}[P^n, \Phi BSO(m)]) < e$ for various values of the parameters.

The proof of the following result occupies most of the rest of this section.

THEOREM (2.2). *Let n be odd, $m \geq 12$, and $\phi_{n,m}$ denote the restriction homomorphism*

$$\mathbf{s}[P^n, \Phi BSO(m)] \to \mathbf{s}[P^{n-1}, \Phi BSO(m)]$$

[2] called $\mathbf{Tel}_1 X$ there

between cyclic 2-groups. Then

$$|\ker(\phi_{n,m})| = \begin{cases} 2 & \textit{if } n \equiv 1 \mod 8 \\ 1 & \textit{otherwise} \end{cases}$$

$$|\operatorname{coker}(\phi_{n,m})| = \begin{cases} 2 & \textit{if } n \equiv 1 \mod 4 \textit{ and } n-m \equiv 0,1,2 \mod 8 \\ 2 & \textit{if } n \equiv 1 \mod 4 \textit{ and } \nu(n+1-m) \geq m/2-2 \\ 1 & \textit{otherwise} \end{cases}$$

Theorem (1.2) follows directly from (2.2) and the following recapitulation of results of [6].

THEOREM (2.3). ([6, 1.7,1.8,1.10]) *If $n \equiv 6, 8 \mod 8$ and $8i+d \geq 9$, then*

$$\nu(\mathbf{s}[P^n, \Phi BSO(8i+d)]) = 4i + \begin{cases} -1 & d=-1 \\ 0 & d=0,1,2,3 \\ 1 & d=4,5 \\ 2 & d=6. \end{cases}$$

If $n \equiv 2, 4 \mod 8$ and $8i+d \geq 9$, then

$$\nu(\mathbf{s}[P^n, \Phi BSO(8i+d)]) = 4i + \begin{cases} 0 & d=0,1 \\ 1 & d=2 \\ 2 & d=3 \\ 3 & d=4,5,6,7. \end{cases}$$

The lengthy proof of Theorem (2.2) will occupy the remainder of this section. We let $n = 2k+1$. Viewing $\mathbf{s}[P, \Phi BSO(m)]$ as

$$\operatorname{im}([P, \Phi BSO(m)] \xrightarrow{j_{m*}} [P, \Phi BSO],$$

it is clear that the kernel of $\phi_{2k+1,m}$ in (2.2) equals the kernel of

$$[P^{2k+1}, \Phi BSO] \xrightarrow{i^*} [P^{2k}, \Phi BSO].$$

The proof of (2.1) implies that this kernel equals that of

$$\operatorname{colim}[P^{2k+1+c2^L}, BSO] \xrightarrow{i^*} \operatorname{colim}[P^{2k+c2^L}, BSO],$$

which, by the calculation of $\widetilde{KO}(P^n)$ in [1], has order 2 if $k \equiv 0 \bmod 4$, and is trivial otherwise. This establishes the kernel part of (2.2).

The cokernel of $\phi_{2k+1,m}(= \mathbf{s}i^*)$ is much more delicate. It involves the exact sequence

$$[P^{2k+1}, \Phi BSO(m)] \xrightarrow{i^*} [P^{2k}, \Phi BSO(m)] \xrightarrow{\alpha^*} v_1^{-1}\pi_{2k}(BSO(m)), \tag{2.4}$$

where α denotes the attaching map. The following proposition is elementary.

PROPOSITION (2.5). *Let $x \in [P^{2k}, \Phi BSO(m)]$ satisfy $j_{m*}(x) \neq 0$, so its equivalence class $[x]$ is a nonzero element in $\mathbf{s}[P^{2k}, \Phi BSO(m)]$.*

- *If $\alpha^*(x) = 0$, then $[x] \in \operatorname{im}(\phi_{2k+1,m})$.*
- *If $\alpha^*(x) \neq 0$ and there is no $y \in \ker(j_{m*})$ such that $\alpha^*(y) = \alpha^*(x)$, then $[x]$ is a nonzero element of $\operatorname{coker}(\phi_{2k+1,m})$.*

The main point here is the necessity of checking for y.

The proof of the cokernel part of (2.2) varies depending on the mod 4 value of k and mod 8 value of m in (2.4).

Case 1: $k \equiv 2 \mod 4$, $m \equiv -1, 0, 1 \mod 8$. Here $v_1^{-1}\pi_{2k}(BSO(m)) = 0$ by [3, 1.2,3.4,3.6] and so by Proposition (2.5) $\phi_{2k+1,m}$ is surjective in (2.2) in this case.

Case 2: $k \equiv 2 \mod 4$, $m \equiv 3, 4, 5 \mod 8$. By §3[3],

$$\nu(\mathbf{s}[P^{8\ell+5}, \Phi BSO(8i+d)]) \leq 4i + \begin{cases} 1 & d = 3 \\ 2 & d = 4, 5. \end{cases}$$

By Theorem (2.3),

$$\nu(\mathbf{s}[P^{8\ell+4}, \Phi BSO(8i+d)]) = 4i + \begin{cases} 2 & d = 3 \\ 3 & d = 4, 5. \end{cases}$$

Thus $\phi_{2k+1,m}$ in (2.2) must have nontrivial cokernel when $m \equiv 3, 4, 5 \mod 8$ (and still $k \equiv 2 \mod 4$). This cokernel can have order at most 2 because $v_1^{-1}\pi_{2k}(BSO(m)) = \mathbf{Z}/2$ if $m \equiv 3, 5 \mod 8$ by [3, 3.10], while $v_1^{-1}\pi_{2k}(BSO(8i+4)) \approx \mathbf{Z}_2 \oplus \mathbf{Z}_2$.

Case 3: $k \equiv 0 \mod 4$, $m \equiv -1, 0, 1 \mod 8$. By §3,

$$\nu(\mathbf{s}[P^{8\ell+1}, \Phi BSO(8i+d)]) \leq 4i + \begin{cases} -1 & d = -1 \\ 0 & d = 0, 1. \end{cases}$$

By (2.3)

$$\nu(\mathbf{s}[P^{8\ell}, \Phi BSO(8i+d)]) = 4i + \begin{cases} -1 & d = -1 \\ 0 & d = 0, 1. \end{cases}$$

We have already proved $\ker(\phi_{8\ell+1,m}) = \mathbf{Z}/2$, and hence $\mathrm{coker}(\phi_{8\ell+1,m}) \neq 0$. We must prove the order of this cokernel is only 2.

By [3, 1.2,1.3,1.4], $v_1^{-1}\pi_{8\ell-1}(SO(m))$ is an extension of two $\mathbf{Z}/2$-vector spaces[4], one in filtration 2 and the other in filtration 4. We will show that the filtration-4 elements are in the image of α^* in (2.4); they are hit not by the stable summand but rather by elements of order 2. This implies that the desired cokernel has order only 2.

The attaching map for the top cell of $P^{8\ell+1}$ is η on the $(8\ell-1)$-cell. By [6, (2.4)],

$$[P^{8\ell}, \Phi BSO(m)] \approx [P^0_{1-8\ell}, \Phi BSO(m)] \approx [M^0(2^{4\ell}), \Phi BSO(m)].$$

Since, by [6, (2.6)], the stable summand of $[M^0(2^{4\ell}), \Phi BSO(m)]$ comes from the bottom cell of the Moore space, α^* in (2.4) is equivalent to

$$\mu_\ell^* : v_1^{-1}\pi_{-1}(BSO(m)) \to v_1^{-1}\pi_{8\ell}(BSO(m)), \tag{2.6}$$

---

[3] As was remarked prior to Theorem (1.3), all the lower bounds of that theorem are immediate from (3.1) and (3.2), and by (2.1), all the non-asterisked "$\leq$" parts of (1.2) follow from this. When we invoke one of these (sgd$(-,-) \leq -$)-results, we will just say "By §3."

[4] This is the first time of many that we will utilize the isomorphism $v_1^{-1}\pi_i(SO(m)) \approx v_1^{-1}\pi_{i+1}(BSO(m))$.

where μ_ℓ is the element of highest Adams filtration in the $(8\ell+1)$-stem, detected by $P^\ell h_1$ in the Adams spectral sequence. This is seen by observing that

$$S^{8\ell} \xrightarrow{\alpha} P^{8\ell} \xrightarrow{\phi^\ell} P^0_{1-8\ell}$$

and

$$S^{8\ell} \xrightarrow{\mu_\ell} S^{-1} \xrightarrow{\deg 1} P^0_{1-8\ell}$$

become equal in $\pi_{8\ell}(P^0_{1-8\ell} \wedge J) \approx \mathbf{Z}_2 \oplus \mathbf{Z}_2$, where each equals the element of highest filtration. Thus, since $v_1^{-1}\pi_*(P) \approx v_1^{-1}\pi_*(P \wedge J)$ for spectra P by [13], the two composites become equal in $v_1^{-1}\pi_{8\ell}(P^0_{1-8\ell})$. Thus they are equal in $v_1^{-1}\pi_{8\ell}(BSO(m))$. Here we have used the 2-local J-spectrum which is the fiber of $\psi^3 - 1 : bo \to \Sigma^4 bsp$. This spectrum played a key role in the early days of v_1-periodic homotopy theory, especially in [13].

In the spectral sequence of [3] converging to $v_1^{-1}\pi_*(SO(m))$, elements in filtration ≥ 2 occur in eta-towers, with their Pontryagin duals described by elements in $QK^1(\mathrm{Spin}(m))/\operatorname{im}(\psi^2)$, occurring with period 4. Dual to the composition (2.6) is

$$E_2^{s+1,t+2+8\ell}(\mathrm{Spin}(m))^\# \xrightarrow{v_1^{4\ell}} E_2^{s+1,t+2}(\mathrm{Spin}(m))^\# \xrightarrow{h_1^\#} E_2^{s,t}(\mathrm{Spin}(m))^\#, \tag{2.7}$$

where v_1^4 is the isomorphism which shifts eta towers to elements with the same name, and $h_1^\#$ stays in the same eta tower. To see this, note that, with $Y = \mathrm{Spin}(m)$, if $g \in \pi_n(Y)$, then $g \circ \mu_\ell (= \mu_\ell^*(g)$ in (2.6)) can be obtained as the composite

$$S^{8\ell+n+1} \hookrightarrow M^{8\ell+n+2}(2) \xrightarrow{A^\ell} M^{n+2}(2) \xrightarrow{\widetilde{\eta}} S^n \xrightarrow{g} Y, \tag{2.8}$$

where A is an Adams map and $\widetilde{\eta}$ an extension over the mod-2 Moore spectrum of $S^{n+1} \xrightarrow{\eta} S^n$. Then (2.7) is dual to the horizontal composition in Diagram (2.9), while (2.8) induces the composition around the top. The vertical maps ∂ are Bockstein homomorphisms for $\cdot 2$.

DIAGRAM (2.9). *Diagram involving Bockstein and h_1*

$$\begin{tikzcd}
 & E_2^{s,s+n+2}(Y;\mathbf{Z}_2) \arrow[r, "v_1^{4\ell}"] \arrow[d, "\partial"] & E_2^{s,s+n+2+8\ell}(Y;\mathbf{Z}_2) \arrow[d, "\partial"] \\
E_2^{s,s+n}(Y) \arrow[ur, "\widetilde{\eta}^*"] \arrow[r, "h_1"] & E_2^{s+1,s+n+2}(Y) \arrow[r, "v_1^{4\ell}"] & E_2^{s+1,s+n+2+8\ell}(Y)
\end{tikzcd}$$

Now the claim about filtration-4 elements y being $\alpha^*(x)$ with x an element of filtration 3 follows from (2.7), since x is the element in an earlier eta-tower with the same name as y. This completes the proof of Case 3.

For the remaining cases, we will need the following result, where $Q(-)$ denotes the indecomposables.

THEOREM (2.10). *For any positive integers n and m, there is a spectral sequence $E_r(n,m)$ converging to $[P^n, \Phi SO(m)]_*$ with*

$$E_2^{s,t}(n,m) = \mathrm{Ext}^s_{\mathcal{A}}(K^*(\Phi\,\mathrm{Spin}(m)), K^*(\Sigma^t P^n)). \tag{2.11}$$

If n is even, then $E_2^{s,2r}(n,m) = 0$, and if n is also sufficiently large, there is a short exact sequence

$$\begin{aligned} & 0 \to \operatorname{Ext}^s_{\mathcal{A}}(QK^1 \operatorname{Spin}(m)/\operatorname{im}(\psi^2), K^1S^{2r+1}) \to E_2^{s,2r+1}(n,m) \\ & \xrightarrow{\delta} \operatorname{Ext}^{s+1}_{\mathcal{A}}(QK^1 \operatorname{Spin}(m)/\operatorname{im}(\psi^2), K^1S^{2r+1}) \to 0. \end{aligned} \tag{2.12}$$

If n is odd and sufficiently large, there is a split short exact sequence
(2.13)
$$0 \to \operatorname{Ext}^{s,n+t}_{\mathcal{A}}(QK^*(\operatorname{Spin}(m))/\operatorname{im}(\psi^2)) \xrightarrow{q^*} E_2^{s,t}(n,m) \xrightarrow{i^*} E_2^{s,t}(n-1,m) \to 0.$$

Several remarks are in order here. (i) We omit 2-adic coefficients from all $K^*(-)$-groups, and will continue to do so. (ii) $\mathcal{A}$ is the category of 2-adic stable Adams modules.([7]) (iii) We have replaced $SO(m)$ by its double cover $\operatorname{Spin}(m)$. This does not change $v_1^{-1}\pi_*(-)$, and indeed $\Phi SO(m) = \Phi \operatorname{Spin}(m)$. But for calculations such as (2.14), it is essential that the underlying space be simply-connected. (iv) Beginning with (2.13), we will often abbreviate $\operatorname{Ext}^s_{\mathcal{A}}(M, K^*S^t)$ as $\operatorname{Ext}^{s,t}_{\mathcal{A}}(M)$. (v). The splitting of (2.13) is just claimed for E_2, not necessarily for the entire spectral sequence.

Proof. By [7, 7.2], the spectrum $\Phi SO(m)$ is $K/2_*$-local, and so the existence of the spectral sequence follows from [7, 10.4].[5] By [8, 9.1], there is an isomorphism in A

$$K^i(\Phi \operatorname{Spin}(m)) \approx \begin{cases} 0 & i = 0 \\ QK^1(\operatorname{Spin}(m))/\operatorname{im}(\psi^2) & i = 1. \end{cases} \tag{2.14}$$

By [1], if n is even, then

$$K^i(P^n) \approx \begin{cases} \mathbf{Z}/2^{n/2} & i = 0 \\ 0 & i = 1 \end{cases}$$

with $\psi^k = 1$ on $K^0(P^n)$.

Let $M_r = K^*(S^{2r+1}) = \begin{cases} \mathbf{Z}_2^\wedge & * = 1 \\ 0 & * = 0 \end{cases}$ with $\psi^k = k^r$. With n still even, there is a short exact sequence in $\mathcal{A}$

$$0 \to M_r \xrightarrow{2^{n/2}} M_r \to K^*(\Sigma^{2r+1}P^n) \to 0. \tag{2.15}$$

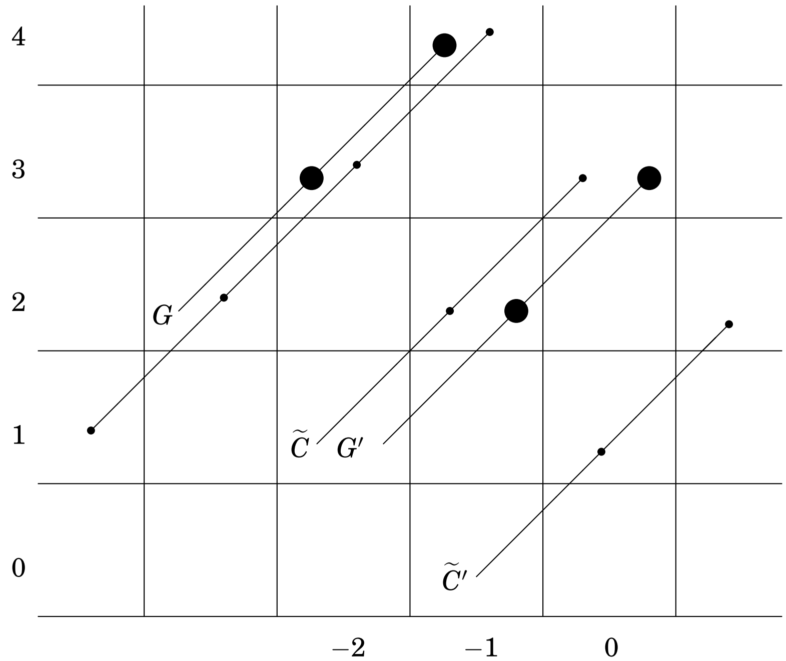
We choose n larger than any of the exponents of Ext groups that occur (roughly $m/2$). Then the long exact sequence with (2.15) in the second variable of $\operatorname{Ext}_{\mathcal{A}}(K^*(\Phi \operatorname{Spin}(m)), -)$ breaks up into short exact sequences (2.12).

If n is odd, the cofibration $P^{n-1} \to P^n \to S^n$ induces a split short exact sequence in $K^*(-)$. In fact, $K^*(S^n)$ and $K^*(P^{n-1})$ are nonzero only in distinct gradings. The split short exact sequence (2.13) is immediate from this. □

By (2.12), if n is even and sufficiently large, the E_2-chart is independent of n, and, using results of [3] about the general form of $\operatorname{Ext}^{**}_{\mathcal{A}}(QK^1 \operatorname{Spin}(m)/\operatorname{im}(\psi^2))$, the chart, in the vicinity of $t-s = -1$, has the form pictured in Diagram (2.16).

[5]Although [7] just deals with odd primes, this result is also valid for the prime 2.

DIAGRAM (2.16). **General form of $E_2^{*,*}(n, m)$ when n is even and large**



The notation here is as follows. As is customary with Adams spectral sequence charts, the group in position $(t-s, s)$ is $E_2^{s,t}$. In [3, esp. 1.3,3.7,3.12], charts for $\mathrm{Ext}_{\mathcal{A}}^{*,*}(QK^1\,\mathrm{Spin}(m))$ are presented for various mod 8 congruences of m. The group $\widetilde{C}$ of Diagram (2.16) is usually[6] a sum of two cyclic groups usually denoted $C_1 \oplus C_2$ in [3]. Our group $\widetilde{C}'$ is a group isomorphic to $\widetilde{C}$ coming from the second half of (2.12). The summand C_1 in $\widetilde{C}'$ is our stable summand $\mathbf{s}E_2^{0,-1}(n, m)$. The groups G and G' have the same order as $\widetilde{C}$, but usually have many more summands; they are also denoted by G in the charts of [3]. The big $\bullet$'s in (2.16) are sums of $\mathbf{Z}_2$'s.

By the proof of [6, 1.7 and 1.10.1], (2.12) splits as spectral sequences, and the stable summand in which we are interested occurs in the summand which comes from δ^{-1}. We may ignore the other summand and, if $n \equiv 6$ or 8 mod 8, think of the spectral sequence for $[P^n, \Phi SO(m)]_*$ as being the spectral sequence for $v_1^{-1}\pi_*(SO(m))$ shifted one unit down and one unit to the right. If $n \equiv 2$ or 4 mod 8, we may think of the spectral sequence for $[P^n, \Phi SO(m)]_*$ as a similar shift of the spectral sequence of [6, 2.16] converging to $v_1^{-1}\pi'_*(SO(m))$. We will review these $v_1^{-1}\pi'_*(-)$-groups later.

When n is odd, the Ext groups from the two parts of (2.13) occur in distinct bigradings. The group $\mathrm{Ext}_{\mathcal{A}}^{s,n+t}(QK^1(\mathrm{Spin}(m))/\operatorname{im}(\psi^2))$ is nonzero if t is even and $s \geq 1$, while, as depicted in Diagram (2.16), $E_2^{s,t}(n-1, m)$ is nonzero if t is odd and $s \geq 0$. For odd n, appended to Diagram (2.16) should be a chart such

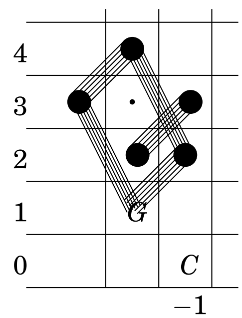
[6] If $m \equiv 0 \mod 4$, there are three summands.

as those of [3] shifted left by n gradings. The issue for α^* in (2.4) is whether the group $\widetilde{C}'$ in (2.16) supports a d_2- or d_4-differential in this new spectral sequence.

Now we return to the consideration of the various cases in the proof of Theorem (2.2).

Case 4: $k \equiv 0 \mod 4$, $m \equiv 3, 4, 5 \mod 8$. Let $k = 4\ell$. We first consider the cases when $m \equiv 3$ or $5 \mod 8$. In this case, the relevant elements of $E_2^{*,*}(8\ell + 1, m)$ are depicted in Diagram (2.17).

DIAGRAM (2.17). **A portion of** $E_2^{*,*}(8\ell + 1, m)$ **when** $m \equiv 3$ **or** $5 \mod 8$



In (2.13), the part in i^{*-1} (resp. $\mathrm{im}(q^*)$) is that in positions (x, y) with $x+y$ odd (resp. even). The indicated d_2-differentials are a consequence of the argument of Case 3; see especially the last paragraph of the proof. We consider the morphism of spectral sequences

$$E_r^{*,*}(8\ell + 1, m) \xrightarrow{i^*} E_r^{*,*}(8\ell, m). \tag{2.18}$$

The result for $\mathbf{s}[P^{8\ell}, \Phi BSO(m)]$ in [6, 1.7,1.8] was obtained from a nonzero d_3-differential from $E_3^{1,-1}$ in the spectral sequence for $v_1^{-1}\pi_*(\mathrm{Spin}(m))$ as established in [3, 3.8], which implies that $d_3 \neq 0$ on $\mathbf{s}E_3^{0,-1}(8\ell, m)$. Hence either $d_2 \neq 0$ or $d_3 \neq 0$ on the generator of C in Diagram (2.17). To know that $\mathrm{coker}(\phi_{8\ell+1,m}) = 0$, we need to know that it is not the case that d_2 is nonzero on the generator of C, and also d_3 nonzero on twice the generator; this follows by naturality using (2.18), since i^* is injective on C and the $\mathbf{Z}_2$ in filtration 3.

If $m \equiv 4 \mod 8$, the same situation applies. There are more target classes for differentials, but those in filtration 4 are killed by d_2-differentials, as indicated in Diagram (2.17), because the relevant new classes from $E_2(S^{m-1})$ occur in the same sort of eta-towers as did those in $E_2(\mathrm{Spin}(m-1))$. (See, e.g., [3, 3.16].) The filtration-3 targets map isomorphically to those in $E_2(8\ell, m)$, and $d_3 \neq 0$ on $\mathbf{s}E_3^{0,-1}(8\ell, m)$, this time by [3, 3.14]. Thus the same naturality argument implies that it is impossible that both d_2 and d_3 are nonzero from $E_2^{0,-1}$. Hence $\mathrm{coker}(\phi_{8\ell+1,m}) = 0$. This completes the proof of Case 4.

Case 5: $k \equiv 2 \mod 4$, $m \equiv 6 \mod 8$. Let $k = 4\ell + 2$ and $m = 8i + 6$. We use the commutative diagram of exact sequences

$$\begin{array}{ccccc} [P^{8\ell+5}, \Phi BSO(8i+5)] & \xrightarrow{i^*} & [P^{8\ell+4}, \Phi BSO(8i+5)] & \xrightarrow{\alpha_1^*} & v_1^{-1}\pi_{8\ell+3}(SO(8i+5)) \\ {\scriptstyle j_1}\downarrow & & {\scriptstyle j_2}\downarrow & & {\scriptstyle j_3}\downarrow \\ [P^{8\ell+5}, \Phi BSO(8i+6)] & \xrightarrow{i'^*} & [P^{8\ell+4}, \Phi BSO(8i+6)] & \xrightarrow{\alpha_2^*} & v_1^{-1}\pi_{8\ell+3}(SO(8i+6)) \end{array}$$

By [6, 1.10], j_2 on stable summands is an isomorphism of $\mathbf{Z}/2^{4i+3}$. By §3,

$$\nu(\mathbf{s}[P^{8\ell+5}, \Phi BSO(8i+5)]) < 4i+3,$$

and hence $\phi_{8\ell+5,8i+5}(= \mathbf{s}i^*)$ is not surjective. By [3, 3.7,3.8,3.10],

$$v_1^{-1}\pi_{8\ell+3}(SO(8i+5)) \approx \mathbf{Z}/2,$$

with generator D. By [3, 3.11,3.12,3.13],

$$v_1^{-1}\pi_{8\ell+3}(SO(8i+6)) \approx \mathbf{Z}/2^{\min(4i+2,\nu(\ell-i)+4)}.$$

(The 2-line group has exponent 1 larger than this, but it supports a nonzero differential.) Thus, with gen denoting a generator of the stable summand, $\alpha_2^*(\text{gen}) = j_3(D)$ and $\alpha_2^*(2 \cdot \text{gen}) = 0$. Hence $|\operatorname{coker}(\phi_{8\ell+5,8i+6})| \le 2$ and it equals 2 if and only if $j_3^{\#}$ sends the generator of $E_2^{2,8\ell+5}(\mathrm{Spin}(8i+6))^{\#}$ to $D \in E_2^{2,8\ell+5}(\mathrm{Spin}(8i+5))^{\#}$.

In the proof of [3, 3.11], which appears near the end of [3, §7], it is proved that the relevant summand of $E_2^{2,8\ell+5}(\mathrm{Spin}(8i+6))^{\#}$ is $\mathbf{Z}/2^{4i+3}$ generated by D_+ if $\nu(\ell-i) > 4i-2$, while if $\nu(\ell-i) \le 4i-2$, it is $\mathbf{Z}/2^{5+\nu(\ell-i)}$ generated by $2^{4i-2-\nu(\ell-i)}D_+ - x_{4i-1}$. Since restriction $j_3^{\#}$ to $\mathrm{Spin}(8i+5)$ sends D_+ to D and x_{4i-1} to x_{4i-1}, we deduce that $j_3^{\#}$ maps onto D if and only if $\nu(\ell-i) \ge 4i-2$, establishing the claim in (2.2) about $\operatorname{coker}(\phi_{8\ell+5,8i+6})$, one of the asterisk cases in (1.2) and (1.3).

Case 6: $k \equiv 0 \mod 4$, $m \equiv 2 \mod 8$. The argument is similar to that of Case 5, although it has one additional complication. We use a diagram of exact sequences analogous to that of Case 5, with dimensions of projective spaces and indices of $\Phi BSO(-)$ decreased by 4. By [6, 1.7,1.8], $\mathbf{s}j_2$ is an isomorphism of $\mathbf{Z}/2^{4i}$. Using §3, $\nu(\mathbf{s}[P^{8\ell+1}, BSO(8i+1)]) < 4i+1$. As we showed at the beginning of the proof of (2.2), $\ker(\phi_{8\ell+1,8i+1}) = \mathbf{Z}/2$, and hence $\phi_{8\ell+1,8i+1}$ cannot be surjective.

What complicates the argument compared to Case 5 is that $v_1^{-1}\pi_{8\ell-1}(SO(8i+1))$ and $v_1^{-1}\pi_{8\ell-1}(SO(8i+2))$ are larger than the corresponding groups that appeared in Case 5. These groups are taken from [3, 1.3,3.12]. Both of these groups have a large $\mathbf{Z}_2$-vector space in filtration 4, which maps isomorphically under j_3. It is not an issue as possible image of α_1^* on the stable summand because, as in Case 3, it is in the image under α_1^* from a similar sum of $\mathbf{Z}_2$'s. From the point of view of the spectral sequence of (2.10), they are already hit by d_2-differentials, and so we don't have to worry about whether they are hit by d_4's.

What is more of a worry is that $E_\infty^{2,8\ell+1}(\mathrm{Spin}(8i+1))$ and $E_\infty^{2,8\ell+1}(\mathrm{Spin}(8i+2))$ have, in addition to, respectively, the $\mathbf{Z}_2$-class D and the larger cyclic summand C' that they had in Case 5, also a summand L, which is the sum of many $\mathbf{Z}_2$'s and maps isomorphically under j_3, while the first group also has an additional $\mathbf{Z}_2$-class labeled x_{4i-3}. The summand L is depicted by the big dots in [3, 1.3,3.12] and has dimension $[\log_2(\frac{4}{3}(4i-1))]$. We will show that α_1^* sends the generator of the stable summand to just the class D. The analysis of whether D hits the element of order 2 in C' proceeds exactly as in Case 5. We obtain that j_3 sends D nontrivially, and hence $\operatorname{coker}(\phi_{8\ell+1,8i+2}) = \mathbf{Z}/2$, if and only if $\nu(\ell - i) \geq 4i - 4$, which translates to the claim of the theorem in this case, the other asterisk case in (1.2) and (1.3).

It remains to verify the claim about α_1^*, which is done by applying Pontryagin duality. By (2.6) and (2.7), $\alpha_1^\#$ is determined by

$$E_2^{2,1}(\mathrm{Spin}(8i+1))^\# \xrightarrow{h_1^\#} E_2^{1,-1}(\mathrm{Spin}(8i+1))^\#.$$

That this sends only the class D nontrivially to the stable summand is proved exactly as in the two paragraphs of [6] which appear shortly after Diagram 2.24 of that paper. The first of the two paragraphs begins "In order to show that $d_3(g_1) = 0$." In summary, a presentation of $E_2^{1,-1}(\mathrm{Spin}(8i+1))^\#$ is given, and, for each basis element b of $E_2^{2,1}(\mathrm{Spin}(8i+1))^\#$, $(h_1)^\#(b)$ is interpreted as an element in that presented group, and it is observed that only $(h_1)^\#(D)$ is nonzero.

Case 7: $k \equiv 0 \mod 4$, $m \equiv 6 \mod 8$. Let $k = 4\ell$ and $m = 8i + 6$. This time the diagram of the sort used in Case 5 does not quite work because j_2 is not surjective, due to a d_3-differential in $[P^{8\ell}, \Phi BSO(8i+5)]$ not present in $[P^{8\ell}, \Phi BSO(8i+6)]$. We can, however, consider an E_2-version of the diagram, where α_1^* and α_2^* are, after dualizing, given by (2.7). The diagram below addresses what amounts to the d_2-differential on $\mathbf{s}E_2^{0,-1}(8\ell+1, 8i+6)$. The d_4-differential on this summand is then eliminated similarly to Cases 3, 4, and 6.

$$\begin{array}{ccccc}
\mathbf{s}E_2^{0,-1}(8\ell, 8i+5)^\# & \xrightarrow{\approx} & \mathbf{s}E_2^{1,-1}(\mathrm{Spin}(8i+5))^\# & \xleftarrow{v_1^{4\ell}h_1^\#} & E_2^{2,8\ell+1}(\mathrm{Spin}(8i+5))^\# \\
\approx\Big\uparrow & & j_2^\#\Big\uparrow\approx & & j_3^\#\Big\uparrow \\
\mathbf{s}E_2^{0,-1}(8\ell, 8i+6)^\# & \xrightarrow{\approx} & \mathbf{s}E_2^{1,-1}(\mathrm{Spin}(8i+6))^\# & \xleftarrow{v_1^{4\ell}h_1^\#} & E_2^{2,8\ell+1}(\mathrm{Spin}(8i+6))^\#
\end{array}$$

As in Case 6, the $v_1^{4\ell}h_1^\#$ on $\mathrm{Spin}(8i+5)$ sends only D nontrivially, and $j_3^\#$ sends the generator of the C'-summand to x_{4i-1}, since $\nu((8\ell+1)-(8i+5)) = 2$. Thus $v_1^{4\ell}h_1^\#$ on $\mathrm{Spin}(8i+6)$ is 0, and hence $\phi_{8\ell+1,8i+6}$ is surjective.

Case 8: $k \equiv 2 \mod 4$, $m \equiv 2 \mod 8$. Let $k = 4\ell + 2$. The argument is similar to that of Case 7, but is complicated by $P^{8\ell+4}$ not being K-equivalent to a Moore spectrum. Let, as in [6, 2.14],

$$T^n = S^n \cup_\eta e^{n+2} \cup_2 e^{n+3}.$$

From [6, (2.11),(2.13)], we have

$$\mathbf{s}[P^{8\ell+4}, \Phi BSO(m)] \approx \mathbf{s}v_1^{-1}\pi'_{-2}(SO(m)), \tag{2.19}$$

where, by [6, (2.17)],

$$v_1^{-1}\pi'_n(X) \approx [T^n, \Phi(X)]. \tag{2.20}$$

The analogue of (2.6) is that the morphism α^* in (2.4) is equivalent to

$$\zeta_\ell^* : v_1^{-1}\pi'_{-1}(BSO(m)) \to v_1^{-1}\pi_{8\ell+4}(BSO(m)),$$

where $\zeta_\ell : S^{8\ell+5} \to T^0$ is the element of highest filtration $(4\ell+2)$ in its stem in the Adams spectral sequence of T^0. It is $\eta\mu_\ell$ on the top cell. The reason for this is similar to the discussion between (2.6) and (2.7). In this case, both

$$S^{8\ell+4} \xrightarrow{\alpha} P^{8\ell+4} \xrightarrow{\phi^\ell} P^4_{1-8\ell}$$

and

$$S^{8\ell+4} \xrightarrow{\zeta_\ell} T^{-1} \xrightarrow{f} P^4_{1-8\ell},$$

where f is, up to periodicity, a restriction of the map in [6, 2.8], become equal in $\pi_{8\ell+4}(P^4_{1-8\ell} \wedge J) \approx \mathbf{Z}_2 \oplus \mathbf{Z}_2$, where each is the element of highest filtration. Note that f has Adams filtration -1. Thus the two composites are equal in $v_1^{-1}\pi_{8\ell+4}(P^4_{1-8\ell})$, and hence, following by any element g of $[P^4_{1-8\ell}, \Phi BSO(m)] \approx [P^{8\ell+4}, \Phi BSO(m)]$, $\alpha^*(g) = \zeta_\ell^*(g \circ f)$ in $\pi_{8\ell+4}(\Phi BSO(m))$. Note that f induces the isomorphism obtained from (2.19) and (2.20).

Let $M^6 \xrightarrow{\widetilde{\zeta}} T^0$ be an extension of ζ. Here M^n is the mod-2 Moore spectrum with top cell in dimension n. We claim that

$$\widetilde{\zeta}^* : K^0(T^0) \to K^0(M^6) \tag{2.21}$$

is the nontrivial morphism from $\mathbf{Z}_2^\wedge$ to $\mathbf{Z}/2$. One way to see this is to obtain $ku_*(D(\widetilde{\zeta}))$ from $ko_*(D(\widetilde{\zeta}))$ by using $bu = bo \cup_\eta \Sigma^2 bo$. Here D denotes the S-dual. There is a cofiber sequence

$$M^{-6} \to D(MC(\widetilde{\zeta})) \to D(T^0).$$

In the chart below, the solid dots are from the M^{-6} and the circles from $D(T^0)$. The differential in the ko_*-chart is due to the η^2 connection. It implies the differential in the ku_*-chart, which is the asserted homomorphism (2.21).

DIAGRAM (2.22). *$ko_*(D(MC(\widetilde{\zeta})))$ and $ku_*(D(MC(\widetilde{\zeta})))$*

From e.g. [4, p.488] or [3, 3.6,3.16], $\mathrm{Ext}_{\mathcal{A}}^{1,n+6}(PK^1(S^n)) \approx \mathbf{Z}/2$. We will name the nonzero class $v_1^2h_1$. In the spectral sequence converging to $v_1^{-1}\pi_*(S^n)$, this element supports a d_3-differential, but in that converging to $v_1^{-1}\pi'_*(S^n)$, it survives to a homotopy class, which is the class ζ discussed above. (See [6, 2.18].) We obtain the following analogue of Diagram (2.9).

DIAGRAM (2.23). *Diagram involving Bockstein and $v_1^2h_1$*

$$\begin{tikzcd}
 & E_2^{s,s+n+6}(Y;\mathbf{Z}_2) \arrow[r, "v_1^{4\ell}"] \arrow[d, "\partial"] & E_2^{s,s+n+6+8\ell}(Y;\mathbf{Z}_2) \arrow[d, "\partial"] \\
E_2^{s,s+n}(Y) \arrow[ur, "\widetilde{\zeta}^*"] \arrow[r, "v_1^2h_1"] & E_2^{s+1,s+n+6}(Y) \arrow[r, "v_1^{4\ell}"] & E_2^{s+1,s+n+6+8\ell}(Y)
\end{tikzcd}$$

Here Y could be any space, but we use $Y = \mathrm{Spin}(m)$. The point of the diagram is that the composition around the top is α^*, while the composition on the bottom sends an eta-tower to one with the same name. The claim about (2.21) was needed to establish commutativity of the triangle.

Now that we have related α^* to $v_1^{4\ell+2}h_1$, we obtain the following analogue of the diagram in Case 7.

$$\begin{tikzcd}
\mathbf{s}E_2^{0,-1}(8\ell+4, 8i+1)^{\#} \arrow[r, "\approx"] & \mathbf{s}E_2^{1,-1}(\mathrm{Spin}(8i+1))^{\#} & E_2^{2,8\ell+5}(\mathrm{Spin}(8i+1))^{\#} \arrow[l, "v_1^{4\ell+2}h_1^{\#}"'] \\
\mathbf{s}E_2^{0,-1}(8\ell+4, 8i+2)^{\#} \arrow[r, "\approx"] \arrow[u, "\approx"] & \mathbf{s}E_2^{1,-1}(\mathrm{Spin}(8i+2))^{\#} \arrow[u, "j_2^{\#}", "\approx"'] & E_2^{2,8\ell+5}(\mathrm{Spin}(8i+2))^{\#} \arrow[l, "v_1^{4\ell+2}h_1^{\#}"'] \arrow[u, "j_3^{\#}"]
\end{tikzcd}$$

The same argument as in Case 7 now implies

$$d_2 = 0 : \mathbf{s}E_2^{0,-1}(8\ell+5, 8i+2) \to E_2^{2,0}(8\ell+5, 8i+2).$$

The d_3-differential on $\mathbf{s}E_3^{0,-1}(8\ell+5, 8i+2)$ is as it was on $\mathbf{s}E_3^{0,-1}(8\ell+4, 8i+2)$, which was shown to be 0 in [6].[7] That $d_4 = 0$ on $\mathbf{s}E_4^{0,-1}(8\ell+5, 8i+2)$ is seen as in most of the previous cases, using Diagram (2.23) to assert that the target was already hit by d_2 applied to eta-towers with the same name.

Case 9: $k \equiv 3 \mod 4$, $m \not\equiv 2 \mod 4$, and $m \geq 12$. We decompose α^* in (2.4) as

$$[P^{2k}, \Phi BSO(m)] \xrightarrow{\widetilde{\alpha}^*} [M^{2k+1}, \Phi BSO(m)] \xrightarrow{i^*} v_1^{-1}\pi_{2k-1}(SO(m)), \tag{2.24}$$

where $M^n = M^n(2)$, and $\widetilde{\alpha}$ is the attaching map for the top two cells of P^{2k+2}. Let $k = 4\ell - 1$. There is a commutative diagram in which rows are cofiber sequences and columns are K-equivalences

$$\begin{tikzcd}
M^{8\ell-1} \arrow[r, "\widetilde{\alpha}"] \arrow[d, "A^{\ell}"'] & P^{8\ell-2} \arrow[r] \arrow[d] & P^{8\ell} \arrow[r] \arrow[d] & M^{8\ell} \arrow[d] \\
M^{-1} \arrow[r, "\alpha'"] & P_{1-8\ell}^{-2} \arrow[r] & P_{1-8\ell}^{0} \arrow[r] & M^0 \\
M^{-1} \arrow[r, "q"] \arrow[u, "="] & M^0(2^{4\ell-1}) \arrow[r, "2"] \arrow[u] & M^0(2^{4\ell}) \arrow[r] \arrow[u] & M^0 \arrow[u, "="']
\end{tikzcd} \tag{2.25}$$

[7] It was done in the paragraph of [6] near the end of Section 2, which begins "We prove now that $d_3 = 0$ on $\widetilde{E}_2^{1,-1}(\mathrm{Spin}(8i+2))$."

The top vertical maps are just the v_1-maps. The middle square on the bottom is from [6, 2.2], which was originally from [11]. The construction in [11] implies commutativity of the lower right square. If this cofiber sequence is pushed one space farther, a commutative square is obtained which is the suspension of the lower left square. Hence the lower left square commutes.

Thus we obtain a commutative diagram

$$\begin{array}{ccc} \mathbf{s}[P^{8\ell-2}, \Phi BSO(m)] & \xrightarrow{\widetilde{\alpha}^*} & [M^{8\ell-1}, \Phi BSO(m)] \\ \approx\big\uparrow & & \approx\big\uparrow \\ \mathbf{s}[P^{-2}_{1-8\ell}, \Phi BSO(m)] & \xrightarrow{\alpha'^*} & [M^{-1}, \Phi BSO(m)] \\ \approx\big\downarrow & & =\big\downarrow \\ \mathbf{s}v_1^{-1}\pi_{-2}(SO(m)) & \xrightarrow{q^*} & [M^{-1}, \Phi BSO(m)], \end{array} \tag{2.26}$$

where q is the collapse map. In the bottom row, $\mathbf{s}[M^0(2^{4\ell-1}), \Phi BSO(m)]$ has been replaced by $\mathbf{s}v_1^{-1}\pi_{-1}(BSO(m)) \approx \mathbf{s}v_1^{-1}\pi_{-2}(SO(m))$ because ℓ can be taken to be arbitrarily large, and so the maps from the top cell of the Moore space are ephemeral. When the $\widetilde{\alpha}^*$ in the top row is followed by i^* into $v_1^{-1}\pi_{8\ell-3}(SO(m))$ to yield (2.24), we obtain from the diagram something agreeing up to isomorphisms with that obtained by applying $\mathbf{s}[-, \Phi BSO(m)]$ to the composite

$$S^{8\ell-2} \hookrightarrow M^{8\ell-1} \xrightarrow{A^\ell} M^{-1} \xrightarrow{q} S^{-1}. \tag{2.27}$$

By [2], this composite is the element of order 2 in the stable image of J in the $(8\ell-1)$-stem; however, we will compute it using (2.27) rather than this imJ description. We will show that the composite

$$\begin{array}{lll} \mathbf{s}E_2^{1,-1}(\mathrm{Spin}(m)) & \underset{q^*}{\xrightarrow{\rho_2}} & E_2^{1,-1}(\mathrm{Spin}(m);\mathbf{Z}_2) \xrightarrow{A^\ell} E_2^{1,8\ell-1}(\mathrm{Spin}(m);\mathbf{Z}_2) \\ & \underset{i^*}{\xrightarrow{\partial}} & E_2^{2,8\ell-1}(\mathrm{Spin}(m)) \end{array} \tag{2.28}$$

is 0.[8] Noting that

$$E_\infty^{4,8\ell+1}(\mathrm{Spin}(m)) = 0 \tag{2.29}$$

by [3, 1.3,3.6,3.7], Theorem (2.2) follows in this case.

We show that the Pontryagin dual of (2.28) is 0. Let

$$C_0 \xrightarrow{d_1} C_1 \xrightarrow{d_2} C_2$$

be the sequence of free $\mathbf{Z}_{(2)}$-modules associated to the sequence of free $\mathbf{Z}_2^\wedge$-modules in [3, 11.9]. Thus $C_0 = F$, $C_1 = F \oplus F \oplus F$, and $C_2 = F \oplus F \oplus F \oplus F$, where F is a free $\mathbf{Z}_{(2)}$-module on $[m/2]$ generators. The transpose of the matrix of d_1 is

$$(0 \quad \Psi^2 \quad \Theta_{4\ell-1}), \tag{2.30}$$

[8]Note that ρ_2 and ∂ are parts of different Bockstein exact sequences, and so it is not automatic that the composite is 0.

and the transpose of the matrix of d_2 is

$$\begin{pmatrix} -2 & \Psi^2 & \Theta_{4\ell-1} & 0 \\ 0 & 0 & 0 & \Theta_{4\ell-1} \\ 0 & 0 & 0 & -\Psi^2 \end{pmatrix}, \tag{2.31}$$

and then the homology at C_s is $\mathrm{Ext}_{\mathcal{A}}^{s,8\ell-1}(PK^1(\mathrm{Spin}(m))/\operatorname{im}(\psi^2))$. Here Ψ^2 (resp. Θ_j) is the matrix of ψ^2 (resp. $\psi^3 - 3^j$) on $PK^1(\mathrm{Spin}(m))$. We are using here that for a rationally acyclic complex of finitely generated free $\mathbf{Z}_{(2)}$-modules, the inclusion induces an isomorphism $H_*(-;\mathbf{Z}_{(2)}) \to H_*(-;\mathbf{Z}_2^\wedge)$. In the remainder of this proof, we will write $\mathbf{Z}$ when we really mean $\mathbf{Z}_{(2)}$.

As observed in [3, proof of 11.3], $E_2^{s,8\ell-1}(\mathrm{Spin}(m))^\#$ is the homology at C^*_{s-1} of the chain complex C^* given by

$$C_0^* \xleftarrow{d_1^*} C_1^* \xleftarrow{d_2^*} C_2^*, \tag{2.32}$$

where $C_s^* = \mathrm{Hom}(C_s, \mathbf{Z})$ and the matrices of d_1^* and d_2^* are those of (2.30) and (2.31). The shift from s to $s-1$ is due to the short exact sequence

$$0 \to \mathbf{Z} \to \mathbf{Q} \to \mathbf{Q}/\mathbf{Z} \to 0.$$

Note that $E_2^{s,4\ell-1}(\mathrm{Spin}(m);\mathbf{Z}/2)^\#$ is the homology at $C_s^*/2$ of the mod 2 reduction of (2.32), and

$$\rho_2^\# : E_2^{1,8\ell-1}(\mathrm{Spin}(m);\mathbf{Z}/2)^\# \to E_2^{1,8\ell-1}(\mathrm{Spin}(m))^\#$$

is the boundary homomorphism δ in the exact sequence of homology groups induced by the short exact sequence of chain complexes

$$0 \to C^* \xrightarrow{2} C^* \to C^*/2 \to 0. \tag{2.33}$$

To see this, note that the commutative diagram

$$\begin{tikzcd} 0 \arrow[r] & \mathbf{Z} \arrow[r, "2"] \arrow[d, "1"'] & \mathbf{Z} \arrow[r] \arrow[d, "\frac{1}{2}"'] & \mathbf{Z}/2 \arrow[r] \arrow[d, "i"'] & 0 \\ 0 \arrow[r] & \mathbf{Z} \arrow[r] & \mathbf{Q} \arrow[r] & \mathbf{Q}/\mathbf{Z} \arrow[r] & 0 \end{tikzcd}$$

induces a commutative diagram

$$\begin{tikzcd} H_1(C^*/2) \arrow[r, "\delta"] \arrow[d, "\rho_2^*"'] & H_0(C^*) \arrow[d, "="'] \\ H_1(C^* \otimes \mathbf{Q}/\mathbf{Z}) \arrow[r, "\approx"] & H_0(C^*), \end{tikzcd}$$

from which the agreement of δ and ρ_2^* is immediate.

The composite which we wish to show is 0 (dual to (2.28)) may now be identified as

$$H_1(C^*_{(4\ell-1)}) \xrightarrow{\rho_{2*}} H_1(C^*_{(4\ell-1)}/2) \xrightarrow{=} H_1(C^*_{(-1)}/2) \xrightarrow{\delta} \mathbf{s}H_0(C^*_{(-1)}). \tag{2.34}$$

Here the parenthesized subscript of C^* is the subscript of Θ, and $C^*/2$ means the mod 2 reduction of C^*. The identity map in the middle is due to the subscript not mattering mod 2, and the fact that A^* is the identity homomorphism of $K^*(M)$. Since, for the same parenthesized subscript, $\operatorname{im}(\rho_2^*) = \ker(\delta)$, we are reduced to proving

$$\ker(H_1(C^*/2) \xrightarrow{\delta_\ell} H_0(C^*_{(4\ell-1)})) \subset \ker(H_1(C^*/2) \xrightarrow{\delta_0} \mathbf{s}H_0(C^*_{(-1)})). \tag{2.35}$$

We will need the following result, culled from [3].

THEOREM (2.36). *Suppose $m \geq 12$.*

- *If $m = 2n+1$, then*

$$H_0(C^*_{(4\ell-1)}) \approx \begin{cases} \mathbf{Z}/2^n \oplus \mathbf{Z}/2^n & n \leq \nu(\ell)+4 \\ \mathbf{Z}/2^e \oplus \mathbf{Z}/2^{\nu(\ell)+4} & n > \nu(\ell)+4 \end{cases} \tag{2.37}$$

with $e > n$. The group is presented by a matrix

$$\begin{pmatrix} 2^{A_1} & 0 \\ u_2 2^{A_2} & 2^n \\ u_3 2^n & 2^v \end{pmatrix}, \tag{2.38}$$

where u_i is odd, $A_i > n$, and $v = \min(\nu(\ell)+4, 2n+1)$. The columns of this matrix correspond to generators ξ_1 and D of $PK^1(\mathrm{Spin}(m))$ under the isomorphism

$$H_0(C^*_{(4\ell-1)}) \approx E_2^{1,8\ell-1}(\mathrm{Spin}(m))^{\#} \approx PK^1(\mathrm{Spin}(m))/(\psi^2, \theta_{4\ell-1}), \tag{2.39}$$

where $\theta_j = \psi^3 - 3^j$. The first row of (2.38) is due to a combination of relations of the form $\psi^2(\xi_i)$ and $\theta_{4\ell-1}(\xi_i)$, while the second row is a combination of such relations together with $\psi^2(D)$ (with coefficient 1), and the third row is a combination of such relations together with $1 \cdot \theta_{4\ell-1}(D)$. The first summand of (2.37) is the stable summand; it corresponds to the first (ξ_1) column of (2.38).

- *If $m = 4a$, then*

$$H_0(C^*_{(4\ell-1)}) \approx \begin{cases} \mathbf{Z}/2^{2a} \oplus \mathbf{Z}/2^{2a-1} \oplus \mathbf{Z}/2^{\nu(a)+2} & 2a \leq \nu(\ell)+5 \\ \mathbf{Z}/2^{e_1} \oplus \mathbf{Z}/2^{e_2} \oplus \mathbf{Z}/2^{e_3} & \textit{otherwise,} \end{cases}$$

with $e_1 > 2a$ and $e_3 \leq e_2 < 2a$. The group is presented by a matrix

$$\begin{pmatrix} 2^{A_1} & 0 & 0 \\ 0 & 2^M & -2^M \\ u_2 2^{A_2} & 2^{2a-1} & 0 \\ 2^{2a-1} & u_3 2^{v_1} & u_4 2^{v_2} \end{pmatrix} \tag{2.40}$$

with u_i odd, $A_i > 2a$, $M = \min(2a-1, \nu(2\ell-a)+3)$, $v_1 = \min'(\nu(a)+2, \nu(\ell)+4)$, and $v_2 = \nu(a)+2$. Here $\min'(A,B) = \min(A,B)$ unless $A = B$, in which case it is greater than either. Under the isomorphisms of (2.39), the columns of (2.40) correspond to generators ξ_1, D_+, and D_-, and of the rows (relations) only the last one involves an odd multiple of $\theta_{4\ell-1}(D)$.

Proof. For the first part, we use [3, 3.1,3.2] and [5, 3.18]. The proof of [5, 3.15] explains how the rows of the presentation matrix are obtained, while [5, §4] derives the inequalities for the exponents in those relations. Actually, [5, 3.18] only proves $A_i \geq n$. The stronger result needed here follows by a more careful analysis of the proof of [5, §4]. It follows from [5, 3.18], refined to say that $eSp(4\ell+1, n) > n+1$ and the coefficients of ξ_1 in [5, (3.19)] and [5, (3.20)] are divisible by 2^{n+1}.

By [3, 8.1], $eSp(-, n)$ is divisible by $(2n+1)!$, which is divisible by 2^{n+1} for $n \geq 2$. The divisibility of [5, (3.20)] is proved using its representation as

$$(n-1)2^{2n-4} + \sum_{j=2}^{n/2} \tbinom{n-j}{j} 2^{2n-4j} \sum_{i \geq j-1} 8^i \tbinom{2\ell-1}{i} S_{i,j}$$

with

$$S_{i,j} = \sum_{t=0}^{j-2} (-1)^t \tbinom{2j-1}{t} (2j-2t-1) \tbinom{j-t}{2}^i$$

given in [5, (4.20)]. The term $(n-1)2^{2n-4}$ is divisible by 2^{n+1} for $n \geq 5$. The other terms are divisible by 2^{2n-j-3} with $2 \leq j \leq n/2$, which will be sufficiently divisible except when (n, j) is (6,3). In this case, the additional divisibility is provided by $S_{2,3} = 30$.

The divisibility of [5, (3.19)] is proved similarly using its representation as

$$(n+1)2^{2n-3} \sum_{j \geq 2} 2^{2n+1-4j} \left(\tbinom{n+2-j}{j} - \tbinom{n-j}{j-2} \right) \sum_{i \geq j-1} 8^i \tbinom{2\ell}{i} S_{i,j},$$

with $S_{i,j}$ as above, from [5, p.54]. The lead term $(n+1)2^{2n-3}$ is divisible by 2^{n+1} for $n \geq 3$. Other terms are divisible by 2^{2n-j-2} with $2 \leq j \leq n/2$, which is divisible by 2^{n+1}.

For the second part, we use [3, 3.3] and its proof in [3, §4]. The classes ξ_i, D, D_+, and D_- in $PK^1(\mathrm{Spin}(m))$ are as in [5, 3.10] and [3, 4.1], but do not play a major role in this paper. □

We remark that the condition $m \geq 12$ is necessary for the divisibilities of the entries of the matrices to hold.

By the definition of δ using (2.33), if $\mathbf{x} = (x_1, x_2, x_3) \in C_1^*/2$ is a cycle representing an element of $H_1(C^*_{(4\ell-1)}/2)$, then

$$\delta(\mathbf{x}) = \tfrac{1}{2}\psi^2(x_2) + \tfrac{1}{2}\theta_{4\ell-1}(x_3), \tag{2.41}$$

viewed as an element in the group presented by one of the matrices of (2.36). Here $x_i \in F^*$ or $F/2^*$. We write δ_0 and δ_ℓ for the boundaries δ associated to $C^*_{(-1)}$ and $C^*_{(4\ell-1)}$, respectively. Note that the relations $\xi_j = j^{4\ell-1}\xi_1$ are used to bring these elements into the 2- or 3-generator form of (2.36). This relation is a consequence of [5, 3.9], which says that modding out by $\psi^j - j^{4\ell-1}$ for $j = 3$ and -1 also accomplishes modding out by $\psi^j - j^{4\ell-1}$ for other odd j.

The matrix (2.38) implies that when $m = 2n+1$, $\mathbf{s}H_0(C^*_{(-1)})$ is isomorphic to $\mathbf{Z}/2^n$ generated by ξ_1, since $v = 2n+1$ in this case, and that in (2.41) with $\ell = 0$, $\delta_0(x_1, x_2, x_3) \neq 0 \in \mathbf{s}H_0(C^*_{(-1)})$ if and only if the D-component of x_3 is odd. This key point may warrant some explanation. The interpretation of the rows of (2.38) given after (2.39) implies that when $\psi^2(x_2)$ or $\theta_{-1}(x_3)$ are written in terms of ξ_1 and D, using $\xi_j = j^{-1}\xi_1$, the ξ_1-component of each will be divisible by 2^{n+1} unless the D-component of x_3 is odd, and when these are multiplied by 1/2, as they are in (2.41), the only way to obtain a nonzero component in the ξ_1-component of the $\mathbf{Z}/2^n$-group presented by (2.38) is then to have this D-component of x_3 be odd.

If the D-component of x_3 is odd, then

$$\delta_\ell(x_1, x_2, x_3) \neq 0 \in H_0(C^*_{(4\ell-1)}), \tag{2.42}$$

since it is $\frac{1}{2}$ times the last row of (2.38) plus perhaps $\frac{1}{2}$ times the other rows. Such a vector is easily seen to be nonzero in the group presented by (2.38), regardless of the value of v. This establishes the contrapositive of (2.35).

The same argument applies when $m = 4a$, using the matrix (2.40). The previous paragraph carries through verbatim, with n replaced by $2a - 1$.

Case 10: $k \equiv 3 \mod 4$, $m \equiv 2 \mod 4$. The method of Case 9 does not apply here, since $\psi^{-1} \neq -1$ in $PK^1(\mathrm{Spin}(m))$ when $m \equiv 2 \mod 4$. However the result here follows by naturality from Case 9.

Let $k = 4\ell + 3$ and $m = 4j + 2$. The morphism $\mathbf{s}E_2^{0,-1}(8\ell + 7, 4j + 1) \to \mathbf{s}E_2^{0,-1}(8\ell+7, 4j+2)$ is bijective by [3, 3.3]. As we have just seen that $d_2 = 0$ on the former, it must also be 0 on the latter. Note that d_3 on $\mathbf{s}E_3^{0,-1}(8\ell+7, 4j+2)$ equals d_3 on $\mathbf{s}E_3^{0,-1}(8\ell+6, 4j+2)$, by the general form of the spectral sequence, and this equals d_3 on $E_3^{1,-1}(\mathrm{Spin}(4j+2))$ by the paragraph after Diagram (2.16) beginning "By the proof." By [3, 3.12], this is zero. As there is nothing for d_4 to hit by (2.29)[9], we deduce that the generator of $E_2^{0,-1}(2k + 1, m)$ is an infinite cycle in this case, establishing Theorem (2.2) in this case.

Case 11: $k \equiv 1 \mod 4$, $m \not\equiv 2 \mod 4$, $m \geq 12$. Let $k = 4\ell + 1$. Similarly to (2.25), we have, using [6, 2.8], a commutative diagram in which rows are cofibrations and columns are K-equivalences.

$$\begin{array}{ccccc}
M^{8\ell+3} & \xrightarrow{\widetilde{\alpha}} & P^{8\ell+2} & \longrightarrow & P^{8\ell+4} \\
\downarrow & & \downarrow & & \downarrow \\
M^3 & \longrightarrow & P^2_{1-8\ell} & \longrightarrow & P^4_{1-8\ell} \\
\uparrow & & \uparrow & & \uparrow \\
\Sigma^{2^{4\ell+1}L}F & \longrightarrow & N^{2^{4\ell+1}L}(2^{4\ell}) & \xrightarrow{2} & N^{2^{4\ell+1}L}(2^{4\ell+1})
\end{array}$$

where $N^n(k) = M^n(k) \cup_\eta e^{n+1} \cup_2 e^{n+2}$, the map labeled 2 has degree 2 on the bottom cell, and $\Sigma^{2^{4\ell+1}L}F$ is the stable fiber of this map. Thus

$$F = M^{-1} \cup_\eta M^1 \cup_2 M^2,$$

and, with $T^n = S^n \cup_\eta e^{n+2} \cup_2 e^{n+3}$ as in Case 8, there is a cofiber sequence

$$T^{-2} \to F \to T^{-1} \xrightarrow{2} T^{-1}. \tag{2.43}$$

[9] which also holds when $m \equiv 2 \mod 4$

Similarly to (2.26), we obtain a commutative diagram, using [6, (2.13)]

$$\begin{array}{ccc}
\mathbf{s}[P^{8\ell+2}, \Phi BSO(m)] & \xrightarrow{\widetilde{\alpha}^*} & [M^{8\ell+3}, \Phi BSO(m)] \\
\approx\uparrow & & \approx\uparrow \\
\mathbf{s}[P^2_{1-8\ell}, \Phi BSO(m)] & \longrightarrow & [M^3, \Phi BSO(m)] \\
\approx\downarrow & & \approx\downarrow \\
\mathbf{s}v_1^{-1}\pi'_{2^{4\ell+1}L-2}(SO(m)) & \longrightarrow & [\Sigma^{2^{4\ell+1}L}F, \Phi BSO(m)].
\end{array}$$

Since ℓ is large, the $\Sigma^{2^{4\ell+1}L}$ may be omitted by periodicity, and so α^* in (2.4) is obtained as the composite

(2.44)
$$\mathbf{s}v_1^{-1}\pi'_{-2}(SO(m)) \to [M^3, \Phi BSO(m)] \xrightarrow{\approx} [M^{8\ell+3}, \Phi BSO(m)] \xrightarrow{i^*} v_1^{-1}\pi_{8\ell+1}(SO(m)).$$

This can be considered as the d_2- and d_4-differentials in the spectral sequence described prior to Case 4. Recall from [6, 2.16] that the E_2-term for $v_1^{-1}\pi'_*(-)$ equals that for $v_1^{-1}\pi_*(-)$.

The cofibration (2.43) yields a short exact sequence

$$0 \to K^{-1}(T^{-1}) \xrightarrow{2} K^{-1}(T^{-1}) \to K^{-1}(F) \to 0$$

which is

$$0 \to \mathbf{Z}_2^\wedge \xrightarrow{2} \mathbf{Z}_2^\wedge \to \mathbf{Z}/2 \to 0.$$

Thus (2.44) is, at the E_2-level, given by

$$(2.45)\quad \mathbf{s}E_2^{1,-1}(\mathrm{Spin}(m)) \xrightarrow{\rho_2} E_2^{1,3}(\mathrm{Spin}(m);\mathbf{Z}/2) \xrightarrow{\approx} E_2^{1,8\ell+3}(\mathrm{Spin}(m);\mathbf{Z}/2) \xrightarrow{\partial} E_2^{2,8\ell+3}(\mathrm{Spin}(m)),$$

similarly to (2.28). We can justify the ρ_2 between distinct bigradings in two ways. (a) $\mathrm{Ext}_{\mathcal{A}}^{s,t}(-;\mathbf{Z}/2)$ has period 4 in t; (b) The morphism is induced by $F \to T^{-1}$, and there is a K-equivalence $F \to M^3$.

Hence, by the same argument used in Case 9 to go from (2.28) to (2.35), showing that $d_2 = 0$ on $\mathbf{s}E_2^{0,-1}(8\ell+3, m)$ is equivalent to proving

$$(2.46)\qquad \ker(H_1(C^*/2) \xrightarrow{\delta'_\ell} H_0(C^*_{(4\ell+1)})) \subset \ker(H_1(C^*/2) \xrightarrow{\delta_0} \mathbf{s}H_0(C^*_{(-1)})).$$

Here $\delta'_\ell(x_1, x_2, x_3) = \frac{1}{2}\psi^2(x_2) + \frac{1}{2}\theta_{4\ell+1}(x_3)$.

The proof that (2.46) holds is similar to that of Case 9, except that the matrix, using $\psi^3 - 3^{4\ell+1}$ instead of $\psi^3 - 3^{4\ell-1}$ has a slightly different form. The matrix is described in Lemma (2.50) when m is odd. One must prove, analogous to (2.42), that if the D-component of x_3 is odd, then $\delta'_\ell(x_1, x_2, x_3) \neq 0 \in H_0(C^*_{(4\ell+1)})$. This is easier than in Case 9 because of the 2^3 in the last row of (2.51). As before, the last row is characterized by being the relation due to $\theta_{4\ell+1}(D)$ plus other terms. Hence $\delta'_\ell(x_1, x_2, x_3)$ will involve 1/2 times the last row of (2.51), which, because of the 2^3 is certainly nonzero in the group presented by (2.51).

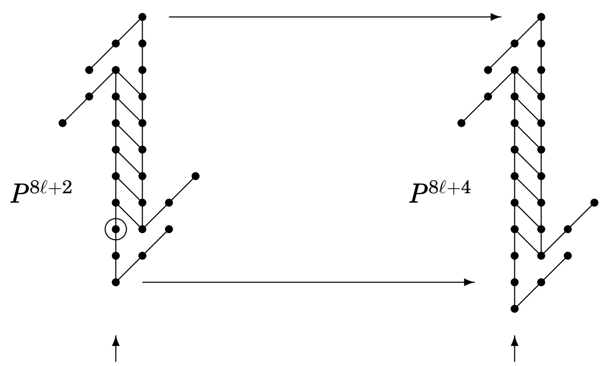
Finally, we must show $d_4 = 0$ on $\mathbf{s}E_4^{0,-1}(8\ell+3, m)$. The composite (2.44) may be viewed as applying $[-, \Phi BSO(m)]$ to

$$(2.47)\qquad S^{8\ell+2} \xrightarrow{\alpha} P^{8\ell+2} \to P^2_{1-8\ell} \to v_1^{-1}P^2_{1-8\ell} \simeq v_1^{-1}N^0(2^{4\ell}).$$

The class of this composite is divisible by 4 in $v_1^{-1}\pi_{4\ell+2}(N^0(2^{4\ell})) \approx v_1^{-1}\pi_{4\ell+2}(P^{8\ell+2})$. Call it 4γ.

To see this divisibility, we use that α goes to 0 in $v_1^{-1}\pi_{8\ell+2}(P^{8\ell+4})$, since it is an attaching map. Diagram (2.48), which is similar to those of [13, pp 94-5], depicts $v_1^{-1}\pi_*(P^{8\ell+2}) \to v_1^{-1}\pi_*(P^{8\ell+4})$ near $* = 8\ell+2$. The group where $* = 8\ell+2$ is indicated with an arrow, and the nonzero element in the kernel of this homomorphism is circled.

DIAGRAM (2.48). $v_1^{-1}\pi_*(P^{8\ell+2}) \to v_1^{-1}\pi_*(P^{8\ell+4})$ **near** $* = 8\ell+2$



This chart also depicts $v_1^{-1}\pi_*(N^0(2^{4\ell}))$, and the circled element equals the composite (2.47) (since the α is nontrivial, because Sq^4 is nonzero in its mapping cone). The inclusion $v_1^{-1}T^{-1} \xrightarrow{i_T} v_1^{-1}N^0(2^{4\ell})$ induces in $\pi_{8\ell+2}(-)$ an injection $\mathbf{Z}/8 \to \mathbf{Z}/8 \oplus \mathbf{Z}/2$.[10]

Let g denote the generator of $v_1^{-1}\pi_{-2}(T^{-1})$, and let $2^e g$ denote an extension of $2^e g$ over an appropriate Moore spectrum. Then (2.47) equals the top row of the commutative diagram (2.49) followed by i_T.

$$\begin{array}{ccccccc} S^{8\ell+3} & \xrightarrow{i} & M^{8\ell+3} & \xrightarrow{A^\ell} & M^3 & \xrightarrow{4g} & v_1^{-1}T^{-1} \\ {\scriptstyle 2}\downarrow & & {\scriptstyle 2}\downarrow & & {\scriptstyle 2}\downarrow & & {\scriptstyle =}\downarrow \\ S^{8\ell+3} & \xrightarrow{i} & M^{8\ell+3}(4) & \xrightarrow{A^\ell} & M^3(4) & \xrightarrow{2g} & v_1^{-1}T^{-1} \end{array} \tag{2.49}$$

Here $2 : M^{8\ell+3} \to M^{8\ell+3}(4)$ from the mod 2 Moore spectrum to the mod 4 Moore spectrum has degree 2 on the bottom cell and degree 1 on the top cell.

Since $E_2^{3,8\ell+4}(\mathrm{Spin}(m))$ and $E_2^{4,8\ell+5}(\mathrm{Spin}(m))$ are $\mathbf{Z}_2$-vector spaces, and there can be no extension from filtration 2 to filtration 3 by naturality, the only way that α^* in (2.44) could hit an element in filtration 4 is if γ^* hits an element of order 4 in filtration 2, and there is a nontrivial extension. We will show that $(2\gamma)^*$ cannot be nonzero in filtration 2.

Since $\alpha^*(= (4\gamma)^*)$ is given by applying $[-, \Phi BSO(m)]$ to the top composite in (2.49), then $(2\gamma)^*$ is given by applying $[-, \Phi BSO(m)]$ to the bottom composite. The E_2-version of this bottom composite is just like (2.45) with $\mathbf{Z}/2$ replaced by $\mathbf{Z}/4$. Thus showing that $(2\gamma)^*$ is 0 in filtration 2 is equivalent to proving the analogue of (2.46) with $C^*/2$ replaced by $C^*/4$.

We need the following lemma.

[10] $v_1^{-1}T^{-1}$ can be defined to be $T^{-1} \wedge v_1^{-1}J$.

LEMMA (2.50). *The matrix, analogous to (2.38) in the interpretations of its rows and columns, which presents $H_0(C^*_{(4\ell+1)})$ for* Spin$(2n+1)$ *with $n > 5$ is*

$$\begin{pmatrix} 2^{A_1} & 0 \\ u_2 2^{A_2} & 2^n \\ u_3 2^n & 2^3 \end{pmatrix} \tag{2.51}$$

with u_i odd and $A_i \geq n+1$.

This is proved similarly to (2.36). It differs in that it involves $4\ell+1$ rather than $4\ell-1$. It is just [5, 3.18] with a lower bound for some exponents being 1 larger than was proved in [5]. As we don't need this refinement here, we will not present the details of the proof, which are extremely similar to those of (2.36).

Now the analogue of (2.46) with 4 instead of 2 is proved by the same method used for 2. Now we have that $\delta_0(x_1, x_2, x_3) \neq 0 \in \mathbf{s}H_0(C^*_{(-1)})$ if and only if the D-component of x_3 is not divisible by 4. Here we need that $A_i \geq n+1$ in (2.38) when $\ell = 0$, which was proved in (2.36). In this case, $\delta'_\ell(x_1, x_2, x_3)$ is nonzero in $H_0(C^*_{(4\ell+1)})$ because it is $\frac{1}{4}$ or $\frac{1}{2}$ times the last row of (2.51) plus $\frac{1}{4}$ times multiples of the other rows. This will be nonzero because of the 2^3 in the second column.

This completes the argument (for Case 11) when m is odd. If $m = 4a$ a similar argument works. A matrix of the same general form as (2.40) presents $H_0(C^*_{(4\ell+1)})$. Its rows and columns have analogous interpretations. As in the case m odd, the key point is a 2^3 which occurs in the last row, second column. This is due to the $(3^{m+1}-1)$-factor in [3, (4.27)]. The m of that paper is our $4\ell+1$. This 2^3 will cause (2.46) to hold, and with the 2 replaced by a 4, just as it did when m is odd.

Case 12: $k \equiv 1 \mod 4$, $m \equiv 2 \mod 4$. Similarly to Case 10, the method of Case 11 does not apply because the chain complex used there required $\psi^{-1} = -1$. Again, we can make the required deductions by naturality. The morphism $\mathbf{s}E_2^{0,-1}(8\ell+3, 4j+1) \to \mathbf{s}E_2^{0,-1}(8\ell+3, 4j+2)$ is bijective by [3, 3.3]. If j is odd, the generator of $E_2^{0,-1}(8\ell+3, 4j+1)$ is a permanent cycle by Case 11, and hence so is its image. Now let j be even. The same naturality argument shows that $d_2 = 0$ on $\mathbf{s}E_2^{0,-1}(8\ell+3, 4j+2)$. That $d_3 = 0$ is proved by the method of Case 10, using that $d_3 = 0$ on $\widetilde{E}_3^{1,-1}(\mathrm{Spin}(4j+2))$ by [6, 2.23]. Finally we consider d_4. We cannot use naturality from $E_4(8\ell+3, 4j+1)$ because it had a nonzero d_3 by [6, 2.23]. Instead we use the argument in Case 11, that the attaching map α equals 4γ. We use naturality from $E_2(8\ell+3, 4j+1)$ to see that $(2\gamma)^*$ must be zero in filtration 2, and deduce as in Case 11 that α^* is 0 in filtration 4.

3. Nonlifting results

In [9], the following result was proven.

THEOREM (3.1). *If u is odd and $2^{4b+\epsilon} > 4k+t$, then*

$$\mathrm{gd}(u2^{4b+\epsilon}\xi_{4k+t}) \geq 4k - 8b + d,$$

where d is given in the following table.

		ϵ			
		0	1	2	3
	1	0	-2	-2	-4
t	2	2	2	0	-4
	3	2	2	0	-4
	4	4	2	2	0

Several more nonlifting results could have been obtained by the same method. The author of [9] did not give careful enough consideration to P_b^t with $t \equiv 1 \bmod 4$ or $b \equiv 2 \bmod 4$. We sketch a proof of the following result. Theorems (3.1) and (3.2) together provide all the nonlifting results in Theorem (1.3), and those of [6, 1.1(2)].

THEOREM (3.2). *If u is odd and $2^{4b+\epsilon} > 4k+t$, then*

$$\mathrm{gd}(u2^{4b+\epsilon}\xi_{4k+t}) \geq 4k - 8b + \delta$$

if $(\epsilon, t, \delta) = (0, 2, 3)$, $(0, 3, 3)$, $(1, 4, 3)$, $(1, 1, 0)$, *or* $(0, 1, 2)$.

Proof. We must show there does not exist an axial map

$$P^{4k+t} \times P^{u2^{4b+\epsilon}-4k+8b-\delta} \to P^{u2^{4b+\epsilon}-1}.$$

This is done by showing that $\psi^3 - 1$ applied to the dual class in

$$ko_{-2}(P_{-4k-t-1}^{-2} \wedge P_{-u2^{4b+\epsilon}+4k-8b+\delta-1} \wedge P^{u2^{4b+\epsilon}-1}) \tag{3.3}$$

is nonzero. This class is called the axial class.

LEMMA (3.4). *Let $X = P_{-4k-t-1}^{-2} \wedge P_{-u2^{4b+\epsilon}+4k-8b+\delta-1}$. Then $ko_*(X \wedge P^{u2^{4b+\epsilon}-1})$ contains summands*

$$ko_*(X \wedge S^{u2^{4b+\epsilon}-1}) \oplus ko_*(X \wedge P^{u2^{4b+\epsilon}-2}).$$

The upper edge of the second of these summands extends one filtration higher than that of the first.

Proof. Let A_1 denote the subalgebra of the mod 2 Steenrod algebra generated by Sq^1 and Sq^2. We use that the Adams spectral sequence converging to $ko_*(X)$ has $E_2 = \mathrm{Ext}_{A_1}(H^*X)$. (We omit writing $\mathbf{Z}_2$ in the second variable.) Let N denote the A_1-module with classes in grading 0, 2, 3, and 5 with $\mathrm{Sq}^2\,\mathrm{Sq}^1\,\mathrm{Sq}^2 \neq 0$, and let N_0 be defined by the short exact sequence of A_1-modules

$$0 \to \Sigma^5\mathbf{Z}_2 \to N \to N_0 \to 0.$$

If M is an A_1-module which is free as a module over the subalgebra A_0 generated by Sq^1, then $\mathrm{Ext}_{A_1}(M \otimes N) = 0$ in filtration > 0, and hence, for $s > 0$, we have

$$\mathrm{Ext}_{A_1}^{s,t}(M \otimes \Sigma^4\mathbf{Z}_2) \approx \mathrm{Ext}_{A_1}^{s,t+1}(M \otimes \Sigma^5\mathbf{Z}_2) \xrightarrow{\approx} \mathrm{Ext}_{A_1}^{s+1,t+1}(M \otimes N_0). \tag{3.5}$$

The first of these groups can correspond roughly to the first summand of the lemma, and the last to the other summand, after adjoining many copies of $\mathrm{Ext}_{A_1}(M \otimes N)$. The filtration shift in (3.5) yields the conclusion of the lemma.

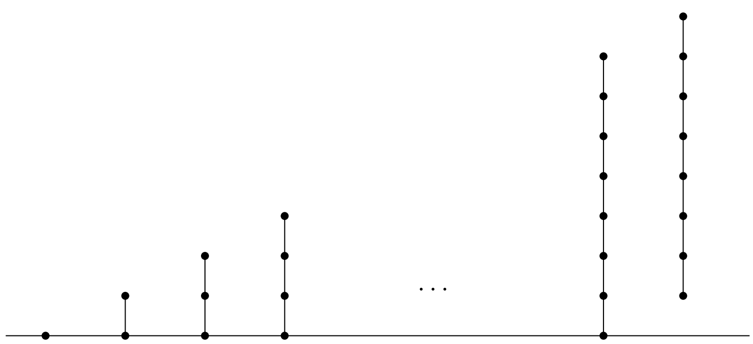
Here we have used that, except in its bottom few cells, the A_1-module $H^*P^{u2^{4b+\epsilon}-2}$ is built by short exact sequences from many copies of $\Sigma^i N$ and one of $\Sigma^{u2^{4b+\epsilon}-5}N_0$. A deviation due to the bottom few cells of $P^{u2^{4b+\epsilon}-2}$ will not alter the Ext groups in the region of interest. Note that H^*X is A_0-free except in the case where $t = 3 = \delta$, in which case it is a direct sum of an A_0-free summand and one that is inconsequential here. $\square$

Using some suspension isomorphisms, the part of (3.3) corresponding to the first summand in (3.4) is

$$ko_{-1}(P^{-2}_{-4k-t-1} \wedge P_{4k-8b+\delta-1}).$$

The subscript of one P is odd[11] and the other $\equiv 2 \bmod 4$. The $P_{4\ell+2}$ is built from copies of N, which, after tensoring with the other P, give no Ext in positive filtration, together with $\langle g_{4\ell+2}, \mathrm{Sq}^2(g_{4\ell+2})\rangle$, which changes bo to bu. Thus the chart for the portion of (3.4) due to the top cell is given by the diagram below, with the bottom class in dimension $-8b + \delta - t - 2$.

DIAGRAM (3.6).



All of our cases[12] have $\delta - t = 1 - 2\epsilon$. Thus the chart starts in $-8b - 2\epsilon - 1$, and its top element in dimension -1 is in filtration $4b + \epsilon$. The summand of (3.3) corresponding to the second summand of (3.4) has top element in filtration $4b + \epsilon + 1$.

According to the third case of Table 12 of [9], the axial class has a component $2 \cdot u2^{4b+\epsilon}$ in this second summand, i.e. at height $4b+\epsilon+1$, and so is nonzero. $\square$

Received June 27, 2005

Final version received September 23, 2005

MARTIN BENDERSKY
HUNTER COLLEGE, CUNY
NY, NY 10021
mbenders@shiva.hunter.cuny.edu

DONALD M. DAVIS
LEHIGH UNIVERSITY
BETHLEHEM, PA 18015
dmd1@lehigh.edu

[11] except for the case $(0, 3, 3)$, which is equivalent to $(0, 2, 3)$ plus an additional split summand
[12] with the exception noted in the previous footnote

References

[1] J. F. Adams, *Vector fields on spheres*, Annals of Math **75** (1962) 603–632.

[2] J. F. Adams, *On the groups $J(X)$, IV*, Topology **5** (1966) 21–71.

[3] M. Bendersky and D. M. Davis, *The v_1-periodic homotopy groups of SO(n)*, Memoirs Amer Math Soc **815** (2004).

[4] M. Bendersky and D. M. Davis, *2-primary v_1-periodic homotopy groups of SU(n)*, Amer Jour Math **114** (1991) 529–544.

[5] M. Bendersky and D. M. Davis, *The 1-line of the K-theory Bousfield-Kan spectral sequence for* $\mathrm{Spin}(2n+1)$, Contemp Math AMS **279** (2001) 37–56 .

[6] M. Bendersky, D. M. Davis, and M. Mahowald, *Stable geometric dimension of vector bundles over even-dimensional real projective spaces*, Trans Amer Math Soc. **358** (2005) 1585–1603. `http://www.lehigh.edu/~dmd1/sgd2.html`

[7] A. K. Bousfield, *The K-theory localization and v_1-periodic homotopy groups of finite H-spaces*, Topology **38** (1999) 1239–1264.

[8] A. K. Bousfield, *On the 2-primary v_1-periodic homotopy groups of spaces*, Topology **44** (2005) 381–413.

[9] D. M. Davis, *Generalized homology and the generalized vector field problem*, Quar Jour Math Oxford **25** (1974) 169–193.

[10] D. M. Davis, S. Gitler, and M. Mahowald, *The stable geometric dimension of vector bundles over real projective spaces*, Trans Amer Math Soc **268** (1981) 39–61.

[11] D. M. Davis and M. Mahowald, *Homotopy groups of some mapping telescopes*, Annals of Math Studies **113** (1987) 126–151.

[12] D. M. Davis and M. Mahowald, *v_1-localizations of torsion spectra and spherically resolved spaces*, Topology **32** (1993) 543–550.

[13] M. Mahowald, *The image of J in the EHP sequence*, Annals of Math **116** (1982) 65–112.

Bol. Soc. Mat. Mexicana (3) Vol. 12, 2006

TROTTER-KATO APPROXIMATIONS OF SEMILINEAR STOCHASTIC EVOLUTION EQUATIONS

Dedicated to Professor Onésimo Hernández-Lerma on his 60^{th} *birthday*

T. E. GOVINDAN

Abstract. This paper deals with a semilinear stochastic evolution equation in a real separable Hilbert space that is related to a McKean-Vlasov type measure-valued evolution equation. The main goal here is to study the Trotter-Kato approximations associated with mild solutions of such equations and also to deduce the convergence of the corresponding probability measures. As an application, we also investigate the dependence of such equations on a parameter.

1. Introduction

Consider the following stochastic process described by a semilinear Itô equation in a real separable Hilbert space H:

$$\begin{aligned} (1.1) \qquad dx(t) &= [Ax(t) + f(x(t), \mu(t))]dt + \sqrt{Q}dw(t), \qquad t \in [0, T], \\ \mu(t) &= \text{probability distribution of } x(t), \\ x(0) &= c, \end{aligned}$$

where $w(t)$ is a given H- valued cylindrical Wiener process; A is the infinitesimal generator of a strongly continuous semigroup $\{S(t) \colon t \geq 0\}$ of bounded linear operators on H; f is an appropriate H-valued function defined on $H \times M_{\lambda^2}(H)$, where $M_{\lambda^2}(H)$ denotes a proper subset of probability measures on H; Q is a positive, symmetric and bounded operator on H satisfying some conditions (see hypotheses (H1) and (H2)); and c is a given H-valued random variable. If the nonlinear term f in equation (1.1) does not depend on the probability distribution $\mu(t)$ of the process x at time t, then the solution process of equation (1.1) is a standard Markov process, and such equations are well studied, see Da Prato and Zabczyk [2] and the references therein. On the other hand, there are situations where the nonlinear drift term f depends not only on the state of the process at time t but also on the probability distribution of the process at that time as indicated in equation (1.1), we refer to McKean [7] and Ahmed and Ding [1] for details.

In fact, Ahmed and Ding [1] investigated the existence and uniqueness of a mild solution of equation (1.1) and also considered the Yoshida approximations [2], p. 382, of such solutions, among others. However, to the best of our knowledge nothing is known on the Trotter-Kato approximations, see Pazy [8],

2000 *Mathematics Subject Classification*: 60H15, 34K40.

Keywords and phrases: stochastic evolution equations, mild solutions, Trotter-Kato approximations, convergence of probability measures, dependence on a parameter.

p. 88, of equation (1.1). This, therefore, motivates us in this work to consider some sequential approximations, that is, Trotter-Kato approximations and its version, so-called the zeroth-order approximations, see Kannan and Bharucha-Reid [6] and Govindan [4] associated with mild solutions of equation (1.1). Using the latter, we shall provide an estimate of the error in the approximation. As an application of such approximations, we shall also investigate some limit theorems on the dependence of equation (1.1) on a parameter, see Gikhman and Skorokhod [3], pp. 50-54.

This paper is organized as follows: In Section 2, we give the preliminaries from [1] as we work in their framework. The Trotter-Kato approximation results are considered in Section 3. In Section 4, we study the dependence of such equations on a parameter. Finally, in Section 5, we consider some examples.

2. Preliminaries

Let $(\Omega, F, \{F_t\}_{t\geq 0}, P)$ denote a complete probability space equipped with a family of nondecreasing sub-sigma algebras $\{F_t\}_{t\geq 0}$. H is a real separable Hilbert space with the scalar product $< \cdot, \cdot >$ and norm $|\cdot|$. $B(H)$ denotes the Borel sigma algebra of subsets of H and $M(H)$ is the space of probability measures on $B(H)$ carrying the usual topology of weak convergence. $C(H)$ denotes the space of continuous functions on H. The notation $< \mu, \varphi >$ means $\int_H \varphi(x)\mu(dx)$ whenever this integral makes sense. Throughout this paper we let $\lambda(x) \equiv 1 + |x|$, $x \in H$, and define the Banach space

$$C_\rho(H) = \{\varphi \in C(H) : \| \varphi \|_{C_\rho(H)} \equiv \sup_{x\in H} \frac{|\varphi(x)|}{\lambda^2(x)} + \sup_{x\neq y} \frac{|\varphi(x) - \varphi(y)|}{|x-y|} < \infty\}.$$

For $p \geq 1$, let $M^s_{\lambda^p}(H)$ be the Banach space of signed measures m on H satisfying $\| m \|_{\lambda^p} \equiv \int_H \lambda^p(x)|m|(dx) < \infty$, where $|m| = m^+ + m^-$ and $m = m^+ - m^-$ is the Jordan decomposition of m. Let $M_{\lambda^2}(H) = M^s_{\lambda^2}(H) \cap M(H)$ be the set of probability measures with second moments on $B(H)$. We put on $M_{\lambda^2}(H)$ a topology induced by the following metric:

$$\rho(\mu, \nu) = \sup\{< \varphi, \mu - \nu >: \varphi \in C_\rho(H) \text{ with } \| \varphi \|_{C_\rho(H)} \leq 1\}.$$

Then $(M_{\lambda^2}(H), \rho)$ forms a complete metric space. Note that the topology just introduced is stronger than the usual topology of weak convergence, see [1]. We denote by $C([0, T], (M_{\lambda^2}(H), \rho))$ the complete metric space of continuous functions from $[0, T]$ to $(M_{\lambda^2}(H), \rho)$ with the metric:

$$D_T(\mu, \nu) = \sup_{t\in[0,T]} \{\rho(\mu(t), \nu(t)) : \mu, \nu \in C([0, T], (M_{\lambda^2}(H), \rho))\}.$$

Let $C([0, T]; L^2(\Omega, F, P; H))$ be the Banach space of continuous maps from $[0, T]$ into $L^2(\Omega, F, P; H))$ satisfying the condition $\sup_{t\in[0,T]} E|x(t)|^2 < \infty$. Let $\wedge_2$ be the closed subspace of $C([0, T]; L^2(\Omega, F, P; H))$ consisting of measurable and F_t-adapted processes $x = \{x(t): t \in [0, T]\}$. Then $\wedge_2$ is a Banach space with the norm topology given by $|x|_{\wedge_2} = (\sup_{t\in[0,T]} E|x(t)|^2)^{1/2}$.

Definition (2.1). a) A stochastic process $x = \{x(t): t \in [0, T]\}$ defined on the probability space (Ω, F, P) is said to be a mild solution of equation (1.1) if the following conditions are satisfied:

(i) $x(t, \omega)$ is Borel measurable as a function from $[0, T] \times \Omega$ to H and $x(t)$ is F_t-measurable for each $t \in [0, T]$;

(ii) $E|x(t)|^2 < \infty$ for all $t \in [0, T]$;

(iii) x satisfies the following integral equation:

$$x(t) = S(t)c + \int_0^t S(t-s)f(x(s), \mu(s))ds + \int_0^t S(t-s)\sqrt{Q}dw(s), \quad P\text{-}a.s. \tag{2.2}$$

b) A stochastic process $x = \{x(t): t \in [0, T]\}$ defined on the probability space (Ω, F, P) is said to be a strong solution of equation (1.1) if $x \in D(A)$, domain of A, P-a.s., $\int_0^T |Ax(t)|^2 dt < \infty$, P-a.s., and for $t \in [0, T]$

$$x(t) = c + \int_0^t [Ax(s) + f(x(s), \mu(s))]ds + \int_0^t \sqrt{Q}dw(s), \qquad P\text{-}a.s. \tag{2.3}$$

We now make the following assumptions, see [1]:

Assumptions (2.4). (H1) (i) A is the infinitesimal generator of a C_0-semigroup $\{S(t): t \geq 0\}$ of bounded linear operators on H of negative type:

$$\| S(t) \| \leq M \exp(-\alpha t), \quad t \geq 0$$

for some positive constants $M \geq 1$ and α, where $\| \cdot \|$ denotes the operator norm;

(ii) Q is a positive, symmetric and bounded operator in H such that the operator Q_t defined by

$$Q_t = \int_0^t S(r)QS^*(r)dr$$

is nuclear for all $t \geq 0$ and $\sup_{t\geq 0} \operatorname{tr} Q_t < \infty$;

(iii) w is an H-valued cylindrical Wiener process defined on (Ω, F, P) with (the incremental) covariance operator an identity operator.

(H2) $f: H \times (M_{\lambda^2}(H), \rho) \to H$ satisfies the Lipschitz and linear growth conditions:

$$|f(x, \mu) - f(y, \nu)| \leq k(|x-y| + \rho(\mu, \nu)), \qquad P\text{-}a.s.$$
$$|f(x, \mu)| \leq l(1 + |x| + \| \mu \|_\lambda), \qquad P\text{-}a.s.$$

where k and l are positive constants.

Remark (2.5). For the interpretations of the hypotheses (H1) and (H2) and the details, we refer to [1]. In (H2), $\| \mu \|_\lambda$ is as defined before with $p = 1$.

Ahmed and Ding [1] proved the following result.

THEOREM (2.6). *Suppose that the conditions (H1) and (H2) are satisfied. Then:*

(a) *For each given initial data c, a random variable in H with probability distribution $\nu \in M_{\lambda^2}(H)$, equation (1.1) has a unique mild solution $x = \{x(t): t \in [0, T]\}$ in $\wedge_2$ and its probability distribution $\mu = \{\mu(t): t \in [0, T]\}$ belongs to $C([0, T], (M_{\lambda^2}(H), \rho))$.*

(b) *For any $p \geq 1$, we have*

$$\sup_{t\in[0,T]} E|x(t)|^{2p} \leq k_{p,T}(1 + E|c|^{2p}),$$

where $k_{p,T}$ is a positive constant.

(c) *There exists a sequence $\{x_\eta\}_{\eta=1}^{\infty}$ of strong solutions of the Yoshida approximating systems such that $\{x_\eta\}_{\eta=1}^{\infty}$ converges to x in $C([0,T]; L^2(\Omega, F, P; H))$ as $\eta \to \infty$.*

COROLLARY (2.7). [1] *The sequence of probability laws $\{\mu_\eta\}_{\eta=1}^{\infty}$ corresponding to $\{x_\eta\}_{\eta=1}^{\infty}$ converges to the probability law μ of x in $C([0,T], (M_{\lambda^2}(H), \rho))$ as $\eta \to \infty$.*

3. Trotter-Kato Approximations

In this section, we shall establish the Trotter-Kato approximation results analogous to Theorem (2.6)(c) and Corollary (2.7).

Consider the family of stochastic semilinear evolution equations

$$\begin{aligned} dx_n(t) &= [A_n x_n(t) + f(x_n(t), \mu_n(t))]dt + \sqrt{Q}dw(t), \qquad t \in [0,T], \\ \mu_n(t) &= \text{probability distribution of } x_n(t), \\ x_n(0) &= c, \end{aligned} \tag{3.1}$$

where A_n, $n = 1, 2, 3, \dots$ is the infinitesimal generator of a strongly continuous semigroup $\{S_n(t)\colon t \geq 0\}$ of bounded linear operators on H.

For each $n \geq 1$, by Theorem (2.6) (a), equation (3.1) has a unique mild solution $x_n \in C([0,T]; L^2(\Omega, F, P; H))$ with the probability measure $\mu_n \in C([0,T], (M_{\lambda^2}(H), \rho))$. Hence $x_n(t)$ satisfies the stochastic integral equation

$$\begin{aligned} x_n(t) = S_n(t)c &+ \int_0^t S_n(t-s) f(x_n(s), \mu_n(s))\,ds \\ &+ \int_0^t S_n(t-s)\sqrt{Q}dw(s), \qquad P\text{-}a.s. \end{aligned} \tag{3.2}$$

In the sequel, we will use the notation $A \in G(M, \alpha)$ for an operator A which is the infinitesimal generator of a C_0-semigroup $\{S(t)\colon t \geq 0\}$ of bounded linear operators on H satisfying $\| S(t) \| \leq M \exp(\alpha t)$, $t \geq 0$ for some positive constants $M \geq 1$ and α.

We now make the following assumptions, see Pazy [8], Theorem 4.5, p.88:

(H3) (i) Let $A_n \in G(M, \alpha)$ for each $n = 1, 2, \dots$;

(ii) As $n \to \infty$, $A_n x \to Ax$ for every $x \in D$, where D is a dense subset of H;

(iii) There exists a γ with $\operatorname{Re}\gamma > \alpha$ for which $(\gamma I - A)D$ is dense in H, then the closure $\bar{A}$ of A is in $G(M, \alpha)$.

THEOREM (3.3). *Suppose that the conditions (H1)-(H3) are satisfied. Let $x_n(t)$ and $x(t)$ be the mild solutions (3.2) and (2.2), respectively. Then for each $T > 0$,*

$$\sup_{0 \leq t \leq T} E|x_n(t) - x(t)|^2 \to 0, \qquad as \qquad n \to \infty.$$

Proof. Considering the equality

$$\begin{aligned} x_n(t) - x(t) &= [S_n(t) - S(t)]c \\ &+ \int_0^t [S_n(t-s) f(x_n(s), \mu_n(s)) - S(t-s) f(x(s), \mu(s))]ds \end{aligned}$$

$$+\int_0^t [S_n(t-s)-S(t-s)]\sqrt{Q}]dw(s), \qquad P\text{-}a.s.$$

we obtain

$$|x_n(t)-x(t)|^2 \le 4\{|S_n(t)c-S(t)c|^2$$

$$+|\int_0^t S_n(t-s)[f(x_n(s),\mu_n(s))-f(x(s),\mu(s))]ds|^2$$

$$+|\int_0^t [S_n(t-s)-S(t-s)]f(x(s),\mu(s))ds|^2$$

$$+|\int_0^t [S_n(t-s)-S(t-s)]\sqrt{Q}dw(s)|^2\}, \qquad P\text{-}a.s. \tag{3.4}$$

Since $A_n \in G(M,\alpha)$ for each $n = 1,2,3,\dots$ and $\bar{A} \in G(M,\alpha)$, $E|[S_n(t)-S(t)]c| \le 2M\exp(\alpha T)E|c|$, uniformly in n and $t \in [0,T]$, where $\{S(t)\colon t \ge 0\}$ is the C_0-semigroup generated by $\bar{A}$. Therefore, by hypothesis, a version of the Trotter-Kato theorem [8], p. 88, yields

$$\sup_{0\le t\le T} E|S_n(t)c-S(t)c|^2 \to 0, \qquad \text{as} \qquad n\to\infty, \tag{3.5}$$

for all $t \ge 0$, $c \in H$ and the limit in (3.5) is uniform in t for t in bounded intervals.

By assumption (H2) and the inequality $\rho^2(\mu_n(s),\mu(s)) \le E|x_n(s)-x(s)|^2$, a consequence of the definition of ρ and Jensen's inequality, we have

$$\sup_{0\le s\le t} E|\int_0^s S_n(s-r)[f(x_n(r),\mu_n(r))-f(x(r),\mu(r))]dr|^2$$

$$\le 2Tk^2\int_0^t \| S_n(t-s)\|^2 \, [E|x_n(s)-x(s)|^2+\rho^2(\mu_n(s),\mu(s))]ds$$

$$\le 4Tk^2M^2\exp(2\alpha T)\int_0^t E|x_n(s)-x(s)|^2 ds\,. \tag{3.6}$$

Next, by Lemma 7.7, [2], p. 194

$$\sup_{0\le s\le t} E|\int_0^s [S_n(s-r)-S(s-r)]\sqrt{Q}dw(r)|^2$$

$$\le \int_0^t E\,\|\,[S_n(t-s)-S(t-s)]\sqrt{Q}\,\|^2_{L_2^0}\, ds\,, \tag{3.7}$$

where $\|\cdot\|_{L_2^0}$ denotes the Hilbert-Schmidt norm.

Using the estimates (3.5)-(3.7), inequality (3.4) reduces to

$$\sup_{0\le s\le t} E|x_n(s)-x(s)|^2 \le \beta(n,T)+16Tk^2M^2\exp(2\alpha T)\int_0^t \sup_{0\le r\le s} E|x_n(r)-x(r)|^2 ds,$$

where

$$\beta(n,T) = 4\sup_{0\le t\le T} E|S_n(t)c-S(t)c|^2+4T\int_0^t E|[S_n(t-s)-S(t-s)]f(x(s),\mu(s))|^2 ds$$

$$+4\int_0^t E \parallel [S_n(t-s)-S(t-s)]\sqrt{Q} \parallel_{L_2^0}^2 ds. \tag{3.8}$$

An application of Gronwall's lemma yields

$$\sup_{0\le s\le t} E|x_n(s)-x(s)|^2 \le \beta(n,T)\exp(16Tk^2M^2\exp(2\alpha T)t), \qquad t\in[0,T].$$

The first term on the right-hand side of (3.8) tends to zero as $n\to\infty$ by (3.5). The second term also tends to zero in view of (3.5) together with the Lebesgue's dominated convergence theorem. Concerning the third term, note that

$$\int_0^t E \parallel [S_n(t-s)-S(t-s)]\sqrt{Q} \parallel_{L_2^0}^2 ds \le 2M^2\exp(2\alpha T)\int_0^t E \parallel \sqrt{Q} \parallel_{L_2^0}^2 ds < \infty.$$

Hence by (3.5) and Lebesgue's dominated convergence theorem this term also tends to zero. Thus $\beta(n,T)\to 0$ as $n\to\infty$. This completes the proof. □

COROLLARY (3.9). *The sequence of probability laws $\{\mu_n\}_{n=1}^\infty$ corresponding to $\{x_n\}_{n=1}^\infty$ converges to the probability law μ of x in $C([0,T],(M_{\lambda^2}(H),\rho))$ as $n\to\infty$.*

Proof. This follows from the fact that

$$D_T(\mu_n,\mu) = \sup_{t\in[0,T]}\rho(\mu_n(t),\mu(t)) \le \sup_{t\in[0,T]} E|x_n(t)-x(t)|^2.$$

Let us next consider the zeroth-order approximations. Consider the stochastic semilinear evolution equation

$$\begin{aligned} dx_\varepsilon(t) &= [A_\varepsilon x_\varepsilon(t)+f(x_\varepsilon(t),\mu_\varepsilon(t))]dt+\sqrt{Q}dw(t), \quad t\in[0,T], \\ \mu_\varepsilon(t) &= \text{probability distribution of } x_\varepsilon(t), \\ x_\varepsilon(0) &= c, \end{aligned} \tag{3.10}$$

where $A_\varepsilon(\varepsilon>0)$ is the infinitesimal generator of a strongly continuous semigroup $\{S_\varepsilon(t)\colon t\ge 0\}$ of bounded linear operators on H.

For each $\varepsilon>0$, one can show by Theorem (2.6)(a) that equation (3.10) has a unique mild solution $x_\varepsilon\in C([0,T];L^2(\Omega,F,P;H))$ with $\mu_\varepsilon\in C([0,T],(M_{\lambda^2}(H),\rho))$ satisfying

$$x_\varepsilon(t) = S_\varepsilon(t)c+\int_0^t S_\varepsilon(t-s)f(x_\varepsilon(s),\mu_\varepsilon(s))ds$$

$$+\int_0^t S_\varepsilon(t-s)\sqrt{Q}dw(s), \qquad P\text{-}a.s. \tag{3.11}$$

Assume now the following condition [4], [6]:

(H4) Let $A_\varepsilon, A\in G(M,\alpha)(\varepsilon>0)$ with $D(A_\varepsilon)=D(A)(\varepsilon>0)$; and $S_\varepsilon(t)\to S(t)$ as $\varepsilon\downarrow 0$, uniformly in $t\in[0,T]$ for each $T>0$. □

In the following result, we shall estimate the error in the approximation. Note that the proof follows mimicking arguments from Theorem (3.3).

THEOREM (3.12). *Suppose that the conditions (H1)-(H2) and (H4) are satisfied. Let $x_\varepsilon(t)$ and $x(t)$ be the mild solutions (3.11) and (2.2), respectively. Then*

$$E|x_\varepsilon(t) - x(t)|^2 \leq \psi(\varepsilon)\phi(t), \qquad t \in [0,T],$$

where $\phi(t)$ is a positive exponentially increasing function and $\psi(\varepsilon)$ is a positive function decreasing monotonically to zero as $\varepsilon \downarrow 0$.

Proof. Consider

(3.13)

$$\begin{aligned} x_\varepsilon(t) - x(t) = [S_\varepsilon(t) - S(t)]c &+ \int_0^t S_\varepsilon(t-s)[f(x_\varepsilon(s),\mu_\varepsilon(s)) - f(x(s),\mu(s))]ds \\ &+ \int_0^t [S_\varepsilon(t-s) - S(t-s)]f(x(s),\mu(s))ds \\ &+ \int_0^t [S_\varepsilon(t-s) - S(t-s)]\sqrt{Q}dw(s), \quad P\text{-}a.s. \end{aligned}$$

We estimate each term on the right-hand side of (3.13):

Since $S_\varepsilon(t) \to S(t)$ as $\varepsilon \downarrow 0$, uniformly in $t \in [0,T]$, there exist an $\varepsilon_1 > 0$ and some constant $K_1 > 0$ such that

(3.14) $$E|S_\varepsilon(t)c - S(t)c|^2 \leq K_1 a_1(\varepsilon), \qquad \text{for all} \qquad t \in [0,T],$$

where $0 < a_1(\varepsilon) \downarrow 0$ as $\varepsilon_1 > \varepsilon \downarrow 0$.

From (3.6), we have

$$\begin{aligned} &E|\int_0^t S_\varepsilon(t-s)[f(x_\varepsilon(s),\mu_\varepsilon(s)) - f(x(s),\mu(s))]ds|^2 \\ &\leq \; 4Tk^2M^2\exp(2\alpha T)\int_0^t E|x_\varepsilon(s) - x(s)|^2 ds. \end{aligned}$$

Proceeding as above in showing (3.14), there exist $\varepsilon_2 > 0$ and some constant $K_2 > 0$ such that

$$E|\int_0^t [S_\varepsilon(t-s) - S(t-s)]f(x(s),\mu(s))ds|^2 \leq T\int_0^t K_2 a_2(\varepsilon)ds \leq T^2 K_2 a_2(\varepsilon),$$

where $0 < a_2(\varepsilon) \downarrow 0$ as $\varepsilon_2 > \varepsilon \downarrow 0$.

Finally, there exist $\varepsilon_3 > 0$ and some constant $K_3 > 0$ such that

$$\begin{aligned} &E|\int_0^t [S_\varepsilon(t-s) - S(t-s)]\sqrt{Q}dw(s)|^2 \\ \leq \; &\int_0^t E \parallel [S_\varepsilon(t-s) - S(t-s)]\sqrt{Q} \parallel_{L_2^0}^2 ds \leq \int_0^t K_3 a_3(\varepsilon)ds \leq TK_3 a_3(\varepsilon), \end{aligned}$$

where $0 < a_3(\varepsilon) \downarrow 0$ as $\varepsilon_3 > \varepsilon \downarrow 0$.

Consequently, for $\varepsilon_0 > \varepsilon > 0$, where $\varepsilon_0 = \min\{\varepsilon_i, i = 1,2,3\}$,

$$E|x_\varepsilon(t) - x(t)|^2 \leq \psi(\varepsilon) + 16Tk^2M^2\exp(2\alpha T)\int_0^t E|x_\varepsilon(s) - x(s)|^2 ds,$$

where $\psi(\varepsilon) = 4\{K_1 a_1(\varepsilon) + T(TK_2 a_2(\varepsilon) + K_3 a_3(\varepsilon))\}$.

Invoking Gronwall's lemma, one obtains

$$E|x_\varepsilon(t) - x(t)|^2 \leq \psi(\varepsilon)\phi(t), \qquad t \in [0,T],$$

where $\phi(t) = \exp(16Tk^2M^2\exp(2\alpha T)t)$. □

COROLLARY (3.15). *The probability law* $\mu_\varepsilon(\varepsilon > 0)$ *corresponding to* x_ε *converges to the probability law* μ *of* x *in* $C([0,T], (M_{\lambda^2}(H), \rho))$ *as* $\varepsilon \downarrow 0$.

4. Dependence of the equation on a parameter

In this section, as an application of the results in Section 3, we consider a classical limit theorem on the dependence of the stochastic evolution equation (1.1) on a parameter. For this, we shall closely follow [3], pp. 50-54.

Consider the family of stochastic semilinear evolution equations

$$\begin{aligned} dx_n(t) &= [A_n x_n(t) + f_n(x_n(t), \mu_n(t))]dt + \sqrt{Q_n}dw(t), \quad t \in [0,T], \\ \mu_n(t) &= \text{probability distribution of } x_n(t), \\ x_n(0) &= c, \end{aligned} \tag{4.1}$$

where $A_n, n = 1, 2, 3, \ldots$ is the infinitesimal generator of a strongly continuous semigroup $\{S_n(t) \colon t \geq 0\}$ of bounded linear operators on H.

Let A_n, $f_n(x,\mu)$ and Q_n satisfy the conditions of Theorem (2.6) for each $n = 1, 2, \ldots$ with the same constants k, l. Then, equation (4.1) for each $n = 1, 2, \ldots$ has a unique mild solution $x_n \in C([0,T]; L^2(\Omega, F, P; H))$ with the probability measure $\mu_n \in C([0,T], (M_{\lambda^2}(H), \rho))$. Hence $x_n(t)$ satisfies the stochastic integral equation

$$\begin{aligned} x_n(t) = S_n(t)c &+ \int_0^t S_n(t-s) f_n(x_n(s), \mu_n(s))ds \\ &+ \int_0^t S_n(t-s)\sqrt{Q_n}dw(s), \qquad P\text{-}a.s. \end{aligned} \tag{4.2}$$

We now make the following assumptions, see [3], p. 52:

(H5) (i) for each $N > 0$ and $\varepsilon > 0$:

$$P\{\sup_{|x|\leq N}(|f_n(x,\mu) - f(x,\mu)|) > \varepsilon\} \to 0 \qquad \text{as} \qquad n \to \infty;$$

(ii) $Q_n \to Q$.

THEOREM (4.3). *Suppose that the conditions (H3) and (H5) are satisfied. Let* $x_n(t)$ *and* $x(t)$ *be the mild solutions (4.2) and (2.2), respectively. Then for each* $T > 0$:

$$\sup_{0\leq t\leq T} E|x_n(t) - x(t)|^2 \to 0, \qquad as \qquad n \to \infty.$$

Proof. Consider

$$\begin{aligned} x_n(t) - x(t) = \Psi(t) &+ \int_0^t S_n(t-s)[f_n(x_n(s), \mu_n(s)) - f_n(x(s), \mu_n(s))]ds \\ &+ \int_0^t S_n(t-s)[f_n(x(s), \mu_n(s)) - f_n(x(s), \mu(s))]ds, \qquad P\text{-}a.s., \end{aligned}$$

where

$$\Psi(t) = [S_n(t) - S(t)]c + \int_0^t [S_n(t-s) - S(t-s)] f(x(s), \mu(s))ds$$

$$+\int_0^t [S_n(t-s)-S(t-s)]\sqrt{Q}]dw(s)$$

$$+\int_0^t S_n(t-s)[f_n(x(s),\mu(s))-f(x(s),\mu(s))]ds$$

$$+\int_0^t S_n(t-s)[\sqrt{Q_n}-\sqrt{Q}]dw(s)\,. \tag{4.4}$$

By assumption (H2) for $L=6Tk^2M^2\exp(2\alpha T)$, we have

$$E|x_n(t)-x(t)|^2 \le 3E|\Psi(t)|^2 + L\int_0^t E|x_n(s)-x(s)|^2 ds.$$

Hence, by Lemma 1 [3], p. 41

$$E|x_n(t)-x(t)|^2 \le 3E|\Psi(t)|^2 + L\int_0^t e^{L(t-s)}E|\Psi(s)|^2 ds.$$

Hence to prove the theorem, it is sufficient to show that $\sup_{0\le t\le T} E|\Psi(t)|^2 \to 0$. The first three terms on the right-hand side of (4.4) tend to zero as shown earlier. To show that the remaining terms also go to zero, consider

$$\begin{aligned} &E\Big|\int_0^t S_n(t-s)[f_n(x(s),\mu(s))-f(x(s),\mu(s))]ds\Big|^2 \\ &\le tM^2e^{2\alpha t}E\int_0^t |f_n(x(s),\mu(s))-f(x(s),\mu(s))|^2 ds \\ &\le 4T^2l^2M^2e^{2\alpha T}[1+k_{1,T}(1+E|c|^2)+\sup_{0\le t\le T}\|\,\mu(t)\,\|_\lambda^2]<\infty. \end{aligned}$$

Hence by Assumption (H5)(i) and Lebesgue's theorem

$$\begin{aligned} &\sup_{0\le t\le T} E\Big|\int_0^t S_n(t-s)[f_n(x(s),\mu(s))-f(x(s),\mu(s))]ds\Big|^2 \\ &\le TM^2e^{2\alpha T}E\int_0^T |f_n(x(s),\mu(s))-f(x(s),\mu(s))|^2 ds \to 0. \end{aligned}$$

Using Lemma 7.7 [2], p. 194 and Assumption (H5)(ii), it follows that

$$\sup_{0\le t\le T} E\Big|\int_0^t S_n(t-s)[\sqrt{Q_n}-\sqrt{Q}]dw(s)\Big|^2 \le M^2e^{2\alpha T}\int_0^t \|\,\sqrt{Q_n}-\sqrt{Q}\,\|_{L_2^0}^2\,ds \to 0.$$

This completes the proof. □

COROLLARY (4.5). *Assume that the coefficients in equation (1.1) depend on a parameter θ which varies through some set of numbers G_1:*

$$\begin{aligned} dx_\theta(t) &= [A_\theta x_\theta(t)+f_\theta(x_\theta(t),\mu_\theta(t))]dt+\sqrt{Q_\theta}dw(t), \quad t\in[0,T], \\ \mu_\theta(t) &= \text{probability distribution of } x_\theta(t), \\ x_\theta(0) &= c\,, \end{aligned} \tag{4.6}$$

where A_θ is the infinitesimal generator of a strongly continuous semigroup $\{S_\theta(t)\colon t\ge 0\}$ of bounded linear operators on H. Assume further that for each $\varepsilon>0$, $N>0$,

$$P\{\sup_{|x|\le N}(|f_\theta(x,\mu)-f_{\theta_0}(x,\mu)|)>\varepsilon\}\to 0 \qquad as \qquad \theta\to\theta_0.$$

Furthermore, let $A_\theta, \theta \in G_1$ and $A_{\theta_0} \in G(M, \alpha)$ with $D(A_\theta) = D(A_{\theta_0})$ and $S_\theta(t) \to S_{\theta_0}(t)$ as $\theta \to \theta_0$, uniformly in $t \in [0, T]$ for each $T > 0$. Lastly, let A_θ, $f_\theta(x, \mu)$ and Q_θ for each θ satisfy the conditions of Theorem (2.6) with the same constants k, l. Then equation (4.6) has a unique mild solution x_θ and satisfies for each $T > 0$:

$$\sup_{0 \le t \le T} E|x_\theta(t) - x_{\theta_0}(t)|^2 \to 0, \qquad as \qquad \theta \to \theta_0.$$

Proof. This follows immediately from an application of Theorem (4.3) to the sequence $\{x_{\theta_n}(t)\}$, where $\theta_n \to \theta$. □

5. Examples

In this section, we consider some examples to illustrate the theory.

Example (5.1). This example is adapted from Da Prato and Zabczyk [2], A.5.4 and Example 5.8. Let us assume for simplicity that the set Θ is bounded and let us consider the wave equation with Dirichlet boundary conditions:

$$\begin{aligned} y_{tt}(t, \psi) &= \Delta_\psi y(t, \psi), \qquad t \ge 0, \psi \in \Theta, \\ y(t, \psi) &= 0, \qquad t > 0, \psi \in \partial\Theta, \\ y(0, \psi) = x_0(\psi), y_t(0, \psi) &= x_1(\psi), \qquad \psi \in \Theta. \end{aligned} \tag{5.2}$$

To write this problem in the abstract form, denote by $\wedge$ the positive self-adjoint operator

$$D(\wedge) = H^2(\Theta) \cap H_0^2(\Theta),$$
$$\wedge y = -\Delta_\psi y \ \text{ for all } \ y \in D(\wedge),$$

and introduce the Hilbert space $H = D(\wedge)^{1/2} \oplus X_2$. For details, see [2], p.402. We define in H the linear operator A as

$$A = \begin{pmatrix} 0 & I \\ -\wedge & 0 \end{pmatrix}$$

with domain $D(A) = D(\wedge) \oplus D(\wedge^{1/2})$. In fact, A is the infintesimal generator of a contraction semigroup $\{S(t)\colon t \ge 0\}$ in H, that is, $\| S(t) \| \le 1, t \ge 0$, given by

$$S(t) = \begin{pmatrix} \cos(\sqrt{\wedge} t) & \frac{1}{\sqrt{\wedge}} \sin(\sqrt{\wedge} t) \\ -\sqrt{\wedge} \sin(\sqrt{\wedge} t) & \cos(\sqrt{\wedge} t) \end{pmatrix}.$$

Consider now a stochastic wave equation in the abstract setting

$$dx(t) = [Ax(t) + \frac{x(t) + \mu(t)}{1 + x(t)}]dt + \sqrt{Q} dw(t), \qquad t \in [0, T], \tag{5.3}$$

with the initial condition as given above; and $w(t)$, $\mu(t)$ and Q are as defined before. Assumption (H1)(ii) holds provided

$$\int_0^T tr[\frac{\sin^2(\sqrt{\wedge} s)}{\wedge} Q] ds < \infty, \qquad \text{and} \tag{5.4}$$

$$\int_0^T tr[\frac{\cos^2(\sqrt{\wedge} s)}{\wedge} Q] ds < \infty\,. \tag{5.5}$$

From [2], the conditions (5.4)-(5.5) are satisfied if and only if $\mathrm{tr} \wedge^{-1} < +\infty$. One can easily check that the Assumption (H2) holds. Therefore, it follows from Theorem (2.6) that there exists a sequence $\{x_\eta\}_{\eta=1}^{\infty}$ of strong solutions of the Yoshida approximating systems such that $\{x_\eta\}_{\eta=1}^{\infty}$ converges to x in $C([0,T]; L^2(\Omega, F, P; H))$ as $\eta \to \infty$. It also follows from Corollary (2.7) that the sequence of probability laws $\{\mu_\eta\}_{\eta=1}^{\infty}$ corresponding to $\{x_\eta\}_{\eta=1}^{\infty}$ converges to the probability law μ of x in $C([0,T], (M_{\lambda^2}(H), \rho))$ as $\eta \to \infty$.

Example (5.6). This example is adapted from Govindan [5]. Consider the heat equation

$$\frac{\partial z(t,x)}{\partial t} = \frac{\partial^2}{\partial x^2} z(t,x), \qquad t \in [0,T], \tag{5.7}$$

$$z(t,0) = z(t,\pi) = 0, \qquad t \in [0,T],$$

$$z(s,x) = \phi(s,x), \phi(.,x) \in C[0,T], \phi(s,.) \in L^2[0,\pi], \qquad 0 \le x \le \pi.$$

Define $A\colon H \to H$, where $H = L^2[0,\pi]$ by $A = \partial^2/\partial x^2$ with domain $D(A) = \{w \in H\colon w, \partial w/\partial x$ are absolutely continuous, $\partial^2 w/\partial x^2 \in H, w(0) = w(\pi) = 0\}$. Then

$$Aw = \sum_{n=1}^{\infty} n^2 (w, w_n) w_n, \qquad w \in D(A),$$

where $w_n(x) = \sqrt{2/\pi}\sin nx$, $n = 1, 2, 3, \ldots$, is the orthonormal set of eigenvectors of A. It is well known that A is the infinitesimal generator of a C_0-semigroup $\{S(t)\colon t \ge 0\}$ in H, and is given by (see [5] and the references therein)

$$S(t)w = \sum_{n=1}^{\infty} \exp(-n^2 t)(w, w_n) w_n, \qquad w \in H,$$

that satisfies $\| S(t) \| \le \exp(-\pi^2 t)$, $t \ge 0$, and hence is a contraction semigroup.

Consider the stochastic heat equation:

$$dx(t) = [Ax(t) + \frac{x(t) + \mu(t)}{1 + x(t)}]dt + \sqrt{Q}dw(t), \qquad t \in [0,T], \tag{5.8}$$

with the initial condition as given above; and $w(t)$, $\mu(t)$ and Q are as defined before in Example (5.1). Define now $A_\varepsilon(\varepsilon > 0)$ by $A_\varepsilon = (1+\varepsilon)\partial^2/\partial x^2$ which is the infinitesimal generator of a of a C_0-semigroup $\{S_\varepsilon(t)\colon t \ge 0\}(\varepsilon > 0)$ in H, and is given by

$$S_\varepsilon(t)w = \sum_{n=1}^{\infty} \exp(-(1+\varepsilon)n^2 t)(w, w_n) w_n, \qquad w \in H,$$

that satisfies $\| S_\varepsilon(t) \| \le \exp(-(1+\varepsilon)\pi^2 t)$, $t \ge 0$, $\varepsilon > 0$ and hence is a contraction semigroup. Clearly,

$$\lim_{\varepsilon \downarrow 0} S_\varepsilon(t)x = S(t)x, \qquad x \in H,$$

uniformly in $t \in [0,T]$. Hence, by Theorem (3.12),

$$E|x_\varepsilon(t) - x(t)|^2 \le \psi(\varepsilon)\phi(t), \qquad t \in [0,T].$$

Acknowledgement

The author thanks sincerely the anonymous referees for their valuable comments that led to the improvement of the paper.

Received October 27, 2005

Final version received March 02, 2006

DEPARTAMENTO DE MATEMÁTICAS,
ESCUELA SUPERIOR DE FÍSICA Y MATEMÁ TICAS DEL IPN
EDIFICIO 9, U. P. ZACATENCO
07730 MÉXICO, D.F.
MÉXICO
tegovindan@yahoo.com

REFERENCES

[1] N. U. AHMED AND X. DING, *A semilinear McKean-Vlasov stochastic evolution equation in Hilbert space*, Stochastic Proc. Appl. **60** (1995), 65–85.

[2] G. DA PRATO AND J. ZABCZYK, Stochastic Equations in Infinite Dimensions, Cambridge University Press, Cambridge, 1992.

[3] I. I. GIKHMAN AND A. V. SKOROKHOD, Stochastic Differential Equations, Springer-Verlag, Berlin, 1972.

[4] T. E. GOVINDAN, *Autonomous semilinear stochastic Volterra integrodifferential equations in Hilbert spaces*, Dynamic Systems Appl. **3** (1994), 51–74.

[5] T. E. GOVINDAN, *Almost sure exponential stability for stochastic neutral partial functional differential equations*, Stochastics **77** (2005), 139–154.

[6] D. KANNAN AND A. T. BHARUCHA-REID, *On a stochastic integrodifferential evolution equation of Volterra type*, J. Integral Eqns. **10** (1985), 351–379.

[7] H. P. MCKEAN, *A class of Markov processes associated with nonlinear parabolic equations*, Proc. N.A.S. **56** (1966), 1907–1911.

[8] A. PAZY, Semigroups of Linear Operators and Applications to Partial Differential Equations, Springer-Verlag, New York, 1983.

Bol. Soc. Mat. Mexicana (3) Vol. 12, 2006

DIFFERENTIATION MATRICES FOR MEROMORPHIC FUNCTIONS

RAFAEL G. CAMPOS AND CLAUDIO MENESES

ABSTRACT. A procedure for obtaining differentiation matrices is extended straightforwardly to yield new differentiation matrices useful for derivatives of complex rational functions. Such matrices can be used to obtain numerical solutions of some singular differential problems defined in the complex domain. The potential use of these matrices is illustrated with the case of elliptic functions.

1. Introduction

In a series of papers ([4]-[6] and references therein), a Galerkin-type method has been used to solve boundary value problems and to find limit-cycles of nonautonomous dynamical systems. The method is based on the discretization of the differential problem by using differentiation matrices that are projections of the derivative in spaces of algebraic polynomials or trigonometric polynomials. This kind of matrices arise naturally in the context of interpolation of functions and yield exact values for the derivative of polynomial functions at certain points selected as nodes. The class of functions that defines the domain of the differential operator determines the kind of differentiation matrix to be used. Thus, to get a discrete form of an operator acting on real functions that drop off rapidly to zero at large distances, we can use the skew-symmetric differentiation matrix

$$D_{jk} = \begin{cases} 0, & i = j, \\ \dfrac{(-1)^{j+k}}{x_j - x_k}, & i \neq j, \end{cases} \tag{1.1}$$

constructed with the N zeros x_j of the Hermite polynomial $H_N(x)$ as lattice points[1]. On the other hand, to get a discrete form of an operator acting on periodic functions [7], the differentiation matrix

$$D_{jk} = \begin{cases} 0, & j = k, \\ \dfrac{(-1)^{j+k}}{2\sin\frac{(x_j - x_k)}{2}}, & j \neq k, \end{cases} \tag{1.2}$$

2000 *Mathematics Subject Classification*: 65D25, 41A05, 42A15.

Keywords and phrases: numerical differentiation, complex interpolation, meromorphic functions, trigonometric polynomials.

[1]Equation (1.1) gives an asymptotic expression for D_x.

should be used. In this case the lattice points can be chosen as the N equidistant points

$$x_j = -\pi + 2\pi j/N, \quad j = 1, 2, \cdots, N. \tag{1.3}$$

Discrete forms of multidimensional differential operators can be obtained by using direct products of suitable one-dimensional differentiation matrices (see [5] for instance).

Other approaches to obtaining differentiation matrices can be found in [10].

The purpose of this paper is to use the complex Hermite interpolation formula to find new differentiation matrices on the complex domain that produce an exact formula for the derivatives of complex rational functions and that can be used to approximate the solution of a singular differential equation in the complex domain. As examples, the derivatives of meromorphic two-periodic functions such as Jacobi elliptic functions and Weierstrass elliptic functions will be used to show the potential of this technique.

2. Interpolatory differentiation matrices

In order to illustrate the procedure for generating differentiation matrices, an already known matrix [3], which gives the derivative of algebraic polynomials of the complex variable z is obtained by using the Hermite interpolation formula.

(2.1) Algebraic polynomials. Let $f(z)$ be an analytic function on a domain G containing a closed rectifiable Jordan curve γ and let z_k be N different points of $I(\gamma)$ defining the polynomial

$$\omega(z) = \prod_{k=1}^{N} (z - z_k).$$

Then the unique polynomial $p(z)$ interpolating $f(z)$ associated to the set of points z_k is given by ([8], [9])

$$p(z) = \frac{1}{2\pi i} \int_\gamma \frac{f(\zeta)}{\omega(\zeta)} \frac{\omega(\zeta) - \omega(z)}{\zeta - z} d\zeta. \tag{2.1.1}$$

The residual function $R(z) = f(z) - p(z)$, $z \in I(\gamma)$ vanishes at z_k, yielding that $f(z_k) = p(z_k)$. To see this, write $R(z)$ as

$$R(z) = \frac{\omega(z)}{2\pi i} \int_\gamma \frac{f(\zeta)}{(\zeta - z)\omega(\zeta)} d\zeta. \tag{2.1.2}$$

Since the integral of the right-hand side of this equation represents an analytic function on $I(\gamma)$ we have that $R(z_k) = 0$.

To obtain the form of $p(z)$, the integral in equation (2.1.1) can be calculated by the residue theorem. Since $\omega(\zeta) - \omega(z)$ is divisible by $\zeta - z$, the integrand has simple poles only at $z_1, \ldots, z_N$. Thus the residue theorem yields the well-known Lagrange interpolation formula

$$p(z) = \sum_{j=1}^{N} f(z_j) \frac{\prod_{k\neq j}^{N} (z - z_k)}{\prod_{k\neq j}^{N} (z_j - z_k)}, \qquad z \in G. \tag{2.1.3}$$

The differentiation matrix D associated to this formula can be obtained if the derivative of (2.1.3) is evaluated at z_i and written as

$$\frac{dp(z_i)}{dz} = \sum_{j=1}^{N} D_{ij}\, f(z_j). \tag{2.1.4}$$

Since

$$\frac{d}{dz}\left[\prod_{k\neq j}^{N}(z-z_k)\Big/\prod_{k\neq j}^{N}(z_j-z_k)\right]_{z=z_i} = \begin{cases} \dfrac{\prod_{k\neq i,j}^{N}(z_i-z_k)}{\prod_{k\neq j}^{N}(z_j-z_k)}, & i\neq j,\\[2ex] \displaystyle\sum_{k\neq i}^{N}\frac{1}{(z_i-z_k)}, & i=j,\end{cases}$$

we get that the differentiation matrix for algebraic polynomials is given by

$$D_{ij} = \begin{cases} \displaystyle\sum_{k\neq i}^{N}\frac{1}{(z_i-z_k)}, & i=j,\\[2ex] \dfrac{\omega'(z_i)}{(z_i-z_j)\;\omega'(z_j)}, & i\neq j. \end{cases} \tag{2.1.5}$$

If $f(z)$ is a polynomial of degree at most $N-1$, $f(z)$ is given identically by $p(z)$ and therefore formula (2.1.4) gives the exact derivative of $f(z)$. Since $f'(z)$ is another polynomial in this class, the derivatives of higher order can be obtained by applying successively the matrix D to the vector of values $f(z_j)$, i.e.,

$$f^{(n)} = D^n f, \qquad n = 0, 1, 2, \ldots. \tag{2.1.6}$$

Here, $f^{(n)}$ is the vector whose entries are $d^n f(z_i)/dz^n$, D^n is the nth power of D and f is just the vector whose entries are $f(z_i)$. The functional form of $f^{(n)}(z)$ can be obtained through an interpolation of the values yielded by (2.1.6). Since any set of N different complex numbers belonging to G yields the same polynomial $f(z)$, this result is independent of the points z_k. On the other hand, if $f(z)$ is not a polynomial of degree at most $N-1$ a residual vector should be added to the right-hand side of (2.1.4) to get $f'(z_i)$. However, such a formula yields an good approximation to $f'(z_i)$ if the absolute value of the M-th term of the Taylor series of $f(z)$ goes to zero rapidly as M is increased.

(2.2) Trigonometric polynomials. The preceding arguments can be modified to consider the interpolation of periodic functions in terms of trigonometric polynomials. Let $f(z)$ be a one-periodic analytic function with period 2π and let G be a domain of the open strip $-\pi < \Re z < \pi$, $-\infty < \Im z < \infty$, containing a closed rectifiable Jordan curve γ.

Since any trigonometric polynomial $\tau(z) = a_0 + \sum_{k=1}^{m}(a_k \cos kz + b_k \sin kz)$ of degree at most m can be written in the form $\tilde{\tau}(s) = s^{-m}q(s)$ under the mapping $s = \varphi(z) = e^{iz}$ where $q(s)$ is a polynomial of degree at most $2m$ in s, we need to take an odd number $N = 2m+1$ of different points $s_k \in \varphi(I(\gamma))$, i.e, $2m+1$ different points $z_k \in I(\gamma)$, to yield an exact interpolation formula in the case in which $f(z)$ is a trigonometric polynomial of degree at most m. This fact can be shown as follows.

Let us take $N = 2m+1$. The set of points s_k, $k = 1, 2, \ldots, N$ define the polynomial $\tilde{\omega}(s) = \prod_{k=1}^{N}(s - s_k)$. The interpolant function $\tilde{p}(s)$ to $\tilde{f}(s) = f(\varphi^{-1}(s))$ corresponding to the set of N points s_k is given by

$$\tilde{p}(s) = \frac{s^{-m}}{2\pi i} \int_{\tilde{\gamma}} \frac{\tilde{f}(\zeta)}{\tilde{\omega}(\zeta)} \frac{s^m \tilde{\omega}(\zeta) - \zeta^m \tilde{\omega}(s)}{\zeta - s} d\zeta, \tag{2.2.1}$$

where $\tilde{\gamma} = \varphi(\gamma)$. Since $[s^m\tilde{\omega}(\zeta) - \zeta^m\tilde{\omega}(s)]/(\zeta - s)$ is a polynomial in s of degree $N - 1 = 2m$, $\tilde{p}(s)$ has the required form $s^{-m}q(s)$, where $q(s)$ is a polynomial of degree at most $2m$, to represent a trigonometric polynomial $\tau(z)$.

To show that $\tilde{f}(s_k) = \tilde{p}(s_k)$, let us consider the residual function $\tilde{R}(s) = \tilde{f}(s) - \tilde{p}(s)$ which is now

$$\tilde{R}(s) = \frac{1}{2\pi i} \frac{\tilde{\omega}(s)}{s^m} \int_{\tilde{\gamma}} \frac{\tilde{f}(\zeta)\zeta^m}{(\zeta - s)\tilde{\omega}(\zeta)} d\zeta.$$

By definition, $\tilde{G} = \varphi(G)$ does not contains points $s_k(\text{mod } 2\pi)$ other than s_k; therefore, the integral of the right-hand side of this equation represents an analytic function in $I(\tilde{\gamma})$ and we have that $\tilde{R}(s_k) = 0$.

Since $s^m\tilde{\omega}(\zeta) - \zeta^m\tilde{\omega}(s)$ is divisible by $\zeta - s$, the poles of the integrand are simple and located at s_k. The residue theorem yields now

$$\tilde{p}(s) = \sum_{j=1}^{N} \tilde{f}(s_j) \left(\frac{s_j}{s}\right)^m \frac{\prod_{k\neq j}^{N}(s - s_k)}{\prod_{k\neq j}^{N}(s_j - s_k)}, \qquad s \in \tilde{G} \tag{2.2.2}$$

and the trigonometric polynomial of degree $m = (N - 1)/2$ that interpolates $f(z)$ is

$$p(z) = \sum_{j=1}^{N} f(z_j) e^{i(N-1)(z_j - z)/2} \frac{\prod_{k\neq j}^{N}(e^{iz} - e^{iz_k})}{\prod_{k\neq j}^{N}(e^{iz_j} - e^{iz_k})}, \qquad z \in G. \tag{2.2.3}$$

Since

$$e^{i(N-1)(z_j - z)/2} \frac{\prod_{k\neq j}^{N}(e^{iz} - e^{iz_k})}{\prod_{k\neq j}^{N}(e^{iz_j} - e^{iz_k})} = \frac{\prod_{k\neq j}^{N} \sin\left(\frac{z - z_k}{2}\right)}{\prod_{k\neq j}^{N} \sin\left(\frac{z_j - z_k}{2}\right)},$$

we obtain from (2.2.3) the Gauss interpolation formula

$$p(z) = \sum_{j=1}^{N} f(z_j) \frac{\prod_{k\neq j}^{N} \sin\left(\frac{z - z_k}{2}\right)}{\prod_{k\neq j}^{N} \sin\left(\frac{z_j - z_k}{2}\right)}, \quad N = 2m + 1, \quad z \in G. \tag{2.2.4}$$

By writing the derivative of this formula in the form given by (2.1.4) we can obtain the differentiation matrix for trigonometric polynomials. Thus the matrix D whose entries are given by

$$(2.2.5) \qquad D_{ij} = \begin{cases} \dfrac{1}{2}\displaystyle\sum_{k\neq j}^{N} \cot\left(\frac{z_j - z_k}{2}\right), & i = j, \\ \\ \dfrac{1}{2}\csc\left(\dfrac{z_i - z_j}{2}\right) \dfrac{\prod_{k\neq i}^{N} \sin(\frac{z_i - z_k}{2})}{\prod_{k\neq j}^{N} \sin(\frac{z_j - z_k}{2})}, & i \neq j, \end{cases}$$

in terms of $N = 2m + 1$ different points $z_k \in G$, is a projection of d/dz in the subspace of trigonometric polynomials of degree at most $(N - 1)/2$. Therefore, if $f(z)$ belongs to this space, $p(z) \equiv f(z)$, and $f^{(n)}(z)$ satisfies again an equation like (2.1.6) but in this case D^n is the nth power of (2.2.5). The form of $f^{(n)}(z)$ can be obtained from the set of values $f^{(n)}(z_i)$ through an interpolation. For a general one-periodic analytic function with period 2π a residual vector should be added to (2.2.4).

It is worth to notice that in the case in which the set of points z_k are real numbers, the matrix (2.2.5) becomes the matrix used previously in [6], [7]. However, if the points lay on a straight line which is parallel to the imaginary axis, D takes the form

$$D_{ij} = \begin{cases} -\dfrac{i}{2}\displaystyle\sum_{k\neq i}^{N} \coth\left(\frac{y_i - y_k}{2}\right), & i = j, \\ \\ -\dfrac{i}{2}\operatorname{csch}\left(\dfrac{y_i - y_j}{2}\right) \dfrac{\prod_{k\neq j}^{N} \sinh(\frac{y_i - y_k}{2})}{\prod_{k\neq i}^{N} \sinh(\frac{y_j - y_k}{2})}, & i \neq j, \end{cases}$$

where $y_k = \Im z_k$, at the same time that the polynomial to differentiate becomes a linear combination of real hyperbolic functions.

(2.3) Rational functions. Equation (2.2.1) suggests the form of a interpolant rational function in the case in which $f(z)$ is a meromorphic function.

Let G be a domain that contains a closed rectifiable Jordan curve γ and let $f(z)$ be a meromorphic function with only one pole at $z = \alpha \notin G$ of order m. Let us choose N different points z_k of $I(\gamma)$ and construct the polynomial $\omega(z) = \prod_{k=1}^{N}(z - z_k)$. Thus the rational function interpolating $f(z)$ corresponding to the set of points z_k is given by

$$(2.3.1) \qquad p(z) = \frac{(z-\alpha)^{-m}}{2\pi i} \int_\gamma \frac{f(\zeta)}{\omega(\zeta)} \frac{(z-\alpha)^m \omega(\zeta) - (\zeta - \alpha)^m \omega(z)}{\zeta - z} d\zeta.$$

This can be shown by the same arguments used in the previous case. Again, $(z-\alpha)^m\omega(\zeta) - (\zeta-\alpha)^m\omega(z)$ is divisible by $\zeta - z$. Let $K_N^m(z, \zeta)$ be such a quotient, i.e.,

$$K_N^m(z,\zeta) = \frac{(z-\alpha)^m \omega(\zeta) - (\zeta - \alpha)^m \omega(z)}{\zeta - z}.$$

Since $K_N^m(z, \zeta)$ is a polynomial of degree $N - 1$ in z, $p(z)$ is a rational function of form $q(z)/(z - \alpha)^m$ that interpolates to $f(z)$ at z_k, where $q(z)$ is a polynomial of degree at most $N-1$. The residual function

$$R(z)=\frac{1}{2\pi i}\frac{\omega(z)}{(z-\alpha)^m}\int_\gamma \frac{f(\zeta)(\zeta-\alpha)^m}{(\zeta-z)\omega(\zeta)}d\zeta.$$

vanish at z_k because the integral is an analytic function in $I(\gamma)$ and $\alpha\notin G$. Therefore, $p(z_k)=f(z_k)$. The residue theorem yields now

$$p(z)=\sum_{j=1}^{N} f(z_j)\left(\frac{z_j-\alpha}{z-\alpha}\right)^m \frac{\prod_{k\neq j}^{N}(z-z_k)}{\prod_{k\neq j}^{N}(z_j-z_k)}, \qquad z\in G. \tag{2.3.2}$$

The derivative of this equation at z_i can be written in the form (2.1.4) where we have now

$$D_{ij}=\begin{cases}\displaystyle\sum_{k\neq i}^{N}\frac{1}{(z_i-z_k)}-\frac{m}{z_i-\alpha}, & i=j,\\[2ex] \displaystyle\frac{(z_j-\alpha)^m/(z_i-\alpha)^m}{z_i-z_j}\,\frac{\omega'(z_i)}{\omega'(z_j)}, & i\neq j.\end{cases} \tag{2.3.3}$$

Obviously, if $f(z)$ is a rational function of the form $q(z)/(z-\alpha)^m$ where $q(z)$ is a polynomial of degree at most $N-1$, $f(z)\equiv p(z)$ and formula (2.1.4) becomes

$$\frac{df(z_i)}{dz}=\sum_{j=1}^{N} D_{ij}f(z_j). \tag{2.3.4}$$

However, the powers of D do not give the derivatives of higher order as in the previous cases since $f'(z)$ does not has the required form: it has a pole of order $m+1$ at $z=\alpha$. Despite this, it is possible to obtain the value of $f^{(n)}(z)$ at z_k by using the matrix

$$D_n(m)=D_{m+n-1}D_{m+n-2}\cdots D_m. \tag{2.3.5}$$

Each matrix D_k is defined by (2.3.3) where the parameter m, defining the order of the pole, is substituted by each value of the index k. Thus, it should be clear that if $f(z)$ has the form given above, (2.1.6) becomes

$$f^{(n)}=D_n(m)f, \qquad n=0,1,2,\ldots. \tag{2.3.6}$$

This equation can be generalized to the case in which $f(z)$ is a rational function of form

$$f(z)=\frac{\sum_{k=0}^{M}a_k z^k}{(z-\alpha_1)^{m_1}(z-\alpha_2)^{m_2}\cdots(z-\alpha_r)^{m_r}} \tag{2.3.7}$$

where $\alpha_l\notin G$, $l=1,2,\cdots,r$. A straightforward calculation gives the generalized form of (2.3.6)

$$f^{(n)}=D_n(m_1,m_2,\ldots,m_r)f, \qquad n=0,1,2,\ldots, \tag{2.3.8}$$

where now $D_n(m_1,m_2,\ldots,m_r)$ stands for the ordered matrix product

$$D_n(m_1,m_2,\ldots,m_r)=\prod_{k=n-1}^{0} D_{m_1+k,m_2+k,\ldots,m_r+k} \tag{2.3.9}$$

and $D_{\mu_1,\mu_2,\dots,\mu_r}$ is the matrix whose entries are given by

$$(D_{\mu_1,\mu_2,\dots,\mu_r})_{ij} = \begin{cases} \displaystyle\sum_{k\neq i}^{N} \frac{1}{(z_i - z_k)} - \sum_{k=1}^{r} \frac{\mu_k}{z_i - \alpha_k}, & i = j, \\ \displaystyle\frac{1}{z_i - z_j}\, \frac{\omega'(z_i)}{\omega'(z_j)} \prod_{k=1}^{r} \left(\frac{z_j - \alpha_k}{z_i - \alpha_k}\right)^{\mu_k}, & i \neq j. \end{cases} \tag{2.3.10}$$

It should be noticed that (2.3.8) is an exact formula whenever $N \geq M + nr - 1$, where M is the degree of the polynomial in the numerator of (2.3.7). The reason is that after each differentiation the numerator of the derivatives of $f(z)$ is a polynomial whose degree grows by r. If the function $f(z)$ to differentiate has poles at $\alpha_1, \dots, \alpha_r$ of orders $n_1, \dots, n_r$ instead $m_1, \dots, m_r$, with $n_1 < m_1, \dots, n_r < m_r$, formula (2.3.8) is still exact whenever $N \geq M + nr - 1 + \sum_{k=1}^{r}(m_k - n_k)$. The reason for this is that in such a case $f(z)$ can be writen in the form (2.3.7) where the numerator is now a polynomial of degree $M + \sum_{k=1}^{r}(m_k - n_k)$. Obviously, in this case is much better to use the differentiation matrix $D_n(n_1, n_2, \dots, n_r)$ instead $D_n(m_1, m_2, \dots, m_r)$ for numerical purposes.

3. Applications

As stated before, the numerical solution of differential problems can be accomplished by the use of differentiation matrices, and in the case of a differential problem in the complex domain, the differentiation matrices introduced in this paper may be useful. To illustrate the potential of their use, we choose two meromorphic cases which are important in applications: Jacobi elliptic functions and Weierstrass elliptic functions. In both cases it is possible to establish the numerical convergence of the results since the derivatives of these functions are known. We also obtain approximate solutions of a differential equation with a regular singularity at $z = 0$. Before beginning these examples is convenient to alert the reader to the fact that the numerical implementation of the matrices for rational functions given above may need a high-precision code: in most cases, the usual 16-digit precision is not enough to obtain accurate results.

(3.1) A rational function. Let us consider the function

$$f(z) = \frac{z^7 + z + 1}{z^{10}}. \tag{3.1.1}$$

According to the results of the last section, to obtain the exact value of the n-th derivative of (3.1.1), [cf. Eq. (2.3.6)], it is necessary to choose $N > 7$ different points $z_k \neq 0$ in the complex plane to build the matrix (2.3.5) where obviously, $\alpha = 0$ and $m = 10$.

As an example, let us take the third derivative of (3.1.1). Table 3.1 shows the max-norm of the vector whose entries are $|[f^{(3)}(z_j) - \sum_{k=1}^{N}(D_3(10))_{jk} f(z_k)]/f^{(3)}(z_j)|$ (the relative error) in terms of N. The nodes were chosen to be as evenly spaced on the ray $z = (1+i)t$, $1/2 < t \leq 1$, i.e., $z_k = (1+i)(1+k/N)/2$, $k = 1, \dots, N$, and the computations were made

Table 1. The norm of the relative error $E_N = |(f''' - p''')/f'''|_\infty$ for (3.1.1) in terms of the number of nodes N.

N	E_N
4	0.657
5	0.136
6	0.155×10^{-1}
7	0.742×10^{-3}
8	$0. \times 10^{-16}$
9	$0. \times 10^{-16}$
10	$0. \times 10^{-16}$
11	$0. \times 10^{-16}$

with the standard 16-digit precision. As it can be seen from Table 3.1, the error vanishes for values of N greater than 7, as expected.

(3.2) Elliptic functions. As is known, an elliptic function is a doubly periodic function which is analytic except at poles [2]-[1] and one of the two simplest cases of elliptic functions corresponds to Jacobi's functions; the other corresponds to Weierstrass' $\wp$-function.

The differentiation matrices given above do not apply to functions like these; however, it is possible to build a matrix which is expected to provide approximate values of $f^{(n)}(z_k)$ along straight lines inside the fundamental paralelograms and near the poles. Since an elliptic function $f(z)$ becomes a one-periodic function if z is constrained to move along a straight line defined by one of the periods, such a matrix can be constructed by using trigonometric polynomials divided by an algebraic polynomial with zeros of suitable orders taken at the poles of $f(z)$. This procedure is equivalent to taking apart the matrix (2.3.10) and incorporating only the singular terms in (2.2.5) to yield the new matrices

$$\tilde{D}_n(m_1, m_2, \ldots, m_r) = \prod_{k=n-1}^{0} \tilde{D}_{m_1+k, m_2+k, \ldots, m_r+k}, \tag{3.2.1}$$

where

(3.2.2)

$$(\tilde{D}_{\mu_1,\mu_2,\ldots,\mu_r})_{ij} = \begin{cases} \displaystyle\sum_{k\neq i}^{N} \cot\left(\frac{z_j - z_k}{2}\right) - \sum_{k=1}^{r} \frac{\mu_k}{z_i - \alpha_k}, & i=j, \\ \displaystyle\frac{1}{2}\csc\left(\frac{z_i - z_j}{2}\right) \prod_{k\neq i}^{N} \frac{\sin(z_i - z_k)/2}{\sin(z_j - z_k)/2} \prod_{k=1}^{r} \left(\frac{z_j - \alpha_k}{z_i - \alpha_k}\right)^{\mu_k}, & i\neq j. \end{cases}$$

To test the numerical performance of this matrix we choose two numerical examples. The first one corresponds to Jacobi's function $f(z) = \mathrm{sn}(z|\,\frac{1}{2}\,)$ which has two periods $4K$ and $2iK'$ and two simple poles at iK' and $2K + iK'$, where

$$K = \int_0^{\pi/2} \frac{d\theta}{\sqrt{1-(\sin^2\theta)/2}}, \qquad K' = K \approx 1.854.$$

Therefore, to build the differentiation matrix we take 2π-periodic trigonometric polynomials, $r = 2$, $\mu_1 = \mu_2 = 1$, $\alpha_1 = iK'$, and $\alpha_2 = 2K + iK'$ in (3.2.2), and to measure the approximation of $f'(z_j)$ by $p'(z_j) = \sum_{k=1}^{N} (\tilde{D}_1(1,1))_{jk} f(z_k)$ we use the max norm. Here, $f'(z) = \mathrm{cn}(z|\,\frac{1}{2}\,)\,\mathrm{dn}(z|\,\frac{1}{2}\,)$. The results are displayed in Figure 1, where the max-norm of the error $f'(z_j) - p'(z_j)$ is plotted against the number of nodes showing numerical convergence. The number of digits of precision used in the calculations is 16 and the nodes are chosen to be evenly spaced on the ray $z = (2+i)t$, $1/2 < t \le 1$, i.e., $z_k = (2+i)(1+k/N)/2$, $k = 1, \ldots, N$. The norm of the error is 1.6×10^{-4} for $N = 10$ and 1.0×10^{-8} for $N = 20$.

Our second example is the Weierstrass $\wp$-function, which has a double pole at the origin and two periods ω_1 and ω_2. The derivative of $\wp(z)$ satisfies

$$(\wp'(z))^2 = 4(\wp(z))^3 - g_2\wp(z) - g_3\,,$$

where g_2 and g_3 are the elliptic invariants which are related to both ω_1 and ω_2 [1].

To approximate the first derivative of $\wp(z)$ at the nodes, we need to take $n = 1$ and $r = 1$ in (3.2.1) and $\mu_1 = 2$ and $\alpha_1 = 0$ in (3.2.2). The precision of the calculations, the nodes, and the kind of trigonometric polynomials used to construct the differentiation matrix are the same as above. The numerical results are displayed in Figure 1 and show again numerical convergence with a small number of nodes. For $N = 10$ the norm of the error is 1.5×10^{-5} and 1.0×10^{-8} for $N = 20$.

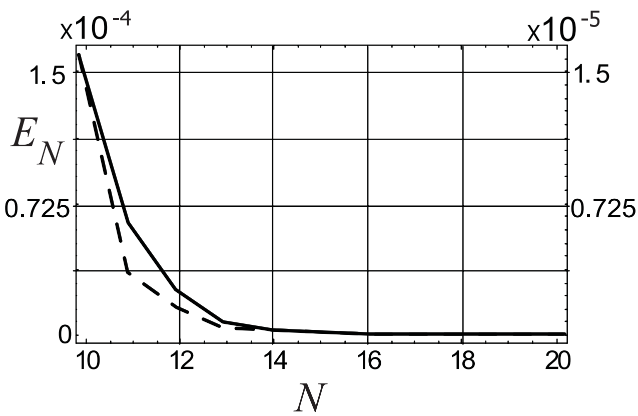


Figure 1: The norm of the error $E_N = |f' - p'|_\infty$ against the number of nodes for the Jacobian elliptic function (solid line) and Weierstrass' $\wp$-function (broken line). The vertical axes are scaled according to Jacobi's case (the left-hand axis) or Weierstrass' case (the right-hand axis).

(3.3) Krummer's equation. As a final example, we consider the Krummer differential equation, written as an eigenvalue problem

$$z\frac{d^2f(z)}{dz^2}+(b-z)\frac{df(z)}{dz}=af(z), \tag{3.3.1}$$

which has a regular singularity at $z=0$ and an irregular singularity at ∞. As is well known, the single-valued solution of this equation is the confluent hypergeometric function $M(a,b,z)={}_1F_1(a,b,z)$. Since this function can be approximated by algebraic polynomials for $b\neq -n$ (n a positive integer), we can obtain approximate solutions of this differential equation by using the differentiation matrix D given by (2.2.5) to approximate the derivative of $M(a,b,z)$ and solving the N-dimensional eigenvalue problem

$$Lf_\lambda=\lambda f_\lambda,\quad L=ZD^2+(b1_N-Z)D \tag{3.3.2}$$

where Z is a diagonal matrix whose nonzero elements are the nodes $z_1,\ldots,z_N$, b is a complex number ($b\neq -n$), 1_N is the identity matrix of dimension N and the eigenvalue λ is the value of a at which $M(a,b,z_k)$ is to be approximated by $(f_\lambda)_k$. Let us denote by M_a the vector whose kth component is $M(a,b,z_k)$. Since the nth coefficient of the power series of $M(a,b,z)$ is given by

$$\frac{a(a+1)(a+2)\ldots(a+n-1)}{b(b+1)(b+2)\ldots(b+n-1)n!},$$

the best approximation obtained for a given set of parameters b, N, z_k, is given by the eigenvector f_λ ($\lambda=a$) corresponding to the eigenvalue with lowest absolute value λ_m. To construct the matrix L in (3.3.2) we choose $N=21$ and $z_k=5(1+i)k/N$, $k=1,\ldots N$. For b we take the cases $b=5/2$ and $b=3+2i$. In order to compare the approximate and exact results, we normalize both vectors $M_{\lambda m}$ and $f_{\lambda m}$ with the max-norm. The calculations were made with 16 digits of precision and the results are displayed in Fig. 2. The absolute errror $\|M_{\lambda m}-f_{\lambda m}\|_0$ is 0.0675659 for $b=5/2$ and 0.0426948 for $b=3+2i$.

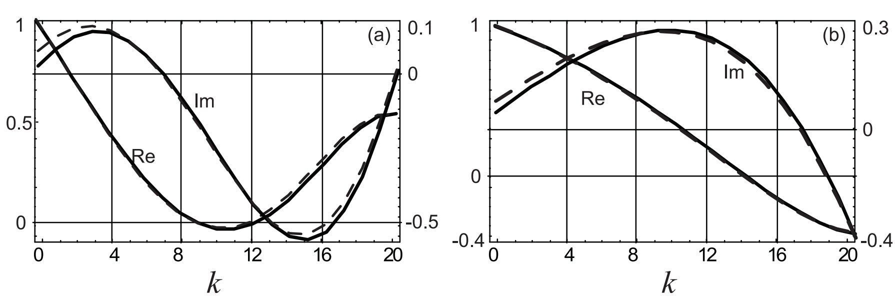


Figure 2: Normalized real (Re) and imaginary (Im) parts of $f_{\lambda m}$ plotted versus their index (solid lines). Case (a) corresponds to $b=5/2$ and $\lambda_m=-0.301513+1.00758i$ and case (b) to $b=3+2i$ and $\lambda_m=-0.381925+0.527533i$. They are compared with the exact values of the Krummer function (broken lines). The matrix L is constructed with 21 nodes $z_k=5(1+i)k/N$, $k=1,\ldots 21$. The scale on the left-hand vertical axis corresponds to the real part and the one on the right-hand to the imaginary part.

4. Concluding remark

According to the results of section 2, the process of interpolation in vector spaces of polynomials of dimension N maps the derivative d/dx into a $N \times N$ matrix D. The fact that a differential operator acting on a vector space of finite dimension can be written as a matrix is not a surprise of course; however, it should be noted that the matrix D yields the derivative of a function by taking the values of the function at N arbitrary (but different) points including the point where the derivative is to be evaluated, i.e., it acts on a function as a nonlocal operator in spite of the local character of a differential operator as the derivative.

Acknowledgment

The authors are grateful to the referees for their valuable comments to improve this paper.

Received July 02, 2004

Final version received November 21, 2006

FACULTAD DE CIENCIAS FÍSICO-MATEMÁTICAS
UNIVERSIDAD MICHOACANA
58060 MORELIA, MICH.
MÉXICO
rcampos@zeus.umich.mx
claudio@fismat.umich.mx

REFERENCES

[1] M. ABRAMOWITZ AND I. A. STEGUN (eds.), Handbook of mathematical fuctions, Dover Publications, New York, 1972.

[2] F. BOWMAN, Introduction to elliptic functions with applications, Dover Publications, New York, 1961.

[3] F. CALOGERO, *Matrices, differential operators and polynomials*, J. Math. Phys. **22** (1981), 919-932.

[4] R. G. CAMPOS, *Solving nonlinear two point boundary value problems*, Bol. Soc. Mat. Mexicana. **3** (1997), 279–297.

[5] R. G. CAMPOS AND L.O. PIMENTEL, *Hydrogen atom in a finite linear space*, J. Comp. Phys. **160** (2000), 179–194.

[6] R. G. CAMPOS Y GILBERTO O. ARCINIEGA, *A limit-cycle solver for nonautonomous dynamical systems*, Rev. Mex. Fis. **52** (2006) 267–271.

[7] R. G. CAMPOS AND L.O. PIMENTEL, *A finite-dimensional representation of the quantum angular momentum operator*, Il Nuovo Cimento B**116** (2001), 31–46.

[8] A. I. MARKUSEVICH, Theory of Functions of a Complex Variable, Chelsea Pub. Co., New York, 1985.

[9] J. L. WALSH, Interpolation and Approximation by Rational Functions in the Complex Domain, American Mathematical Society, Colloquium Publications, Vol. XX, Providence, Rhode Island, 1969.

[10] B. D. WELFERT, *Generation of pseudospectral differentiation matrices I*, SIAM J. Numer. Anal. **34** (1997), 1640–1657

Bol. Soc. Mat. Mexicana (3) Vol. 12, 2006

UNIFORM CONVERGENCE OF VALUE ITERATION POLICIES FOR DISCOUNTED MARKOV DECISION PROCESSES

DANIEL CRUZ-SUÁREZ AND RAÚL MONTES-DE-OCA

ABSTRACT. This paper deals with infinite horizon Markov Decision Processes (MDPs) on Borel spaces. The objective function considered, induced by a nonnegative and (possibly) unbounded cost, is the expected total discounted cost. For each of the MDPs analized, the existence of a unique optimal policy is assumed. Conditions that guarantee both pointwise and uniform convergence on compact sets of the minimizers of the value iteration algorithm to the optimal policy are provided. The theory developed in this paper is illustrated with three examples: an inventory/production system, the linear regulator problem, and a nonlinear additive-noise system with unbounded cost function.

1. Introduction

This paper is related to infinite horizon Markov decision processes (MDPs) on Borel spaces (see [5], [12],[13] and [16]).

A nonnegative and possibly unbounded cost function is considered. The expected total discounted cost is supposed to be the objective function. For such Markov decision process (MDP), the existence of a unique stationary optimal policy f^* is assumed (see [8] for conditions to guarantee the uniqueness of optimal policies in discounted MDPs).

Consider an MDP with such description. Denote the state and the action spaces by X and A, respectively. Let $A(x)$, $x \in X$ be the admissible action sets, and let Q represent the transition probability law. Take V^* as the optimal value function, and for each $n = 1, 2, \cdots$, let V_n and f_n denote the minimum and the minimizer corresponding to the step n of the value iteration algorithm (see [12]), respectively.

This paper deals mainly with establishing conditions that ensure the uniform convergence on compact sets of $\{f_n\}$ to f^* (see Theorem (3.41) below).

Additionally, for the pointwise convergence of $\{f_n\}$ to f^*, conditions which are fewer and weaker than those presented in the paper for the uniform convergence on compact sets of $\{f_n\}$ to f^* are provided (see Section 5 below).

The uniform convergence on compact sets of $\{f_n\}$ to f^* guarantees, given a compact set $\varsigma \subset X$ and $\varepsilon > 0$, the existence of a positive integer $N(\varsigma, \varepsilon)$ such that $f_{N(\varsigma,\varepsilon)}(x)$ is in the ε-neighborhood of $f^*(x)$ for all $x \in \varsigma$. Therefore this integer $N(\varsigma, \varepsilon)$ can be interpreted as an extended version of the standard concept known in the literature of MDPs as the Forecast Horizon (FH) (see [6], [7], [10], [13], [19], [20], [21], and [22]). The FH is a positive integer N^* such that $f_n = f^*$, for all $n \geq N^*$. Note that the FH allows us to obtain a

2000 *Mathematics Subject Classification*: 90C40, 93E20.

Keywords and phrases: discounted Markov decision process, optimal policy, value iteration algorithm and its minimizers, convergence to the optimal policy, forecast horizon.

strong class of convergence of $\{f_n\}$ to f^*, i.e. the limit f^* is attained in N^* steps. This excludes important control problems in which $f_n \neq f^*$ for all n, for instance, the linear-quadratic model (see [5] and [12]). In this paper the linear-quadratic model satisfies practically all the conditions proposed.

The results presented here are based mainly on the continuity of the cost function, of V^* and $V_n, n = 1, 2, \cdots$, as well as on the continuity on $\mathbb{K} = \{(x, a) \mid x \in X, a \in A(x)\}$ of

$$\int V_n(y) Q\left(dy \mid \cdot, \cdot\right),$$

$n = 1, 2, \cdots$, and also of

$$\int V^*(y) Q\left(dy \mid \cdot, \cdot\right)$$

(these integrals are associated with the right-hand side of the Optimality Equation, see [5], [11], [12],[13] and [16]).

Due to the fact that the cost function is assumed nonnegative, $V_n \uparrow V^*$ (see [12]) and $\int V_n(y) Q\left(dy \mid \cdot, \cdot\right) \uparrow \int V^*(y) Q\left(dy \mid \cdot, \cdot\right)$. In this situation, the assumptions of Dini's Theorem (see [14]), are satisfied. This important fact will be used in the proof of the results (see Lemma (3.28) below).

Several examples are presented to illustrate the theory developed in this paper. These examples are: an inventory/production system (see [11]), the linear-quadratic model (see [5]), and a nonlinear additive-noise system with unbounded cost function.

Now some comments about previous results in the literature will be provided.

The study of the pointwise convergence has been dealt with in Stokey and Lucas [23] for MDPs on Euclidean spaces, compact and convex actions sets and assuming strict concavity with respect to the actions in the right-hand side of the Optimality Equation (Stokey and Lucas [23] deal with rewards instead of costs). Furthermore, if X is compact, the uniform convergence of $\{f_n\}$ to f^* has been obtained in [23].

It is important to mention the result obtained by Schäl [17] (see also [15]) which permits, in the case of compact admissible actions sets, the conclusion that for each $x \in X$, $f^*(x)$ is an accumulation point of the sequence $\{f_n(x)\}$.

Finally, for general optimization problems with applications to deterministic MDPs, the convergence of optimal solutions of finite horizon problems to the optimal solutions of the infinite horizon problems is analyzed in [19].

The paper is organized as follows. Section 2 provides the preliminaries, i.e. the basic theory that will be used in the paper. In Section 3, the main assumptions and the theorem of the uniform convergence are provided. In Section 4 examples are presented. In Section 5, some remarks about pointwise convergence are given.

2. Preliminaries

Let $(X, A, \{A(x)\colon x \in X\}, Q, c)$ be a discrete-time, stationary ***Markov decision model*** (see [12] for notation and terminology) which consists of the state space X, the control or action set A, the admissible action sets $A(x), x \in X$, the transition law Q, and the one-stage cost c.

The sets X and A are assumed to be Borel spaces, with Borel $\sigma-$algebras $\mathcal{B}(X)$ and $\mathcal{B}(A)$, respectively. Moreover, for every $x \in X$ there is a nonempty Borel set $A(x) \subset A$ whose elements are the feasible control actions when the state of the system is x. Define $\mathbb{K} := \{(x, a) \mid x \in X, a \in A(x)\}$. The transition law $Q(B \mid x, a), B \in \mathcal{B}(X), x \in X, a \in A(x)$ is a stochastic kernel on X, given $\mathbb{K}$ (that is, $Q(\cdot \mid x, a)$ is a probability measure on X for every $(x, a) \in \mathbb{K}$; and $Q(B \mid \cdot)$ is a measurable function on $\mathbb{K}$ for every $B \in \mathcal{B}(X)$). Finally, $c\colon \mathbb{K} \to \mathbb{R}$ is a measurable function which represents the cost-per-stage.

A *policy* π is a (measurable, possibly randomized) rule for choosing actions, and at each time $t = 0, 1, \cdots$, the control prescribed by π may depend on the current state as well as on the history of previous states and actions. The set of all policies is denoted by Π. Given the initial state $x_0 = x$, any policy π defines the unique probability distribution of the state-action processes $\{(x_t, a_t)\}$. For details see, for instance, [12]. This probability distribution is denoted by P_x^π, whereas E_x^π stands for the corresponding expectation operator. Let $\mathbb{F}$ be the set of all measurable functions $f\colon X \to A$, such that $f(x) \in A(x)$ for every $x \in X$. A policy $\pi \in \Pi$ is *stationary* if there exists $f \in \mathbb{F}$ such that, under π, the control $f(x_t)$ is applied at each time $t = 0, 1, \cdots$. The set of all stationary policies is identified with $\mathbb{F}$.

The focus here is on the *expected total discounted cost* defined as

$$V(\pi, x) = E_x^\pi[\sum_{t=0}^{\infty} \alpha^t c(x_t, a_t)], \tag{2.1}$$

when the policy $\pi \in \Pi$ is used, and $x \in X$ is the initial state. In (2.1), $\alpha \in (0, 1)$ is a given discount factor.

A policy π^* is said to be *optimal* if

$$V(\pi^*, x) = V^*(x),$$

$x \in X$, where

$$V^*(x) := \inf_{\pi \in \Pi} V(\pi, x), \tag{2.2}$$

$x \in X$, is the so-called *optimal value* function.

Now some assumptions and results to be used in the next sections will be listed.

Assumption (2.3). a. The one-stage cost c is nonnegative, lower semicontinuous (l.s.c.) and inf-compact on $\mathbb{K}$. (Recall that c is inf-compact on $\mathbb{K}$ if the set

$\{a \in A(x) \mid c(x, a) \leq \bar{s}\}$ is compact for every $x \in X$ and $\bar{s} \in \mathbb{R}$.)

b. The transition law Q is strongly continuous, i.e.,

$$\mu'(x, a) := \int \mu(y) Q\left(dy \mid x, a\right)$$

is continuous and bounded on $\mathbb{K}$ for every measurable bounded function $\mu\colon X \to \mathbb{R}$.

c. There is a policy π such that $V(\pi, x) < \infty$ for each $x \in X$.

Definition (2.4). The *value iteration* functions are defined as

$$V_n(x) = \min_{a \in A(x)} \left[c(x,a) + \alpha \int V_{n-1}(y) Q\left(dy \mid x, a\right) \right], \tag{2.5}$$

for all $x \in X$ and $n = 1, 2, \cdots$, with $V_0(\cdot) = 0$.

Remark (2.6). Using Assumption (2.3) it is possible to demonstrate (see [12]) that for each $n = 1, 2, \cdots$, V_n is a measurable function and there exists a stationary policy $f_n \in \mathbb{F}$ such that the minimum in (2.5) is attained, i.e.,

$$V_n(x) = c\left(x, f_n(x)\right) + \alpha \int V_{n-1}(y) Q\left(dy \mid x, f_n(x)\right), \tag{2.7}$$

$x \in X$.

LEMMA (2.8). *([12], Theorem 4.2.3) Suppose that Assumption (2.3) holds. Then*

a. *The optimal value function V^* defined in (2.2) is the (pointwise) minimal solution of the* **Optimality Equation** *(OE), i.e. for all $x \in X$:*

$$V^*(x) = \min_{a \in A(x)} \left[c(x,a) + \alpha \int V^*(y) Q\left(dy \mid x, a\right) \right], \tag{2.9}$$

and if u is another solution to the OE, then $u(\cdot) \geq V^*(\cdot)$.

There is also $f^* \in \mathbb{F}$ such that

$$V^*(x) = c\left(x, f^*(x)\right) + \alpha \int V^*(y) Q\left(dy \mid x, f^*(x)\right), \tag{2.10}$$

$x \in X$, and f^* is optimal.

b. *For every $x \in X$, $V_n(x) \uparrow V^*(x)$ as $n \to +\infty$.*

Assumption (2.11). Suppose that f^* given in (2.10) is unique.

Remark (2.12). See [8] for conditions to ensure the uniqueness of optimal policies of discounted MDPs.

Throughout the paper, MDPs that satisfy Assumptions (2.3) and (2.11) are considered. Assumptions (2.3) and (2.11) will not be mentioned in each Lemma or Theorem in this paper, but they are supposed to hold.

LEMMA (2.13). *Let $\{g_n\}$ be a sequence of continuous real-valued functions on a metric space $\left(Y, d_Y\right)$. Let g be a continuous real-valued function on Y. Then $g_n(z_n) \to g(z), n \to +\infty$ for every $z \in Y$ and every sequence $\{z_n\}$ in Y which converges to z if and only if $\{g_n\}$ converges uniformly on compact sets to g.*

Proof. Suppose that $g_n(z_n) \to g(z), n \to +\infty$ for every $z \in Y$ and every sequence $\{z_n\}$ in Y which converges to $z \in Y$, but there is a compact set $\varsigma \subset Y$ such that $\{g_n\}$ does not converge uniformly to g on ς. Consequently, there exist $\varepsilon > 0$, a sequence of positive integers $\{n_k\}$, $n_1 < n_2 < \cdots$, and a sequence $\{y_k\} \subset \varsigma$, such that

$$\left|g_{n_k}(y_k) - g(y_k)\right| \geq \varepsilon, \tag{2.14}$$

$k = 1, 2, \cdots$. It is possible to assume without losing generality that there exists $y \in \varsigma$ such that $y_k \to y$, $k \to +\infty$ (recall that ς is a compact set). Now take $z_{n_k} := y_k$, $k = 1, 2, \cdots$, and observe that $z_{n_k} \to y$, $k \to +\infty$. By the hypothesis, it follows that $g_{n_k}\left(z_{n_k}\right) \to g(y), k \to +\infty$. Therefore, by the continuity of g, it results that

$$\left|g_{n_k}\left(z_{n_k}\right) - g\left(z_{n_k}\right)\right| \leq \left|g_{n_k}\left(z_{n_k}\right) - g(y)\right| + \left|g(y) - g\left(z_{n_k}\right)\right| \to 0, \tag{2.15}$$

$k \to \infty$.

On the other hand, putting $y_k = z_{n_k}$, $k = 1, 2 \cdots$, in (2.14), it follows that

$$\left|g_{n_k}\left(z_{n_k}\right) - g\left(z_{n_k}\right)\right| \geq \varepsilon > 0, \tag{2.16}$$

for all $k = 1, 2, \cdots$, which is a contradiction to (2.15).

Conversely, assume that $\{g_n\}$ converges uniformly on compact sets to g. Take $z \in Y$ and let $\{z_n\}$ be a sequence in Y which converges to z. Put $\varsigma := \{z_n\} \cup \{z\}$. Hence ς is a compact set and, by hypothesis, $\{g_n\}$ converges uniformly on ς to g. Thus, this fact, the continuity of g, and the inequality

$$|g_n(z_n) - g(z)| \leq |g_n(z_n) - g(z_n)| + |g(z_n) - g(z)|, \tag{2.17}$$

$n = 1, 2, \cdots$, imply that $g_n(z_n) \to g(z), n \to +\infty$. This completes the proof of Lemma (2.13). □

Remark (2.18). Lemma (2.13) is a particular case of result 7.5 of Chapter XII in [9]. In fact, Lemma (2.13) holds for g_n and g which are defined in a topological space Y (that is 1^{st} countable) and take values in a metric space Z (see [9]). The proof of this Lemma is presented here for completeness of the paper.

Definition (2.19). Let X and Y be (nonempty) Borel spaces. A multifunction Ψ from X to Y is said to be

a. *upper semicontinuous* (u.s.c.) if $\{x \in X \mid \Psi(x) \cap F \neq \emptyset\}$ is closed in X for every closed $F \subset Y$;

b. *lower semicontinuous* (l.s.c.) if $\{x \in X \mid \Psi(x) \cap G \neq \emptyset\}$ is open in X for every open set $G \subset Y$;

c. *continuous* if it is both u.s.c. and l.s.c.

A convenient characterization of a compact-valued u.s.c. multifunction will be given in the following result (for the proof of this result see Theorem 16.20 and Theorem 16.21, p. 534 in [1]). It will be used in Sections 3 and 4 below.

LEMMA (2.20). *Let X and Y be (nonempty) Borel spaces. Let Ψ be a multifunction from X to Y. Suppose that $\Psi(x) \neq \emptyset$ for every $x \in X$ and, moreover, Ψ is compact-valued. Then the following statements are equivalent:*

a. *Ψ is u.s.c.*

b. *If $x_n \to x, n \to +\infty$ in X and $y_n \in \Psi(x_n), n = 1, 2, \cdots$, there exist a subsequence $\{y_{n_k}\}$ of $\{y_n\}$ and $y \in \Psi(x)$ such that $y_{n_k} \to y, k \to +\infty$.*

3. Uniform Convergence on Compact Sets

In this section the assumptions and the theorem which ensure the uniform on compact sets convergence of minimizers of the value iteration algorithm of discounted MDPs to the optimal policies are presented.

The following notation will be used in the rest of the paper.

Notation (3.1). a. For $x \in X$, denote

$$\widehat{A}(x) := A_{V^*(x)}(x) = \{a \in A(x) \mid c(x,a) \leq V^*(x)\},$$

where V^* is the function given in (2.2).

b. For each $n = 1, 2, \cdots$,

$$G_n(x,a) := c(x,a) + \alpha \int V_{n-1}(y) Q\left(dy \mid x, a\right), \tag{3.2}$$

$(x,a) \in \mathbb{K}$, where V_{n-1} is the function given in (2.5),

c. $G(x,a) := c(x,a) + \alpha \int V^*(y) Q\left(dy \mid x,a\right), (x,a) \in \mathbb{K}$.

d. Denote

$$\widehat{\mathbb{K}}_{\varsigma} := \{(x,a) \in \mathbb{K} \mid x \in \varsigma, a \in \widehat{A}(x)\}, \tag{3.3}$$

where $\varsigma \subset X$ is a nonempty compact set.

Assumption (3.4). a. The multifunction $x \to A(x)$ is u.s.c. and closed-valued.

b. $c(\cdot,\cdot)$ is a continuous function on $\mathbb{K}$.

c. $V_n(\cdot)$, $n = 1, 2, \cdots$, and $V^*(\cdot)$ are continuous functions on X.

d. The integrals

$$\int V_n(y) Q\left(dy \mid \cdot, \cdot\right), n = 1, 2, \cdots, \tag{3.5}$$

and

$$\int V^*(y) Q\left(dy \mid \cdot, \cdot\right) \tag{3.6}$$

are finite and continuous functions on $\mathbb{K}$.

Assumption (3.7). a. A, the control space, is a compact set.

b. The multifunction $x \mapsto A(x)$ is compact-valued.

c. The one-stage cost c is strictly unbounded, that is, there exist nondecreasing sequences of compact sets $X_n \uparrow X, n \to +\infty$, and $A_n \uparrow A, n \to +\infty$ such that $\Lambda_n := X_n \times A_n$ is a subset of $\mathbb{K}$, and

$$\lim_{n \to \infty} \inf_{(x,a) \in \Lambda_n^c} c(x,a) = +\infty, \tag{3.8}$$

where Λ_n^c denotes the complement of Λ_n.

Remark (3.9). a. Notice that the inf-compactness on $\mathbb{K}$ of the cost function c implies that $\widehat{A}(x)$ is compact for every $x \in X$; moreover, it is direct to obtain that $f_n(x)$ and $f^*(x)$ belong to $\widehat{A}(x)$, for every $x \in X$ and every $n = 1, 2, \cdots$.

b. Observe that $\mathbb{K}$ is closed as an immediate consequence of Assumption (3.4a) and Proposition 7, p. 110 [2] (see also Proposition D3 in Appendix D, p. 182 [12]). In addition, it is easily seen that Assumptions (3.4a), (3.4b) and (3.4c) imply that $\widehat{\mathbb{K}}_{\varsigma}$ is closed in $X \times A$ for every nonempty compact set $\varsigma \subset X$.

LEMMA (3.10). *Suppose that Assumptions (3.4a), (3.4b), and (3.4c) hold. Then each Assumption (3.7a), (3.7b), or (3.7c) implies that $\widehat{\mathbb{K}}_{\varsigma}$ is a compact set for every nonempty compact set $\varsigma \subset X$.*

Proof. Suppose that Assumptions (3.4a), (3.4b), and (3.4c) hold. Let ς be an arbitrary, fixed, nonempty compact subset of X. Observe that $\widehat{\mathbb{K}}_\varsigma$ is closed in $X \times A$ (see Remark (3.9(b))). Therefore, a suitable compact set $\widehat{J} \subset X \times A$ such that $\widehat{\mathbb{K}}_\varsigma \subset \widehat{J}$ will be shown for each Assumption (3.7a), (3.7b), or (3.7c). This allows to conclude the compactness of $\widehat{\mathbb{K}}_\varsigma$.

Firstly, suppose Assumption (3.7a) holds. In this case, taking $\widehat{\mathbb{K}}_\varsigma \subset \widehat{J} := \varsigma \times A$, it results that $\widehat{\mathbb{K}}_\varsigma$ is compact.

Secondly, suppose Assumption (3.7b) holds. Now observe that $\cup_{x\in\varsigma} A(x)$ is a compact set (see [4] p. 72). As

$$\widehat{\mathbb{K}}_\varsigma \subset \widehat{J} := \varsigma \times \cup_{x\in\varsigma} A(x), \tag{3.11}$$

then $\widehat{\mathbb{K}}_\varsigma$ is a compact set.

Thirdly, suppose Assumption (3.7c) holds. It will be shown that there exists a positive integer n_0 such that $\widehat{\mathbb{K}}_\varsigma \subset \Lambda_{n_0}$. By contradiction, assume that for each $n = 1, 2, \cdots$, there exists $(x_n, a_n) \in \widehat{\mathbb{K}}_\varsigma$, such that $(x_n, a_n) \notin \Lambda_n$. Observe that

$$c(x_n, a_n) \le V^*(x_n) \le L := \sup_{x\in\varsigma} V^*(x) < +\infty, \tag{3.12}$$

for all $n = 1, 2, \cdots$ (recall that Assumption (3.4c) holds, hence V^* is a continuous function).

Therefore

$$\limsup_{n\to\infty} c(x_n, a_n) \le L. \tag{3.13}$$

Finally, as

$$\inf_{(x,a)\in\Lambda_n^c} c(x, a) \le c(x_n, a_n), \tag{3.14}$$

$n = 1, 2, \cdots$, then

$$+\infty = \lim_{n\to\infty} \inf_{(x,a)\in\Lambda_n^c} c(x, a) \le \limsup_{n\to\infty} c(x_n, a_n), \tag{3.15}$$

which is a contradiction to (3.13). Thus, there exists n_0 such that $\widehat{\mathbb{K}}_\varsigma \subset \widehat{J} := \Lambda_{n_0}$, which implies that $\widehat{\mathbb{K}}_\varsigma$ is a compact set.

Since ς is arbitrary, the desired result follows. □

COROLLARY (3.16). *Suppose that Assumptions (3.4a), (3.4b) and (3.4c) hold. Then each of Assumptions (3.7a), (3.7b) or (3.7c) implies that the multifunction* $x \to \widehat{A}(x)$ *is u.s.c.*

Proof. Observe that $\widehat{A}(x)$ is compact for every $x \in X$ (see Remark (3.9a)). Now suppose $x_n \mapsto x$ in X and take $a_n \in \widehat{A}(x_n)$, $n = 1, 2, \cdots$. Let $\varsigma = \{x_n\} \cup \{x\}$. Notice that ς is a compact set, so from Lemma (3.10), $\widehat{\mathbb{K}}_\varsigma$ is also compact. Therefore, since $(x_n, a_n) \in \widehat{\mathbb{K}}_\varsigma$, $n = 1, 2, \cdots$, then there exist a subsequence $\{(x_{n_k}, a_{n_k})\}$ of $\{(x_n, a_n)\}$ and $(x, a) \in \widehat{\mathbb{K}}_\varsigma$ such that

$$(x_{n_k}, a_{n_k}) \to (x, a), \tag{3.17}$$

$k \to \infty$. In particular, notice that $a_{n_k} \to a$, $k \to +\infty$ and $a \in \widehat{A}(x)$. Hence, from Lemma (2.20), the result follows. □

LEMMA (3.18). *Suppose that Assumption (3.4) and one of Assumptions (3.7a), (3.7b), or (3.7c) hold. Then the stationary optimal policy f^* is a continuous function.*

Proof. Suppose that f^* is not continuous. Then there exist $x \in X$ and a sequence $\{x_n\}$ in X such that

$$x_n \to x, \tag{3.19}$$

but

$$f^*(x_n) \nrightarrow f^*(x). \tag{3.20}$$

Then there exist $\varepsilon > 0$ and a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that

$$d\left(f^*\left(x_{n_k}\right), f^*(x)\right) \geq \varepsilon, \tag{3.21}$$

for all $k = 1, 2, \cdots$. (d denotes the metric in the control space A.)

Let $y_{n_k} := f^*\left(x_{n_k}\right), k = 1, 2, \cdots$. Observe that $y_{n_k} \in \widehat{A}\left(x_{n_k}\right)$, for all $k = 1, 2, \cdots$, and that the multifunction $x \mapsto \widehat{A}(x)$ is compact-valued, and it is also u.s.c. as a consequence of Corollary (3.16). Now, from Lemma (2.20) there exist a subsequence $\{y_{n_{k_l}}\}$ of $\{y_{n_k}\}$ and $y \in \widehat{A}(x)$ such that

$$y_{n_{k_l}} \to y. \tag{3.22}$$

Then, from (3.21) it results that

$$d(y_{n_{k_l}}, f^*(x)) \geq \varepsilon, \tag{3.23}$$

for all $l = 1, 2, \cdots$. Letting $l \to +\infty$ in (3.23), it follows that

$$d\left(y, f^*(x)\right) \geq \varepsilon. \tag{3.24}$$

On the other hand, (2.10) implies that

$$V^*(x_{n_{k_l}}) = c(x_{n_{k_l}}, y_{n_{k_l}}) + \alpha \int V^*(y)\, Q(dy \mid x_{n_{k_l}}, y_{n_{k_l}}), \tag{3.25}$$

for all $l = 1, 2, , \cdots$. Hence, letting $l \to +\infty$ and using Assumption (3.4) in (3.25), it results that

$$V^*(x) = c(x, y) + \alpha \int V^*(z)\, Q\left(dz \mid x, y\right), \tag{3.26}$$

but

$$V^*(x) = c\left(x, f^*(x)\right) + \alpha \int V^*(z)\, Q\left(dz \mid x, f^*(x)\right). \tag{3.27}$$

Therefore, from (3.26), (3.27) and the uniqueness of f^*, it follows that $y = f^*(x)$, which is a contradiction to (3.24). This completes the proof of Lemma (3.18). □

LEMMA (3.28). *Suppose that Assumption (3.4) holds. Then $\{G_n\}$ converges uniformly to G on every nonempty compact subset of $\mathbb{K}$.*

Proof. Let Θ be an arbitrary, fixed, nonempty compact subset of $\mathbb{K}$. Note that from Lemma (2.8) it results that $\{G_n\}$ converges pointwise on Θ to G, in fact, $G_n \uparrow G, n \to +\infty$. In addition, from Assumption (3.4), G and $G_n, n = 1, 2, \cdots$, are continuous functions. Then, by Dini 's Theorem ([14] p. 239), $G_n \to G$ uniformly on Θ. Since Θ is arbitrary, the result follows. □

LEMMA (3.29). *Suppose that Assumption (3.4) and one of Assumptions (3.7a), (3.7b) or (3.7c) hold. Then, for each nonempty compact set $\varsigma \subset X$, the following holds: for every $\varepsilon > 0$, there exists $\delta > 0$ such that, for all $(x, a) \in \widehat{\mathbb{K}}_\varsigma$, if*

$$\left|G\left(x, f^*(x)\right) - G(x, a)\right| < \delta, \tag{3.30}$$

then

$$d\left(f^*(x), a\right) < \varepsilon. \tag{3.31}$$

Proof. Let ς be an arbitrary, fixed, nonempty compact subset of X. The proof is by contradiction. Suppose that there exists $\varepsilon > 0$, such that for every $\delta > 0$, there is $(x_\delta, a_\delta) \in \widehat{\mathbb{K}}_\varsigma$, which satisfies

$$\left|G\left(x_\delta, f^*(x_\delta)\right) - G(x_\delta, a_\delta)\right| < \delta, \tag{3.32}$$

and

$$d\left(f^*(x_\delta), a_\delta\right) \geq \varepsilon. \tag{3.33}$$

Let $\delta = \frac{1}{n}$, $n = 1, 2, \cdots$. Then there exist $(x_n, a_n) \in \widehat{\mathbb{K}}_\varsigma$, such that

$$\left|G\left(x_n, f^*(x_n)\right) - G(x_n, a_n)\right| < \frac{1}{n}, \tag{3.34}$$

and

$$d\left(f^*(x_n), a_n\right) \geq \varepsilon, \tag{3.35}$$

for all $n = 1, 2, \cdots$. Now, since $\widehat{\mathbb{K}}_\varsigma$ is a compact set, there exist a subsequence $\left\{\left(x_{n_k}, a_{n_k}\right)\right\}$ of $\{(x_n, a_n)\}$ and $(x, a) \in \widehat{\mathbb{K}}_\varsigma$, such that $\left(x_{n_k}, a_{n_k}\right) \to (x, a)$, $k \to \infty$. Notice that

$$\left|G\left(x_{n_k}, f^*\left(x_{n_k}\right)\right) - G\left(x_{n_k}, a_{n_k}\right)\right| < \frac{1}{n_k} \tag{3.36}$$

for all $k = 1, 2, \cdots$. Therefore, letting $k \to \infty$ and using Assumption (3.4) and Lemma (3.18) in (3.36), it follows that

$$G(x, a) = G\left(x, f^*(x)\right). \tag{3.37}$$

This implies that

$$a = f^*(x), \tag{3.38}$$

(recall the uniqueness of f^*).

Nevertheless, from (3.35)

$$d\left(f^*\left(x_{n_k}\right), a_{n_k}\right) \geq \varepsilon, \tag{3.39}$$

for all $k = 1, 2, \cdots$.

So, if $k \to \infty$ in the last inequality and using Lemma (3.18),

$$d\left(f^*(x), a\right) \geq \varepsilon, \tag{3.40}$$

which is a contradiction to (3.38). Since ς is arbitrary, the proof of Lemma (3.29) is finished. □

Now, the uniform convergence on compact sets of minimizers of the value iteration algorithm of Discounted MDPs to optimal policies will be proved.

THEOREM (3.41). *Suppose that Assumption (3.4) and one of Assumptions (3.7a), (3.7b), or (3.7c) hold. Then $\{f_n\}$ converges uniformly on compact sets to f^*.*

Proof. Let ς be an arbitrary, fixed, non-empty compact of X. Observe that $f_n(x)$ minimizes to $G_n(x,\cdot)$ on $A(x)$, and $f^*(x)$ is the (unique) minimizer to $G(x,\cdot)$ on $A(x)$, for all $x \in \varsigma$, and for all n. Therefore,

$$\begin{aligned} 0 &\leq G\left(x, f_n(x)\right) - G\left(x, f^*(x)\right) \\ &\leq G\left(x, f_n(x)\right) - G_n\left(x, f_n(x)\right) \\ &\quad + G_n\left(x, f^*(x)\right) - G\left(x, f^*(x)\right) \\ &\leq 2 \sup_{(x,a)\in\widehat{\mathbb{K}}_\varsigma} \left|G(x,a) - G_n(x,a)\right|, \end{aligned} \tag{3.42}$$

for all $x \in \varsigma$, and for all n.

Now, let $\varepsilon > 0$. By Lemma (3.29) applied to ς, there exists $\delta > 0$ such that for all $(x,a) \in \widehat{\mathbb{K}}_\varsigma$, if

$$\left|G\left(x, f^*(x)\right) - G(x,a)\right| < \delta, \tag{3.43}$$

then

$$d\left(f^*(x), a\right) < \varepsilon. \tag{3.44}$$

Furthermore, Lemma (3.28) guarantees the existence of a positive integer R such that

$$2 \sup_{(x,a)\in\widehat{\mathbb{K}}_\varsigma} \left|G(x,a) - G_n(x,a)\right| < \delta, \tag{3.45}$$

if $n \geq R$. Now, from (3.42) and (3.45), if $n \geq R$,

$$\left|G\left(x, f^*(x)\right) - G\left(x, f_n(x)\right)\right| < \delta, \tag{3.46}$$

for all $x \in \varsigma$.

Combining (3.46), (3.43) and (3.44) it follows that

$$d\left(f^*(x), f_n(x)\right) < \varepsilon, \tag{3.47}$$

for all $n \geq R$, and for all $x \in \varsigma$. So, $f_n \to f^*$ uniformly on ς.

Since ς is arbitrary, the result follows. □

4. Examples

Remark (4.1). Consider MDPs that satisfy Assumptions (2.3), (2.11), (3.4a) and (3.4b) and that both integrals (3.5) and (3.6) are finite.

Concerning the continuity required in Assumptions (3.4c) and (3.4d), for $V_n, n = 1, 2, \cdots, V^*$, and the integrals (3.5) and (3.6), observe the following.

a. The continuity mentioned trivially holds for discrete models (i.e., MDPs for which both X and A are finite or denumerable sets).

b. For bounded models (i.e., MDPs with bounded cost functions and compact admissible action sets), the continuity of the integrals (3.5) and (3.6) follows directly from the strong continuity of the transition law Q. If moreover, the multifunction $x \mapsto A(x)$ is continuous (see Definition (2.19c)), then the continuity of $V_n, n = 1, 2, \cdots$, and V^* is an immediate consequence of Proposition

D.3(c) p. 130 in [11], using the continuity of the cost function c and of the integrals (3.5) and (3.6), and equations (2.5) and (2.9).

For such a bounded model, see Example (4.2) below.

c. In [8] two convexity conditions have been presented (see Conditions C1 and C2 in [8]), each of which guarantees (in particular) that V_n, $n = 1, 2, \cdots, V^*$, and the integrals (3.5) and (3.6) are convex functions (see Lemma 6.2 p. 433 and its proof in [8]). If, in addition, the spaces X, A and $\mathbb{K}$ are open sets, then the continuity required in Assumptions (3.4c) and (3.4d) is obtained (see Theorem 3, p. 113 [3]). This is illustrated in Examples (4.9) and (4.15) below.

An inventory/production model is presented below (see [11] p. 10 for the precise description of this example in the context of inventory/production area).

Example (4.2). Let M be a fixed positive constant. Let $X = A = \left[0, M\right]$, $A(x) = \left[0, M - x\right]$, $x \in X$, and consider

$$x_{t+1} = [x_t + a_t - \xi_t]^+ , \tag{4.3}$$

$t = 0, 1, \cdots$, where $z^+ := \max\{0, z\}$. Here $\xi_0, \xi_1, \cdots$ are i.i.d. random variables taking values in $S = [0, \infty)$, and with common density Δ.

Assumption (4.4). a. Δ is a bounded continuous function. (Notice that the distribution function $\widehat{G}$ of ξ is a continuous function, where ξ is a generic element of the sequence $\{\xi_t\}$.)

b. c is non-negative, continuous and strictly convex on $\mathbb{K}$ (observe that $\mathbb{K}$ in this example is a compact set); also c is an increasing function in the first variable.

LEMMA (4.5). a. *Example (4.2) satisfies Assumptions (2.3), (2.11), (3.7a), and (3.7b).*

b. *The Assumption (3.4) holds.*

Proof. a. Example (4.2) clearly satisfies Assumptions (2.3a) and (2.3c). In reference to Assumption (2.3b), notice that if $\mu \colon X \to \mathbb{R}$ is a measurable and bounded function, then a simple computation shows that

$$\begin{aligned} \int \mu(y) Q\left(dy \mid x, a\right) \quad &= \quad \mu(0)\left[1 - \widehat{G}(x+a)\right] \\ &\quad + \int I_{[0,x+a]}(u)\, \mu(u)\, \Delta(x + a - u)\, du, \end{aligned} \tag{4.6}$$

$(x, a) \in \mathbb{K}$, where $I_{[\cdot]}$ denotes the indicator function of the subset $[\cdot]$.

Since $\widehat{G}$ is a continuous function, it follows that $\mu(0)\left[1 - \widehat{G}(x+a)\right]$ is a continuous function on $\mathbb{K}$. Now, as μ is a bounded function and Δ is a bounded continuous function, it follows directly, using the Dominated Convergence Theorem, that

$$\int I_{[0,x+a]}(u)\, \mu(u)\, \Delta(x + a - u)\, du \tag{4.7}$$

is a continuous function on $\mathbb{K}$. Hence

$$\int \mu(y) Q\left(dy \mid \cdot, \cdot\right) \tag{4.8}$$

is a continuous function on $\mathbb{K}$.

Observe that by means of straightforward computations it is possible to verify that this Example satisfies Condition C1 in [8], therefore, the uniqueness of the optimal policy holds.

Finally, notice that Assumptions (3.7a) and (3.7b) trivially hold.

b. Firstly, it is evident that $A(x)$ is a closed set for each $x \in X$. Secondly, since c is bounded, it follows that for each $n = 1, 2, \cdots$, V_n and also V^* are bounded and they are also measurable from Assumption (2.3) (which holds from item a). Therefore, from the fact that V_n and V^* are bounded, it follows that the integrals (3.5) and (3.6) are finite; and the strong continuity of Q allows to conclude that (3.5) and (3.6) are continuous on $\mathbb{K}$. Finally, direct computations show that the multifunction $x \mapsto A(x)$ is continuous (i.e. $x \mapsto A(x)$ is both u.s.c. and l.s.c. (see Definition (2.19) and Lemma (2.20)). Hence, the continuity of the integrals (3.5) and (3.6), the continuity of the cost c, (2.5), (2.9) and Proposition D.3 (c) p. 130 in [11] imply that V_n and V^* are continuous. □

Example (4.9). Consider a simple linear system

$$x_{t+1} = \widehat{\gamma} x_t + \delta a_t + \xi_t, \quad t = 0, 1, \cdots, \tag{4.10}$$

with quadratic cost

$$c(x, a) = qx^2 + ra^2, \tag{4.11}$$

$x, a \in \mathbb{R}$. Here $X = A = A(x) = \mathbb{R}$, for all $x \in X$.

Assumption (4.12). a. $\widehat{\gamma}\delta \neq 0$, both q and r are positive.

b. The disturbances ξ_t, $t = 0, 1, \cdots$ are i.i.d. random variables with values in $S = \mathbb{R}$. Moreover, suppose that ξ_0 has a continuous density Δ, zero mean value and a finite variance $\sigma^2 > 0$.

Remark (4.13). Assumptions (2.3) and (2.11) have been proved in Example 4.8 in [8] (in particular, Condition C2 in [8] holds for this model). This Example satisfies trivially Assumptions (3.4a), (3.4b), and (3.7c). On the other hand, it is easy to verify that $W(x) = \bar{l}\left(x^2 + \left(\alpha/(1-\alpha)\right)\sigma^2\right)$, with $\bar{l} := q + r\left(\widehat{\gamma}^2/\delta^2\right)$, is an upper bound for $V_n(x)$, $n = 1, 2, \cdots$, and $V^*(x)$, $x \in X$. In fact, $W(x) = V\left(f, x\right)$, $x \in X$, for $f(x) = -\left(\widehat{\gamma}/\delta\right)x$, $x \in X$ (see Lemma 4.9(c) in [8]). Since

$$\int W(y)\, Q\left(dy \mid x, a\right) = \bar{l}\left[\left(\widehat{\gamma}x + \delta a\right)^2 + \sigma^2\right] + \frac{\bar{l}\alpha\sigma^2}{1-\alpha} < \infty, \tag{4.14}$$

$(x, a) \in \mathbb{R}^2$, it follows that the integrals (3.5) and (3.6) are finite. Finally, since for this Example the Condition C2 in [8] holds, hence from Lemma 6.2 (and its proof) in [8] it follows that V_n and V^* are convex on $\mathbb{R}$, and the integrals (3.5) and (3.6) are convex on $\mathbb{R}^2$. Consequently, Assumptions (3.4c) and (3.4d) hold (see Remark (4.1)).

Now an example of a nonlinear additive-noise system with unbounded cost function will be presented.

Example (4.15). Consider

$$x_{t+1} = x_t + a_t^2 + \xi_t, \quad t = 0, 1, \cdots, \tag{4.16}$$

with cost

$$c(x,a) = h(x) + g(a), \tag{4.17}$$

where $h(x) = e^x$, $g(a) = 2a^4 + a + 1$, $x, a \in \mathbb{R}$, and $X = A = A(x) = \mathbb{R}$, for all $x \in X$.

Assumption (4.18). The disturbances ξ_t, $t = 0, 1, \cdots$ are i.i.d. random variables with values in $S = \mathbb{R}$. Moreover, suppose that ξ_0 has a continuous density Δ, and

$$k := \int e^s \Delta(s)\, ds < \infty, \tag{4.19}$$

with $0 < \alpha k < 1$.

Remark (4.20). Clearly, c is non-negative (in fact, $h(x) > 0, x \in X$ and $g(a) \geq 5/8, a \in A$, therefore $c(x,a) > 5/8$ for all $(x,a) \in \mathbb{K}$), and continuous, and Assumptions (3.4a) and (3.4b) hold. The inf-compactness of c in $\mathbb{K}$ and Assumption (3.7c) follow directly from the fact that

$$\lim_{a \to +\infty} g(a) = \lim_{a \to -\infty} g(a) = +\infty. \tag{4.21}$$

Also this Example trivially satisfies Condition C1 in [8] (notice that this Example does not satisfy Condition C2 in [8]). Hence Assumption (2.11) follows.

LEMMA (4.22). *a. Q defined by (4.16) is strongly continuous.*
b. There is $\pi \in \Pi$ such that $V(\pi, x) < \infty$.

Proof. a. (4.16) and the (well-known) Change of Variable Theorem permit to get that

$$Q\left(B \mid x, a\right) = \int I_B\left[x + a^2 + s\right] \Delta(s)\, ds = \int_B \Delta(s - \left(x + a^2\right))ds \tag{4.23}$$

for $x \in X, a \in A$ and $B \in \mathcal{B}(X)$. Hence from Assumption (4.18) and from Example C.6 in Appendix C in [12], it results that Q is strongly continuous.

b. Let $f \in \mathbb{F}$ given $f(x) = 0$, for all $x \in X$. Then it is possible to prove by straightforward induction argument that

$$E_x^f\left[c\left(x_t, f\right)\right] = k^t e^x + 1, \tag{4.24}$$

for each $t = 0, 1, \cdots$, and $x \in X$, where k was defined in Assumption (4.18). Now, from (4.24) and Assumption (4.18), it results that, for each $x \in X$:

$$V\left(f, x\right) = \sum_{t=0}^{\infty} \alpha^t E_x^f c\left(x_t, f\right) = \frac{e^x}{1 - \alpha k} + \frac{1}{1 - \alpha} < \infty. \tag{4.25}$$

□

Remark (4.26). Finally, similar to the previous Example (see Remark (4.13)), Assumptions (3.4c), and (3.4d) follow from Lemma 6.2 (and its proof) in [8].

5. Remarks on the Pointwise Convergence of Minimizers

Remark (5.1). As a consequence of Theorem (3.41), it is easily obtained that $\{f_n\}$ converges pointwise to f^* on X. Actually, for the pointwise convergence of $\{f_n\}$ to f^*, fewer and weaker assumptions are needed than those proposed for the uniform convergence on compact sets of $\{f_n\}$ to f^* (see also Remark (5.14) below); in fact, it is sufficient with a weaker version of Assumption (3.4) (see Assumption (5.2) below), and Assumption (3.7) is not necessary.

Assumption (5.2). a. For each $x \in X$, $c(x, \cdot)$ is a continuous function on $A(x)$.

b. For each $x \in X$,

$$\int V_n(y) Q\left(dy \mid x, \cdot\right), n = 1, 2, \cdots, \tag{5.3}$$

$$\int V^*(y) Q\left(dy \mid x, \cdot\right), \tag{5.4}$$

are finite and continuous functions on $A(x)$.

LEMMA (5.5). *Suppose that Assumption (5.2) holds. Then for each $x \in X$, $G_n(x, \cdot) \to G(x, \cdot)$ uniformly on every nonempty compact subset of $A(x)$.*

Proof. It is similar to the proof of Lemma (3.28). □

THEOREM (5.6). *Suppose that Assumption (5.2) holds. Then $f_n(x) \to f^*(x)$, $n \to +\infty$ for each $x \in X$.*

Proof. The proof is by contradiction. Suppose that there exists $x \in X$ such that $f_n(x) \nrightarrow f^*(x)$. Let $\{f_{n_k}(x)\}$ be a subsequence of $\{f_n(x)\}$, and $\varepsilon > 0$ such that

$$d\left(f_{n_k}(x), f^*(x)\right) \geq \varepsilon, \tag{5.7}$$

for all $k = 1, 2, \cdots$.

Since $\widehat{A}(x)$ is a compact set (see Remark (3.9a)), there exist $a_x \in \widehat{A}(x)$, and a subsequence $\{f_{n_{k_l}}(x)\}$ of $\{f_{n_k}(x)\}$, such that

$$f_{n_{k_l}}(x) \to a_x, \tag{5.8}$$

$l \to +\infty$. Observe that from (5.7) it results that

$$d\left(a_x, f^*(x)\right) \geq \varepsilon. \tag{5.9}$$

Now note that by Assumption (5.2), Lemma (2.13), and Lemma (5.5) specialized to the compact $\widehat{A}(x)$, it follows that

$$\lim_{l \to +\infty} G_{n_{k_l}}(x, f_{n_{k_l}}(x)) = G(x, a_x). \tag{5.10}$$

On the other hand, for each $l \geq 1$,

$$G_{n_{k_l}}(x, f_{n_{k_l}}(x)) = V_{n_{k_l}}(x). \tag{5.11}$$

Hence from (2.10), (5.10), (5.11) and Lemma (2.8b), it follows that

$$\begin{aligned}\lim_{l\to+\infty} G_{n_{k_l}}(x, f_{n_{k_l}}(x)) &= V^*(x) \\ &= G\left(x, f^*(x)\right) \\ &= G(x, a_x). \end{aligned} \tag{5.12}$$

Then by uniqueness of f^* it results that $f^*(x) = a_x$, which is a contradiction to (5.9). This completes the proof of Theorem (5.6) □

Now a corollary with two results which represent special cases of Theorem (5.6) will be shown. They are related to the concept known in the literature of MDPs as the Forecast Horizon (see [6], [7], [10], [13], [19], [20], [21] and [22]).

COROLLARY (5.13). *If A is a finite set or a denumerable set, then for each $x \in X$ there exists a positive integer $N^*(x)$ such that $f_n(x) = f^*(x)$, for all $n \geq N^*(x)$, supposing that integrals (5.3) and (5.4) are finite.*

Proof. Fix $x \in X$. Suppose that A is a finite set or a denumerable set with the discrete metric. Since in this case Assumption (5.2) trivially holds, it follows that $f_n(x) \to f^*(x)$. Then there exists a positive integer $N^*(x)$ such that $f_n(x) \in \{f^*(x)\}$ for all $n \geq N^*(x)$ (recall that in the discrete metric, $\{f^*(x)\}$ is an open set). Therefore $f_n(x) = f^*(x)$ for all $n \geq N^*(x)$. Since x is arbitrary, the result follows. □

Remark (5.14). Observe that under Assumptions (2.3) and (2.11), the pointwise convergence of $\{f_n\}$ to f^* is a direct consequence of Theorem 4.6.5, p. 67 in [12] provided the control space A is locally compact. On the other hand, Proposition 12.2 in [18] yields the pointwise convergence for the case of admissible compact actions sets and under Assumption (2.11).

Acknowledgement

The authors wish to thank two anonymous referees and the associate editor for helpful comments and for suggesting us Remark (5.14).

Received September 29, 2004

Final version received December 02, 2005

DANIEL CRUZ SUÁREZ
DIVISIÓN ACADÉMICA DE CIENCIAS BÁSICAS
UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO
APDO. POSTAL 5
86690 CUNDUACÁN, TAB.
MÉXICO
daniel.cruz@basicas.ujat.mx

RAÚL MONTES-DE-OCA
DEPARTAMENTO DE MATEMÁTICAS UNIVERSIDAD AUTÓNOMA METROPOLITANA-IZTAPALAPA
AV. SAN RAFAEL ATLIXCO 186, COL. VICENTINA
09340 MÉXICO D.F.
MÉXICO
momr@xanum.uam.mx

References

[1] C. D. Aliprantis, and K. C. Border, Infinite Dimensional Analysis, 2nd edition, Springer, Berlin, 1999.

[2] J. P. Aubin, and I. Ekeland, Applied Nonlinear Analysis, John Wiley, New York, 1984.

[3] J. P. Aubin, Applied Abstract Analysis, John-Wiley, New York, 1977.

[4] J. P. Aubin, Mathematical Methods of Game and Economic Theory, North-Holland, New York, 1979.

[5] D. P. Bertsekas, Dynamic Programming: Deterministic and Stochastic Models, Prentice-Hall, NJ, 1987.

[6] C. Bes, S. P. Sethi, *Concepts of forecast and decision horizons: Applications to dynamic stochastic optimization problems*, Math. Oper. Res. **13** (2), (1988), 295–310.

[7] S. Chand, S. P. Sethi, and J. M. Proth, *Existence of forecast horizons in undiscounted discrete-time lot size models*, Oper. Res. **38** (5), (1990), 884–892.

[8] D. Cruz-Suárez, R. Montes-de-Oca, and F. Salem-Silva, *Conditions for the uniqueness of optimal policies of discounted Markov decision processes*, Math. Methods Oper. Res. **60** (3), (2004), 415–436.

[9] J. Dugundji, Topology, Allyn and Bacon Inc., Boston, 1966.

[10] W. J. Hopp, J. C. Bean, and R. L. Smith, *A new optimality criterion for nonhomogeneous Markov decision processes*, Oper. Res., **35** (6), (1987), 875–883.

[11] O. Hernández-Lerma, Adaptive Markov Control Processes. Springer-Verlag, New York 1989.

[12] O. Hernández-Lerma, and J. B. Lasserre, Discrete-Time Markov Control Processes. Springer-Verlag, New York 1996.

[13] O. Hernández-Lerma, and J. B. Lasserre, Further Topics on Discrete-Time Markov Control Processes. Springer-Verlag, New York, 1999.

[14] J. L. Kelley, General Topology, Springer-Verlag, 1975.

[15] H.-L. Langen, *Convergence of dynamic programming models*, Math. Oper. Res. **6** (4), (1981), 493-512.

[16] M. L. Puterman, Markov Decision Processes: Discrete Stochastic Dynamic Programming. Wiley, New York 1994.

[17] M. Schäl, *A selection theorem for optimization problems*, Arch. Math., V. XXV, (1974), 219–224.

[18] M. Schäl, *Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal*, Z. Wahrsch. Verw. Geb. **32** (1975), 179–196.

[19] I. E. Schochetman, and R. L. Smith, Infinite horizon optimization, Math. Oper. Res. **14** (3), (1989), 559–574.

[20] S. P. Sethi, and S. Bhaskaran, *Conditions for the existence of decision horizons for discounted problems in a stochastic enviroment: a note*, Oper. Res. Lett. **4** (2), (1985), 61–64.

[21] J. F. Shapiro, *Turnpike planning horizons for a markovian decision model*, Manage. Sci. **14** (5), (1968), 292–300.

[22] R. L. Smith, and R. Q. Zhang, *Infinite horizon production planning in time-varying systems with convex production and inventory costs*. Manage. Sci. **44** (9), (1998), 1313–1320.

[23] N. L. Stokey, and R. L. Jr. Lucas, Recursive Methods in Economic Dynamics. Harvard University Press, Cambridge, MA, 1989.