Transilvania University of Braşov

HABILITATION THESIS

Stochastic Processes, Analysis, and Applications

Mihai Nicolae Pascu

Specialization: Mathematics

2014

To Bianca, Paul, Luca, Nicolae,
Olivia, Ana and Ştefan

Rezumat

Prezenta Teză de Abilitare prezintă o parte din cele mai recente şi semnificative rezultate ale autorului în domeniul Proceselor Stochastice, în principal legate de construcţia, proprietăţile, şi aplicaţiile mişcării Browniene, şi Analiza Complexă, în special în domeniul Teoriei Geometrice a Funcţiilor.

Materialul prezentat în această teză este împărţit în şase capitole, şi reflectă interesul autorului pentru cele două arii de cercetare distincte menţionate mai sus, şi, atunci când este posibil, a legăturilor dintre cele două.

Prima parte a tezei reflectă interesul autorului în domeniul Procelor Stochastice, şi este formată din trei capitole. Ideea din spatele acestei împărţiri este că mişcarea Browniană este invariantă la trei transformări geometrice: scalare, reflecţie/simetrie, şi translaţie. Corespunzător fiecărei din aceste proprietăţi de invarianţă autorul a introdus (sau a contribuit la dezvoltarea teoriei cunoscute a) cuplajelor corespunzătoare de mişcări Browniene, şi le-a folosit ca unelte în demonstrarea unor rezultate importante în Analiză. Astfel, Capitolul 1 prezintă construcţia cuplajului prin scalare de mişcări Browniene reflectate, Capitolul 2 prezintă contribuţia autorului la construcţia generală a cuplajului în oglindă de mişcări Browniene reflectate, iar Capitolul 3 prezintă un rezultat foarte recent al autorului în colaborare cu I. Popescu, cu privire la cuplajele de mişcări Browniene reflectate cu distanţă fixă, cuplaje ce pot fi privite ca o extensie a cuplajului prin translaţie la domenii generale.

Partea a doua a tezei prezintă câteva din rezultatele autorului în Teoria geometrică a funcţiilor, şi conţine trei capitole. Capitolul 4 prezintă o extensie a Principiului maximului modulului din Analiza Complexă la cazul funcţiilor ne-analitice, şi câteva aplicaţii ale acestuia. Capitolul 5 prezintă o metodă de construcţie a celei mai bune aproximări univalente a unei funcţii analitice în anumite subclase de funcţii (stelate, respectiv convexe), iar Capitolul 6 prezintă rezultate referitoare la univalenţa perturbărilor funcţiilor analitice şi alte rezultate conexe.

Toate rezultatele demonstrate în teză aparţin autorului, fiind publicate în lucrările precizate mai jos. Structura şi conţinutul tezei sunt următoarele.

Capitolul 1. Cuplaje prin scalare de mişcări Browniene reflectate.

În acest capitol prezentăm construcţia *cuplajului prin scalare* de mişcări Browniene reflectate, cuplaj introdus de către autor. Constricţia este dată mai întâi în cazul discului unitate, iar apoi este extinsă la domenii $C^{1,\alpha}$ netede ($0 < \alpha < 1$) cu ajutorul transformărilor conforme. Ca o aplicaţie, obţinem o rezolvare parţială a faimoasei conjecturi a *Punctelor Fierbinţi* a lui Jeffrey Rauch (1974), care afirmă că o funcţie proprie Neumann corespunzătoare celei de a doua valori proprii Neumann a operatorului Laplacian într-un domeniu convex verifică o variantă tare a principiului de maxim. Mai general, arătăm aici că o a doua funcţie proprie Neumann antisimetrică a unui domeniu cu o axă de simetrie este de fapt monotonă de-a lungul geodezicelor hiperbolice, şi prezentăm o altă extensie a acestui rezultat la cazul când linia nodală este un arc de cerc.

Materialul din acest capitol se bazează pe rezultatele obţinute de autor în [10] (cu R. Bañuelos şi M. Pang) şi [64].

Capitolul 2. Cuplajul în oglindă de mişcări Browniene reflectate.

Cuplajul în oglindă de mişcări Browniene a fost introdus de W. S. Kendall în [50] în cazul mişcărilor Browniene pe un manifold Riemannian complet având curbură Ricci ne-negativă, şi a fost considerat de către F. Y. Wang în [85] în cazul proceselor reflectate. În [25], şi mai recent în [4] şi [5], K. Burdzy et al. au dat o construcţie detaliată a cuplajului în oglindă de mişcări Browniene reflectate într-un domeniu neted din $\mathbb{R}^n$ ($n \geq 2$). În acest capitol prezentăm o extensie a acestei construcţii, datorată autorului, la cazul general când cele două mişcări Browniene trăiesc în domenii diferite. Ca aplicaţii, prezentăm o rezolvare a conjecturii *Laugesen-Morpurgo* asupra monotoniei diagonalei nucleului Neumann al căldurii în bila unitate din $\mathbb{R}^n$ ($n \geq 1$), o demonstraţie unitară a două din cele mai importante rezultate referitoare la validitatea conjecturii lui Chavel privind monotonia în raport cu domeniul a nucleului Neumann al căldurii (rezultate datorate lui I. Chavel [29], respectiv lui W. S. Kendall [51]), şi o nouă demonstraţie a acesteia în cazul domeniilor ce verifică condiţia intermediară a bilei, astfel încât domeniul interior este stelat în raport cu centrul bilei.

Materialul din acest capitol se bazează pe rezultatele obţinute de autor în [65], [66] (cu M. E. Gageonea), şi [67] (cu M. A. Nicolaie).

Capitolul 3. Cuplaje de mişcări Browniene cu distanţa fixă.

Recent, împreună cu I. Popescu (IMAR şi Georgia Tech University), am investigat problema existenţei cuplajelor de mişcări Browniene cu distanţă fixă pe manifolduri complete. În acest capitol prezentăm câteva rezultate foarte recente, care arată spre exemplu că în cazul sferei 2-dimensionale este posibil să construim un cuplaj având distanţă fixă, precum şi cuplaje pentru care distanţa dintre cele două mişcări Browniene creşte, respectiv scade, la o rată exponenţială. Aceste rezultate sunt apoi extinse, mai întâi la cazul manifoldurilor Riemanniene având curbură constantă, iar apoi la cazul general al manifoldurilor Riemanniene având curbură Ricci mărginită inferior şi curbură secţională mărginită superior.

Ca aplicaţii, prezentăm o soluţie a unei versiuni stochastice a problemei *Leul şi Omul* a lui Rado, şi o demonstraţie a principiului de maxim pentru gradientul funcţiilor armonice.

Materialul din acest capitol se bazează pe rezultatele obţinute de autor în [73] (cu I. Popescu).

Capitolul 4. Un principiu de maxim al modulului pentru funcţii ne-analitice definite în discul unitate

Principiul maximului modulului din Analiza Complexă afirmă că maximul modulului unei funcţii analitice ne-constante definite într-un domeniu simplu conex nu poate fi atins într-un punct interior al domeniului, şi este uşor de văzut că acest principiu este fals dacă se renunţă la ipoteza de analiticitate a funcţiei. În acest capitol arătăm că putem extinde principiul maximului modulului la anumite clase de funcţii ne-analitice definite în discul unitate. Ca şi consecinţe obţinem o nouă demonstraţie a principiului clasic al modulului pentru funcţii analitice, condiţii simple asupra coeficienţilor dezvoltării în serie pentru care principiul de maxim al modulului are loc, precum şi aplicaţii la cazul funcţiilor reale de două variabile reale.

Materialul din acest capitol se bazează pe rezultatele obţinute de autor în [36] (cu M. E. Gageonea, S. Owa, şi N. R. Pascu) şi [37] (cu M. E. Gageonea şi N. R. Pascu).

Capitolul 5. Aproximări univalente ale funcţiilor analitice

Univalenţa unei funcţii analitice este o problemă importantă a Teoriei geometrice a funcţiilor, şi există multe condiţii suficiente de univalenţă în literatură (a se vedea spre exemplu monografiile [34], [74], sau [75]). Atunci când o funcţie analitică nu este univalentă, în probleme practice este deseori de interes să se găsească o "cea mai bună" aproximare a funcţiei printr-o funcţie univalentă.

În acest capitol introducem o măsură a ne-univalenţei unei funcţii analitice, pe care o folosim pentru a găsi cea mai bună aproximare a unei funcţii analitice normate în anumite subclase de funcţii univalente (funcţii stelate, respectiv convexe). Arătăm că problemele corespunzătoare pot fi reduse la anumite probleme semi-infinite de programare pătratică, pe care le rezolvăm explicit, conducând astfel la o metodă de găsire a celei mai bune aproximări stelate, respectiv a celei mai bune aproximări convexe. Rezultatele obţinute conţin algoritmi constructivi pentru determinarea explicită a măsurilor de (ne)stelaritate, respectiv de (ne)convexitate a unei funcţii analitice, precum şi pentru găsirea celei mai bune aproximări stelate corespunzătoare, respectiv a celei mai bune aproximări convexe, fiind adecvate pentru implementare numerică şi pentru aplicaţii practice.

Materialul din acest capitol se bazează pe rezultatele obţinute de autor în [71] and [72] (ambele cu N. R. Pascu).

Capitolul 6. Vecinătăţi ale funcţiilor univalente

Continuând studiul funcţiilor univalente, în acest capitol considerăm problema perturbărilor funcţiilor univalente. Ca o măsură a (ne)univalenţei unei funcţii introducem constanta $K(f, D)$ asociată unei funcţii $f : D \subset \mathbb{C} \to \mathbb{C}$ analitice într-un domeniu D, şi o folosim pentru a arăta că o mică perturbare a unei funcţii univalente este de asemenea o funcţie univalentă. Ca şi consecinţă, arătăm că o funcţie univalentă are o vecinătate constând în întregime din funcţii univalente.

Ca şi aplicaţii ale rezultatului principal obţinem o consecinţă ce este echivalentă cu criteriul clasic de univalenţă Noshiro-Warschawski-Wolff, şi prezenăm o aplicaţie în termeni de serii Taylor.

Materialul din acest capitol se bazează pe rezultatele obţinute de autor în [70] (cu N. R. Pascu).

Capitolul 7. Realizări şi planuri de dezvoltare a carierei

Ultimul capitol al tezei este împărţit în două secţiuni. Prima secţiune conţine o prezentare a realizărilor ştiinţifice şi profesionale ale autorului, şi planuri de evoluţie şi de dezvoltare a carierei profesionale. În a doua secţiune prezentăm câteva probleme deschise legate de cercetările prezentate în capitolele anterioare, precum şi câteva idei ce pot conduce la o posibilă rezolvare a acestora.

Teza se încheie cu o listă de referinţe bibliografice.

Braşov, Mai 2014

Mulţumiri

Îmi exprim recunoştinţa pentru toţi cei care de-a lungul anilor şi-au pus amprenta şi au contribuit la formarea şi dezvoltarea mea personală şi profesională: profesori, colegi, colaboratori, cercetători, studenţi, şi în sfârşit, dar nu în cele din urmă, tuturor membrilor dragii mele familii.

Cum lista celor cărora sunt îndatorat pentru a le mulţumi este mult prea lungă şi există riscul ca această listă să nu fie completă sau să nu fie la zi, voi spune doar

“Vă mulţumesc TUTUROR!”,

minimizând astfel probabilitatea de a uita să mulţumesc cuiva.

Summary

The present Habilitation Thesis presents a part of the most recent and significant results of the author in the area of Stochastic Processes, mostly related to the construction, properties, and applications of couplings of Brownian motions, and Complex Analysis, especially in the area of Geometric Function Theory.

The material presented in this thesis is divided into six chapters, in accordance with the interest of the author for the two distinct areas of research mentioned above, and, whenever possible, of the interplay between the two.

The first part of the thesis, reflecting the interest of the author in Stochastic Processes, contains three chapters. The underlining idea for this division is that the Brownian motion is invariant under three geometric transformations: scaling, reflection/symmetry, and translation. Corresponding to each of these invariance properties, the author introduced (or contributed to the development of the known theory of) corresponding couplings of reflecting Brownian motions, and used them as tools in proving important results in Analysis. As such, the Chapter 1 presents the construction of the scaling coupling of reflecting Brownian motions, Chapter 2 presents the contribution of the author to the general construction of the mirror coupling of reflecting Brownian motion, and Chapter 3 presents a very recent result of the author with I. Popescu regarding fixed distance couplings of reflecting Brownian motion, which may be viewed as an extension of the translation coupling to general domains.

The second part of the thesis presents some of the results of the author in the area of Geometric function theory, and contains three chapters. Chapter 4 presents an extension of the classic Maximum modulus principle from Complex Analysis to the case of non-analytic functions, and some of its applications. Chapter 5 presents a method for constructing the best univalent approximations of analytic functions in certain subclasses of functions (starlike, respectively convex), and Chapter 6 presents results related to the univalency of perturbations of analytic functions and other connected results.

All the results proved in the thesis belong to the author, being published in the papers indicated below. The structure and the contents of the thesis is the following.

Chapter 1. Scaling coupling of reflecting Brownian motions.

In this chapter we present the construction of the *scaling coupling* of reflecting Brownian motions, introduced by the author. The construction is first given in the case of the unit disk, and then extended to smooth $C^{1,\alpha}$) ($0 < \alpha < 1$) domains by means of conformal maps. As an application, we derive a partial resolution of the famous *Hot Spots* conjecture of Jeffrey Rauch (1974), which asserts that the second Neumann eigenfunctions of the Laplacian in convex domains satisfy a strong maximum principle. More generally, we prove here that antisymmetric second Neumann eigenfunctions of domains with an axis of symmetry are in fact monotone along hyperbolic geodesics, and we present another extension of this result to the case when the nodal line is an arc of a circle.

The material in this chapter is based on the results obtained by the author in [10] (with R. Bañuelos and M. Pang) and [64].

Chapter 2. Mirror coupling of reflecting Brownian motions.

The *mirror coupling* of Brownian motions was introduced W. S. Kendall in [50] in the case of Brownian motions on a complete Riemannian manifold with nonnegative Ricci curvature, and was considered by F. Y. Wang in [85] in the case of reflected processes. In [25], and more recently in [4] and [5], K. Burdzy et al. gave a detailed construction of the mirror coupling of reflecting Brownian motions in a smooth domain in $\mathbb{R}^n$ $(n \geq 2)$. In this chapter we present an extension this construction, due to the author, to the general case when the two reflecting Brownian motions live in different domains. As applications, we present a resolution of the *Laugesen-Morpurgo* conjecture on the monotonicity of the diagonal of the Neumann heat kernel in the unit ball in $\mathbb{R}^n$ $(n \geq 1)$, a unifying proof of the two main results concerning the validity of Chavel's conjecture on the domain monotonicity of the Neumann heat kernel (due to I. Chavel [29], respectively W. S. Kendall [51]), and a new proof of it for domains satisfying the intermediate ball condition, such that the inner domain is star-shaped with respect to the center of the ball.

The material in this chapter is based on the results obtained by the author in [65], [66] (with M. E. Gageonea), and [67] (with M. A. Nicolaie).

Chapter 3. Fixed-distance coupling of reflecting Brownian motions.

Recently, with I. Popescu (IMAR and Georgia Tech University), we investigated the problem of the existence of fixed-distance couplings of reflecting Brownian motions on complete manifolds. In this chapter we present some very recent results, which show for example that in the case of the 2-dimensional sphere it is possible to construct a fixed-distance coupling, and also couplings for which the distance between the two Brownian motions increases, respectively decreases, at an exponential rate. These results are then extended, first to the case of Riemannian manifolds of constant curvature, and then to the general case of Riemannian manifolds with Ricci curvature bounded below and sectional curvature bounded above.

As applications, we present a solution to a stochastic version of the *Lion and Man* problem of Rado, and a proof of the maximum principle for the gradient of harmonic functions.

The material in this chapter is based on the results obtained by the author in [73] (with I. Popescu).

Chapter 4. A maximum modulus principle for non-analytic functions defined in the unit disk

The maximum modulus principle from Complex Analysis asserts that the maximum modulus of a non-constant analytic function defined in a simply connected domain cannot be attained at an interior point of the domain, and it is readily seen to be false if we dispense of the analyticity of the function. In this chapter we show that we can extend the maximum modulus principle to certain classes of non-analytic functions defined in the unit disk. As corollaries we obtain a new proof of the classical maximum modulus principle for analytic functions, simple conditions on the coefficients of the series development under which the maximum modulus principle holds, as well as as applications to the case of real-valued functions of two variables.

The material in this chapter is based on the results obtained by the author in [36] (with M. E. Gageonea, S. Owa, and N. R. Pascu) and [37] (with M. E. Gageonea and N. R. Pascu).

Chapter 5. Univalent approximations of analytic functions

The univalency of an analytic function is an important problem of the Geometric function theory, and there are many sufficient conditions for univalency in the literature (see for example the monographs [34], [74] or [75]). When an analytic function is not univalent, in practical problems it is often of interest to find a "best approximation" of it by a univalent function.

In the present chapter we introduce a measure of the non-univalency of an analytic function, and we use it in order to find the best approximation of a normed analytic function in certain subclasses of univalent functions (starlike, respectively convex functions). We show that the corresponding problems can be reduced to certain semi-infinite quadratic programming problems, which we solve explicitly, thus leading to a method for finding the best starlike, respectively convex approximation. Our results provide constructive algorithms for finding explicitly the measures of the (non)starlikeness, respectively of the (non)convexity of an analytic function, as well as for finding the corresponding best starlike approximation, respectively the best convex approximation, being suitable for numeric implementation and practical applications.

The material in this chapter is based on the results obtained by the author in [71] and [72] (both with N. R. Pascu).

Chapter 6. Neighborhoods of univalent functions

Continuing the study of univalent functions, in this chapter we consider the problem of perturbations of univalent functions. As a measure of the (non)univalency of a function we introduce the constant $K(f, D)$ associated with a function $f : D \subset \mathbb{C} \to \mathbb{C}$ analytic in a domain D, and we use it in order to show that a small perturbation of a univalent function is again a univalent function. As a consequence, we show that a univalent function has a neighborhood consisting entirely of univalent functions.

As applications of the main result, we derive a corollary which is shown to be equivalent to the classical Noshiro-Warschawski-Wolff univalence criterion, and we present an application in terms of Taylor series.

The material in this chapter is based on the results obtained by the author in [70] (with N. R. Pascu).

Chapter 7. Achievements and plans for further career development

The last chapter of the thesis is divided into two sections. The first section contains a presentation of the scientific and professional achievements of the author, and future plans of evolution and development of the professional career. In the second section we present some open problems related to the research presented in the previous chapters, together with some ideas which may lead to a possible resolution of them.

The thesis concludes with a list of bibliographical references.

Braşov, May 2014

Acknowledgments

I express my gratitude to all the people whom, along the years, put their mark and contributed to my personal and professional development: professors, colleagues, collaborators, researchers, students, and lastly, but not last, to all the members of my beloved family.

Since the list of people whom I should be rightfully thanking is much too long and there is always the risk that this list is incomplete or not up-to-date, I will just say

"Thank you ALL!",

thus minimizing the probability of forgetting to thank someone.

Contents

Chapter 1

Scaling coupling of Reflecting Brownian motions

In this chapter we present the construction of the *scaling coupling* of reflecting Brownian motions in smooth planar domains, coupling introduced by the author in [64]. We first give the construction in the case of the unit disk in $\mathbb{R}^2$, and then we extend the construction to general $C^{1,\alpha}$ $(0 < \alpha < 1)$ smooth planar domains by using the conformal invariance of Brownian motion. As applications of this construction, we obtain monotonicity properties of antisymmetric second Neumann eigenfunctions of smooth symmetric convex domains in R^2, which in turn give a resolution of the celebrated "Hot Spots" conjecture of Jeffrey Rauch (presented at a conference at Tulane University in 1974) for a certain class of convex planar domains.

1.1 Introduction

There are mainly known types of coupling of reflecting planar Brownian motions in the literature: *synchronous coupling* and *mirror coupling* (see [25] for a discussion of these couplings). At the core of the construction of these couplings are the translation invariance of the (free) Brownian motion, respectively the invariance of Brownian motion under orthogonal transformations (in particular under reflection in a hyperplane). In the present chapter we use the scaling invariance of free Brownian motion in order to construct the corresponding coupling of reflecting Brownian motions in smooth domains, called the *scaling coupling*.

The structure of this chapter is the following: in Section 1.2 we review some basic facts from complex analysis needed in the paper and we set the notation needed in the sequel.

Next, in Section 1.3 we give the construction of the scaling coupling in the case of unit disk; we show that given a reflecting Brownian motion Z_t in the unit disk U, starting at $z_0 \in \overline{U} - \{0\}$, the formula:

$$\frac{1}{\sup\limits_{s\leq t} |Z_s|} Z_s$$

defines a time change of a reflecting Brownian motion $\widetilde{Z}_t$ in U, starting at $\widetilde{z}_0 = \frac{1}{|z_0|} z_0 \in \partial U$. We define $(Z, \widetilde{Z})$ as a coupling of reflecting Brownian motions starting at z_0 and $\widetilde{z}_0 = \frac{1}{|z_0|} z_0$, respectively. By means of automorphisms of the unit disk, the construction is then extended to any pair of starting points $z_0, \widetilde{z}_0 \in \overline{U}$, not both on the boundary of U. The coupling is uniquely defined by the choice of an additional point lying on the hyperbolic line in U determined by z_0 and $\widetilde{z}_0$, not separating them, which can be viewed as the parameter of the coupling.

In Section 1.4 the construction is extended to smooth $C^{1,\alpha}$ domains $(0 < \alpha < 1)$ by means of conformal maps. The restriction to the class of $C^{1,\alpha}$ domains is necessary in order to insure the conformality at the boundary of U of a mapping from the unit disk onto a $C^{1,\alpha}$ domain, hence to insure that the image of a reflecting Brownian motion in U under a conformal map is a reflecting Brownian motion in the image domain.

Scaling coupling of reflecting Brownian motions in $C^{1,\alpha}$ domains is a coupling in the following generalized sense: there exist a.s. finite stopping times τ and $\widetilde{\tau}$ (with respect to the filtration $(\mathcal{F}_t^Z)$, respectively $(\mathcal{F}_t^{\widetilde{Z}})$) such that

$$Z_{t+\tau} = \widetilde{Z}_{t+\widetilde{\tau}},$$

for all $t \geq 0$. The usual coupling of diffusions can be viewed as a particular case of the above, namely the case when $\tau = \widetilde{\tau}$ a.s. and $\mathcal{F}_t^Z = \mathcal{F}_t^{\widetilde{Z}}$ for all $t \geq 0$.

We show that in the case of convex domains, we can choose the coupling so that we have $\tau \leq \widetilde{\tau}$ a.s., which shows that for $t \geq \widetilde{\tau}$, $\widetilde{Z}_t$ "follows" the path of $Z_{t+\tau-\widetilde{\tau}}$. Moreover, we show that the inequality $\widetilde{\tau} \leq \tau$ is characteristic to the class of convex domains.

The above property of scaling coupling is used to prove strict monotonicity properties of antisymmetric second Neumann eigenfunctions of the Laplacian of a convex $C^{1,\alpha}$ domain D $(0 < \alpha < 1)$ having a line of symmetry, along the family of hyperbolic lines in D which intersect this line of symmetry.

The proof uses the expansion of the function

$$u(t,x) = P(\tau^x > t)$$

in terms of the mixed Dirichlet-Neumann eigenfunctions for the Laplacian on $D^+ = D \cap \{(x,y) : y > 0\}$, with Dirichlet conditions on the part of ∂D^+ lying on the horizontal axis and Neumann conditions on the remaining part of the boundary of D^+ (τ^x denotes the lifetime of reflecting Brownian motion in D^+ starting at $x \in \overline{D^+}$, killed on hitting the horizontal axis).

Using the properties of the scaling coupling, we show a monotonicity property of τ^x (Proposition 1.5.3) along a certain family of curves (hyperbolic lines in D, defined as conformal images of diameters in the unit disk), which gives the monotonicity of $u(t,x) = P(\tau^x > t)$ as a function of x on the indicated family of curves in D^+.

Using the fact that an antisymmetric second Neumann eigenfunction φ for D is a first mixed Dirichlet-Neumann eigenfunction for D^+ (the nodal domain of φ, since the eigenfunction is assumed to be antisymmetric with respect to the horizontal axis), and using the eigenfunction expansion of $u(t,x)$, the monotonicity of $u(t,x)$ translates into the monotonicity of the second Neumann eigenfunction φ for D, along the same family of curves. In particular this shows that an antisymmetric second Neumann eigenfunction of a convex $C^{1,\alpha}$ domain attains its maximum only at the boundary of the domain.

Under our particular setting, this is exactly the object of the hot spots conjecture (due to Jeffrey Rauch, 1974), which states that a second Neumann eigenfunction of a bounded simply connected domain (later modified to convex domains) satisfies a strong maximum/minimum principle, that is, it attains its maximum and minimum over the domain only at the boundary of the domain.

1.2 Preliminaries

We denote the unit disk in $\mathbb{R}^2$ by $U = \{z : |z| < 1\}$.

A curve $\Gamma \subset \mathbb{R}^2$ is said to be of class $C^{1,\alpha}$ $(0 < \alpha < 1)$ if it has a parametrization $w(t)$ that is continuously differentiable, $w' \neq 0$ and w' is Lipschitz of order α, that is, for some $M > 0$ and for all t, t' we have:

$$|w'(t) - w'(t')| \leq M\,|t - t'|^{\alpha}$$

A domain $D \subset \mathbb{R}^2$ is said to be a $C^{1,\alpha}$ domain $(0 < \alpha < 1)$ if its boundary is a Jordan curve Γ of class $C^{1,\alpha}$.

It is known([76], pp. 48–49) that if $f : U \to D$ is a conformal map of the unit disk onto the $C^{1,\alpha}$ domain D $(0 < \alpha < 1)$, then f and f' have continuous extensions to $\overline{U}$, f' is Lipschitz of order α on $\overline{U}$ and $f' \neq 0$ on $\overline{U}$.

We recall that an analytic function f is called convex in U if it maps U conformally onto a convex domain.

For a $C^{1,\alpha}$ domain D $(0<\alpha<1)$ we denote by ν_D the inward unit normal vector field on ∂D, the boundary of D.

By reflecting Brownian motion in D we mean reflecting Brownian motion with respect to the normal vector field ν_D.

Consider an arbitrarily fixed probability space $(\Omega, \mathcal{F}, P)$. Whenever a (reflecting) Brownian motion B_t on $(\Omega, \mathcal{F}, P)$ is considered, we denote by $(\mathcal{F}_t^B)$ the filtration (satisfying the usual conditions) with respect to which B_t is adapted.

We define reflecting Brownian motion in D as a solution of the stochastic differential equation:

$$X_t = X_0 + B_t + \frac{1}{2}\int_0^t \nu_D(X_s)dL_s, \qquad t \geq 0. \tag{1.2.1}$$

Formally we have:

Definition 1.2.1. *X_t is a reflecting Brownian motion in D starting at $x_0 \in \overline{D}$ if it satisfies (2.2.1), where:*

(a) B_t is a 2-dimensional Brownian motion started at 0,

(b) L_t is a continuous nondecreasing process which increases only when $X_t \in \partial D$,

(c) X_t is $(\mathcal{F}_t^B)$-adapted, and almost surely $X_0 = x_0$ and $X_t \in \overline{D}$ for all $t \geq 0$.

Remark 1.2.2. *Bass and Hsu ([15]) showed that in $C^{1,\alpha}$ domains there exists exactly one solution to (2.2.1) for a given Brownian motion B_t. Also, by ([14]) the process X_t corresponds to the Dirichlet form $\mathcal{E}(f,f) = \frac{1}{2}\int_{\overline{D}} |\nabla f|^2$.*

1.3 The case of the unit disk

In this section we give the construction of the scaling coupling in the case of the unit disk. The key of our construction is the following:

Theorem 1.3.1. *Let Z_t be a reflecting Brownian motion in U starting at $z_0 \in \overline{U} - \{0\}$ and let $|z_0| \leq a \leq 1$ be arbitrarily fixed. The process $\tilde{Z}_t$ defined by:*

$$\tilde{Z}_t = \frac{1}{M_{\gamma_t}} Z_{\gamma_t}, \qquad t \geq 0, \tag{1.3.1}$$

where:

$$M_t = a \vee \sup_{s \leq t} |Z_s|,$$

$$C_t = \int_0^t \frac{1}{M_s^2} ds, \tag{1.3.2}$$

$$\gamma_t = \inf\{s > 0 : C_s \geq t\}, \qquad t \geq 0, \tag{1.3.3}$$

is a $(\mathcal{F}_{\gamma_t}^Z)$-adapted reflecting Brownian motion in U starting at $\widetilde{z}_0 = \frac{z_0}{a}$.

Proof. We apply Itô's formula to the semimartingale Z_t and the nondecreasing process M_t with $f(x,y) = \frac{x}{y}$. If

$$Z_t = Z_0 + B_t + \frac{1}{2}\int_0^t \nu_U(Z_s)dL_s$$

is the semimartingale representation of Z_t given by Definition 2.2.1, we have:

$$\begin{aligned} \frac{Z_t}{M_t} &= \frac{Z_0}{M_0} + \int_0^t \frac{1}{M_s} dZ_s - \int_0^t \frac{1}{M_s^2} Z_s dM_s \\ &= \frac{Z_0}{M_0} + \int_0^t \frac{1}{M_s} dB_s + \frac{1}{2}\int_0^t \frac{1}{M_s}\nu_U(Z_s) dL_s - \int_0^t \frac{1}{M_s^2} Z_s dM_s. \end{aligned} \tag{1.3.4}$$

If $\tau = \inf\{s : |Z_s| = 1\}$, note that L_s is constant on $[0, \tau]$ and $M_s \equiv 1$ on $[\tau, \infty)$; further, when M_s is increasing, $\frac{Z_s}{M_s}$ is on ∂U, and therefore $-\frac{Z_s}{M_s} = \nu_U(\frac{Z_s}{M_s})$. The difference of the last two integrals in the last equality above can thus be written:

$$\begin{aligned} &\frac{1}{2}\int_{t\wedge\tau}^t \frac{1}{M_s}\nu_U(Z_s) dL_s - \int_0^{t\wedge\tau} \frac{1}{M_s^2} Z_s dM_s \\ =& \frac{1}{2}\int_{t\wedge\tau}^t \nu_U(\frac{Z_s}{M_s}) dL_s + \frac{1}{2}\int_0^{t\wedge\tau} \nu_U(\frac{Z_s}{M_s}) d\log M_s^2 \\ =& \frac{1}{2}\int_0^t \nu_U(\frac{Z_s}{M_s}) dL_s + \frac{1}{2}\int_0^t \nu_U(\frac{Z_s}{M_s}) d\log M_s^2 \\ =& \frac{1}{2}\int_0^t \nu_U(\frac{Z_s}{M_s}) d\tilde{L}_s, \end{aligned}$$

where $\tilde{L}_s = L_s + \log M_s^2$ is readily seen to be a nondecreasing process which increases only if either L_s or M_s do; this happens only when $\frac{Z_s}{M_s}$ is on the boundary of U.

Note that since $0 < |z_0| \leq M_t \leq 1$ for all $t \geq 0$, C_t defined by (1.3.2) is strictly increasing and $C_t \to \infty$ a.s. It follows that γ_t defined by (1.3.3) is continuous and increasing, and substituting in (1.3.4) we obtain:

$$\tilde{Z}_t =: \frac{Z_{\gamma_t}}{M_{\gamma_t}} = \frac{Z_0}{M_0} + \int_0^{\gamma_t} \frac{1}{M_s} dB_s + \frac{1}{2}\int_0^{\gamma_t} \nu_U(\frac{Z_s}{M_s}) d\tilde{L}_s.$$

If we set $\widetilde{B}_t := \int_0^{\gamma_t} \frac{1}{M_s} dB_s$, then $\langle \widetilde{B}^i, \widetilde{B}^j \rangle_t = \int_0^{\gamma_t} \frac{1}{M_s^2} d\langle \widetilde{B}^i, \widetilde{B}^j \rangle_s = \delta_{ij} \int_0^{\gamma_t} \frac{1}{M_s^2} ds = \delta_{ij} C_{\gamma_t} = \delta_{ij} t$, $i, j \in \{1, 2\}$, and thus $\widetilde{B}_t$ is a 2-dimensional Brownian motion starting at 0.

We have thus shown:

$$\begin{aligned} \tilde{Z}_t &= \tilde{Z}_0 + \widetilde{B}_t + \frac{1}{2}\int_0^{\gamma_t} \nu_D(\frac{Z_s}{M_s}) d\tilde{L}_s \\ &= \tilde{Z}_0 + \widetilde{B}_t + \frac{1}{2}\int_0^t \nu_D(\frac{Z_{\gamma_u}}{M_{\gamma_u}}) d\tilde{L}_{\gamma_u} \\ &= \tilde{Z}_0 + \widetilde{B}_t + \frac{1}{2}\int_0^t \nu_D(\tilde{Z}_u) d\,\overline{L}_u, \end{aligned}$$

where $\overline{L}_u := \tilde{L}_{\gamma_u}$ is a nondecreasing process which increases only when $\tilde{Z}_u := \frac{Z_{\gamma_u}}{M_{\gamma_u}}$ is at the boundary of U. This proves the theorem. $\square$

Remark 1.3.2. *In $\mathbb{R}$, essentially the same proof shows that if Z_t is a 1-dimensional Brownian motion starting at $z_0 > 0$, and $M_t = a \vee \sup_{s \le t} Z_s$ $(a \ge z_0)$, then $\frac{1}{M_t} Z_t$ is a time change of a reflecting Brownian motion on $(-\infty, 1]$, the time change being given by (1.3.2)–(1.3.3) above.*

The above construction also applies to higher dimensions, to give a scaling coupling of reflecting Brownian motions in the unit sphere in $\mathbb{R}^n$, $n \ge 3$. However, since the conformal images of the unit sphere in $\mathbb{R}^n$ $(n \ge 3)$ are again spheres, we cannot use the conformal mapping arguments presented in the following section in order to extend the construction to more general domains.

Definition 1.3.3. *We call the pair $(Z_t, \tilde{Z}_t)$ constructed above a scaling coupling of reflecting Brownian motions in U, starting at $z_0 \in \overline{U} - \{0\}$, respectively at $\tilde{z}_0 := \frac{1}{a} z_0 \in \overline{U}$.*

1.4 The case of smooth domains

In the present section we will extend the construction of the scaling coupling from the case of the unit disk to the case of smooth $C^{1,\alpha}$ domains $(0 < \alpha < 1)$. We will need the following proposition, showing that the conformal image of a reflecting Brownian motion in the unit disk is a time change of a reflecting Brownian motion in the image domain. More precisely, we have:

Proposition 1.4.1. *Let Z_t be a reflecting Brownian motion in U starting at $z_0 \in \overline{U}$. If f is a conformal map of U onto the $C^{1,\alpha}$ domain D $(0 < \alpha < 1)$, then $W_t = f(Z_{\alpha_t})$ is a $(\mathcal{F}^Z_{\alpha_t})$-adapted reflecting Brownian motion in D starting at $f(z_0)$, where:*

$$\alpha_t = \inf\{s : A_s \ge t\}$$

and

$$A_t = \int_0^t |f'(Z_s)|^2 \, ds.$$

Proof. Recall that since D is a $C^{1,\alpha}$ domain, $f \in C^1(\overline{D})$. If

$$Z_t = Z_0 + B_t + \frac{1}{2} \int_0^t \nu_U(Z_s) dL_s$$

is the semimartingale representation of Z_t given by Definition 2.2.1, by applying Itô's formula with $f = (u, v)$, we have:

$$f(Z_t) = f(Z_0) + \int_0^t (\nabla u, \nabla v)(Z_s) dB_s + \frac{1}{2} \int_0^t (\nabla u \cdot \nu_U, \nabla v \cdot \nu_U)(Z_s) dL_s.$$

By Levy's theorem ([11]), the stochastic integral above is a time change of a 2-dimensional Brownian motion, the increasing process being $A_t = \int_0^t |f'(Z_s)|^2 \, ds$. Replacing t by $\alpha_t = \inf\{s : A_s \ge t\}$, the stochastic integral above becomes a $(\mathcal{F}^Z_{\alpha_t})$-adapted Brownian motion (note that by Definition 2.2.1 we have $\mathcal{F}^Z_t = \mathcal{F}^B_t$ for all $t \ge 0$).

So it suffices to show that when L_s increases, $(\nabla u \cdot \nu_U, \nabla v \cdot \nu_U)(Z_s)$ has no tangential component to the boundary of D.

However this follows since by preliminary remarks, f has a conformal extension to $\overline{U}$; one can use Cauchy-Riemann equations and the geometric interpretation of the argument of f' to show that whenever L_s increases we have:

$$(\nabla u \cdot \nu_U, \nabla v \cdot \nu_U)(Z_s) = |f'(Z_s)| \, \nu_D(f(Z_s)),$$

which concludes the proof. □

Before carrying out the general construction of scaling coupling, we introduce the notion of hyperbolic line in $C^{1,\alpha}$ domains ($0<\alpha<1$), as follows:

Definition 1.4.2. *i) We define a hyperbolic line in U as being a line segment or an arc of a circle contained in $\overline{U}$ which meets orthogonally the boundary of U. We denote by $\mathcal{H}_U$ the family of all hyperbolic lines in U. If z_1, z_2 are two distinct points on a hyperbolic line $l \in \mathcal{H}_U$, we define the hyperbolic segment with endpoints z_1 and z_2 (denoted by $[z_1, z_2]$) as the part of l between (and including) z_1 and z_2.*

ii) For a $C^{1,\alpha}$ ($0<\alpha<1$) domain D, we define a hyperbolic line/segment in D as the conformal image of a hyperbolic line/segment in U. We denote by $\mathcal{H}_D$ the family of all hyperbolic lines in D.

Simple geometric considerations show the following:

Proposition 1.4.3. *Let D be a $C^{1,\alpha}$ ($0<\alpha<1$) domain.*

i) Given two distinct points $\overline{D}$, there exists a unique hyperbolic line in D passing through them.

ii) For an arbitrarily chosen diameter d of U, we have

$$\mathcal{H}_U = \{\varphi(d) : \varphi \in \mathcal{A}\},$$

where $\mathcal{A}$ is the family of all automorphisms of U. If f is an arbitrarily chosen conformal map of U onto D, then

$$\mathcal{H}_D = \{f \circ \varphi(d) : \varphi \in \mathcal{A}\}.$$

Remark 1.4.4. *Part i) of the previous proposition shows that given any two distinct points z_1 and z_2 in a $C^{1,\alpha}$ domain D ($0<\alpha<1$), the hyperbolic segment $[z_1, z_2]$ in D is uniquely determined, hence the notion of hyperbolic segment is well defined in the above definition. We will denote by z_1z_2 the unique hyperbolic line passing through z_1 and z_2.*

We give now the construction of scaling coupling for general $C^{1,\alpha}$ domains. Let D be a $C^{1,\alpha}$ domain ($0<\alpha<1$) and let $w_0, \widetilde{w}_0 \in \overline{D}$ distinct, not both on ∂D, be arbitrarily fixed. By Proposition 1.4.3, there is a unique hyperbolic line $w_0\widetilde{w}_0$ in D, passing through w_0 and $\widetilde{w}_0$. Consider a point w_1 on $w_0\widetilde{w}_0 - [w_0, \widetilde{w}_0]$, $w_1 \notin \partial D$. Let $f : U \to D$ be the unique conformal map of U onto D (given by the Riemann mapping theorem) with $f(0) = w_1$ and $\arg f'(0) = 0$. Let $z_0 = f^{-1}(w_0)$ and $\widetilde{z}_0 = f^{-1}(\widetilde{w}_0)$. Note that by definition, $f^{-1}(w_0\widetilde{w}_0)$ is a hyperbolic line in U, and since $0 = f^{-1}(w_1) \in f^{-1}(w_0\widetilde{w}_0)$, it follows that $f^{-1}(w_0\widetilde{w}_0)$ is in fact a diameter of U. Note that by the choice of w_1, we have $0 \notin [z_0, \widetilde{z}_0]$, and therefore $|z_0| \neq |\widetilde{z}_0|$. Without loss of generality we can assume that $|z_0| < |\widetilde{z}_0|$.

Let Z_t be a reflecting Brownian motion in U starting at z_0. Define processes W_t, $\widetilde{W}_t$ by:

$$W_t = f(Z_{\alpha_t}), \qquad t \geq 0, \tag{1.4.1}$$

$$\widetilde{W}_t = f\left(\frac{1}{M_{\beta_t}} Z_{\beta_t}\right), \qquad t \geq 0, \tag{1.4.2}$$

where $M_t = \left|\frac{z_0}{\widetilde{z}_0}\right| \vee \sup\limits_{s \leq t} |Z_s|$, $t \geq 0$ and:

$$A_t = \int_0^t |f'(Z_s)|^2 \, ds, \qquad \alpha_t = \inf\{s : A_s \geq t\}, \qquad t \geq 0, \tag{1.4.3}$$

$$B_t = \int_0^t \frac{1}{M_s^2} \left| f'\left(\frac{Z_s}{M_s}\right) \right|^2 ds, \qquad \beta_t = \inf\{s : B_s \geq t\}, \qquad t \geq 0. \tag{1.4.4}$$

Theorem 1.4.5. *$W_t, \widetilde{W}_t$ defined by (1.4.1)–(1.4.4) are $(\mathcal{F}^Z_{\alpha_t})$, respectively $(\mathcal{F}^Z_{\beta_t})$-adapted reflecting Brownian motions in D, starting at w_0, respectively $\widetilde{w}_0$.*

Proof. That W_t is a $(\mathcal{F}^Z_{\alpha_t})$-adapted reflecting Brownian motion in D follows from Proposition 1.4.1.

To prove that $\widetilde{W}_t$ is a $(\mathcal{F}^Z_{\beta_t})$-adapted reflecting Brownian motion, note that by Lemma 1.3.1, $\frac{Z_t}{M_t}$ is a time change γ_t (given by (1.3.3)) of a reflecting Brownian motion in $\overline{U}$, starting at $\widetilde{z}_0$.

By Proposition 1.4.1, $f(\frac{Z_{\gamma_t}}{M_{\gamma_t}})$ is a time change $\tilde{\alpha}_t$ of a reflecting Brownian motion in D, starting at $f(\widetilde{z}_0) = \widetilde{w}_0$, where

$$\tilde{\alpha}_t = \inf\{s : \tilde{A}_s \geq t\} \text{ and } \tilde{A}_t = \int_0^t \left| f'(\frac{Z_{\gamma_s}}{M_{\gamma_s}}) \right|^2 ds, \qquad t \geq 0. \tag{1.4.5}$$

In order to prove the claim it suffices to show that the combined effect of the two time changes γ_t and $\tilde{\alpha}_t$ is the time change β_t given by (1.4.4), that is $\gamma_{\tilde{\alpha}_t} = \beta_t$ for all $t \geq 0$.

C_u given by (1.3.3) is a bijection on $[0, \infty)$, with inverse $C^{-1} = \gamma$. With the substitution $s = C_u$ in the definition of $\tilde{A}_t$, we obtain:

$$\begin{aligned} \tilde{A}_{C_t} &= \int_0^{C_t} \left| f'(\frac{Z_{\gamma_s}}{M_{\gamma_s}}) \right|^2 ds = \\ &= \int_0^t \left| f'(\frac{Z_u}{M_u}) \right|^2 \frac{dC_u}{du} du = \\ &= \int_0^t \frac{1}{M_u^2} \left| f'(\frac{Z_u}{M_u}) \right|^2 du = \\ &= B_t, \end{aligned}$$

for all $t \geq 0$. This shows that $\tilde{A}_{C_t} = B_t$ for all $t \geq 0$, or equivalently, by taking inverses, we have $\gamma_{\tilde{\alpha}_t} = \beta_t$ for all $t \geq 0$, as needed. □

Definition 1.4.6. *For a $C^{1,\alpha}$ domain D $(0 < \alpha < 1)$ and arbitrarily fixed distinct points $w_0, \widetilde{w}_0 \in \overline{D}$ (not both on ∂D), $w_1 \in w_0\widetilde{w}_0 - [w_0, \widetilde{w}_0]$ (not on ∂D), the pair $(W_t, \widetilde{W}_t)$ defined by (1.4.1) -(1.4.4) is called a scaling coupling of reflecting Brownian motions in D starting at $w_0 \in \overline{D}$ and $\widetilde{w}_0 \in \overline{D}$, respectively.*

Remark 1.4.7. *The above construction of scaling coupling for a $C^{1,\alpha}$ domain with starting points $(w_0, \widetilde{w}_0)$ relied on the choice of a conformal map from the unit disk U onto D. As it is known, the choice is uniquely determined by the values of $f(0)$ and $\arg f'(0)$. However, the choice of just $w_1 = f(0)$ determines uniquely the conformal map, up to a rotation of the unit disk. By the angular symmetry of the construction of scaling coupling in the case of the unit disk, it follows that the construction is invariant under rotations of the unit disk, and therefore the construction does not depend of the choice of $\arg f'(0)$ (we chose $\arg f'(0) = 0$ for simplicity).*

It follows that given two distinct point in $\overline{D}$ (not both on the boundary of D), the scaling coupling with these starting points is uniquely determined once a choice for $f(0)$ (lying on the hyperbolic line passing through them, and not separating them) has been made. We will therefore refer to $w_1 = f(0)$ as the parameter of the scaling coupling.

In order to derive the main property of the scaling coupling in the case of convex domains, we need the following characterization of convexity:

Proposition 1.4.8. *Let $f : U \to D$ be a conformal map of U onto the simply connected domain D. The following are equivalent:*

$$D \textit{ is convex;} \tag{1.4.6}$$

$$\left| r f'(re^{i\theta}) \right| \textit{ is an increasing function of } r \in [0, 1), \textit{ for all } 0 \leq \theta < 2\pi. \tag{1.4.7}$$

Proof. Since $f'(0) \neq 0$, without loss of generality we can assume that $f(0) = f'(0) - 1 = 0$.

Note that the domain D is convex iff the function f is convex, which (under the condition $f(0) = f'(0) - 1 = 0$) is equivalent (see [34], pp.42) to:

$$\mathrm{Re}\left(1 + z\frac{f''(z)}{f'(z)}\right) > 0, \quad z \in U.$$

In polar coordinates, $z = re^{i\theta}$, we obtain equivalent that

$$\mathrm{Re}\left(\frac{1}{r} + e^{i\theta}\frac{f''(re^{i\theta})}{f'(re^{i\theta})}\right) > 0, \quad 0 \leq \theta < 2\pi, \quad 0 < r < 1.$$

Note that since

$$\begin{aligned}\mathrm{Re}\left(\frac{1}{r} + e^{i\theta}\frac{f''(re^{i\theta})}{f'(re^{i\theta})}\right) &= \frac{1}{r} + \mathrm{Re}\left(\frac{\partial}{\partial r}\log f'(re^{i\theta})\right)\\ &= \frac{1}{r} + \frac{\partial}{\partial r}\ln\left|f'(re^{i\theta})\right|\\ &= \frac{\partial}{\partial r}\ln\left|rf'(re^{i\theta})\right|,\end{aligned}$$

the previous statement is equivalent to (1.4.7), as needed. □

The main feature of the scaling coupling is given by the following:

Proposition 1.4.9. *With the notation of Theorem 1.4.5, there exist almost surely finite stopping times τ_1 and τ_2 with respect to the filtration $(\mathcal{F}^Z_{\alpha_t})$, respectively $(\mathcal{F}^Z_{\beta_t})$, such that for all $t \geq 0$ we have:*

$$\alpha_{t+\tau_1} = \beta_{t+\tau_2}$$

Moreover, if the domain D is convex, with probability one we have $\beta_t \leq \alpha_t$, for all $t \geq 0$.

Proof. Set $\tau = \inf\{s : |Z_s| = 1\}$, $\tau_1 = A_\tau$ and $\tau_2 = B_\tau$. Obviously τ is an a.s. finite stopping time with respect to the filtration $(\mathcal{F}^Z_t)$, and we have $M_s \equiv 1$ for all $s \geq \tau$.

It follows that τ_1 and τ_2 are also a.s. finite stopping times with respect to the filtration $(\mathcal{F}^Z_{\alpha_t})$, respectively $(\mathcal{F}^Z_{\beta_t})$, and that for all $t \geq \tau$ we have:

$$A_t - \tau_1 = \int_\tau^t \left|f'(Z_s)\right|^2 ds = B_t - \tau_2.$$

Since $\alpha_t = A_t^{-1}$, $\beta_t = B_t^{-1}$, this implies the first part of the proposition.

For the second part, note that since D is convex, Proposition 1.4.8 shows that:

$$|Z_s f'(Z_s)| \leq \left|\frac{Z_s}{M_s} f'(\frac{Z_s}{M_s})\right|,$$

hence we obtain:

$$A_t = \int_0^t \left|f'(Z_s)\right|^2 ds \leq \int_0^t \frac{1}{M_s^2}\left|f'(\frac{Z_s}{M_s})\right|^2 ds = B_t,$$

and therefore we have $\alpha_t \geq \beta_t$, for all $t \geq 0$. □

Remark 1.4.10. *The pair $(W_t, \widetilde{W}_t)$ in Definition 1.4.6 is a coupling in the following extended sense: there exist a.s. finite stopping times τ_1 and τ_2 with respect to the filtration $(\mathcal{F}^Z_{\alpha_t})$, respectively $(\mathcal{F}^Z_{\beta_t})$, such that:*

$$W_{t+\tau_1} = \widetilde{W}_{t+\tau_2}, \qquad \textit{for all } t \geq 0. \tag{1.4.8}$$

The usual coupling of diffusions can be thought as a particular case of the above, namely the case when $\tau_1 = \tau_2$ a.s. (and the two filtrations coincide).

The scaling coupling is readily seen to satisfy (1.4.8) by using

$$\tau_1 = \int_0^\tau |f'(Z_s)|^2 \, ds$$

$$\tau_2 = \int_0^\tau \frac{1}{M_s^2} \left| f'(\frac{Z_s}{M_s}) \right|^2 ds,$$

where $\tau = \inf\{s > 0 : |Z_s| = 1\}$.

Moreover, in the particular case of convex domains, we have $\tau_1 \leq \tau_2$ a.s., which shows that for $t \geq \tau_2$, $\widetilde{W}_t$ "follows" the path of $W_{t+\tau_1-\tau_2}$.

Remark 1.4.11. *Note that by the equivalence in Proposition 1.4.8 and using the support theorem for Brownian motion, it follows that the class of domains for which the inequality $\beta_t \leq \alpha_t$ holds almost surely for all $t \geq 0$ (and all starting points w_0, $\widetilde{w}_0$), coincides with the class of convex $C^{1,\alpha}$ domains.*

1.5 Hot Spots Conjecture

Results on eigenvalues and eigenfunctions

We will review first some basic facts about eigenfunctions and eigenvalues. We make the remark that for the convenience of arguments involving Brownian motion, we will be using $\frac{1}{2}\Delta$ instead of the Laplace operator Δ. The results hold for the Laplacian Δ by scaling.

Fix an arbitrarily $C^{1,\alpha}$ domain D $(0 < \alpha < 1)$.

We say that λ is an eigenfunction for $\frac{1}{2}\Delta$ in D, if there exists a nontrivial solution $\varphi \in C^2(D) \cap C^1(\overline{D})$ to

$$\frac{1}{2}\Delta\varphi + \lambda\varphi = 0. \tag{1.5.1}$$

φ is then called an eigenfunction corresponding to the eigenvalue λ.

If $\frac{\partial\varphi}{\partial\nu_D} = 0$ on ∂D, the we will refer to λ and φ as being a *Neumann eigenvalue* and *Neumann eigenfunction*, respectively.

If $\frac{\partial\varphi}{\partial\nu_D} = 0$ only on a nonempty proper open subset of ∂D, and $\varphi = 0$ on the remaining part of ∂D, we refer to λ and φ as being a *mixed Dirichlet-Neumann eigenvalue* and *eigenfunction*, respectively.

It is known (see [28], pp. 46) that the set of Neumann/mixed Dirichlet-Neumann eigenvalues forms an unbounded sequence

$$0 \leq \lambda_1 < \lambda_2 < \lambda_3 < ... \nearrow \infty.$$

We will refer to λ_i as the i^{th} Neumann/mixed Dirichlet-Neumann eigenvalue for D and to the eigenfunctions belonging to the eigenspace corresponding to λ_i as the i^{th} Neumann/mixed Dirichlet-Neumann eigenfunctions for D.

Recall that the nodal set of an eigenfunction φ is the set $\varphi^{-1}(\{0\}) = \{x \in \overline{D} : \varphi(x) = 0\}$, and a nodal domain of φ is a component of $\overline{D} - \varphi^{-1}(\{0\})$.

The Courant Nodal Domain theorem ([28], pp. 19) asserts that for each $k \geq 1$, the number of nodal domains of a k^{th} eigenfunction of a simply connected domain is less than or equal to k.

As an immediate consequence we have that a first eigenfunction has constant sign, and a second eigenfunction has precisely 2 nodal domains. Moreover, λ_1 is characterized as being the only eigenvalue with eigenfunction of constant sign ([28], pp. 20).

It is also known, that in the case of a second Neumann eigenfunction the nodal set is a smooth (C^∞) curve, called the nodal line, and that there are no closed nodal lines ([6], pp.128).

The hot spots conjecture (due to Jeffrey Rauch, 1974) is a strong maximum principle for second Neumann eigenfunctions of a simply connected domain D, and it can be formulated as follows:

Conjecture 1.5.1. *For every second Neumann eigenfunction φ_2 of D, and for all $y \in D$, we have:*

$$\min_{x\in\partial D} \varphi_2(x) < \varphi_2(y) < \max_{x\in\partial D} \varphi_2(x). \tag{1.5.2}$$

By an abuse of language, if (1.5.2) holds for a second Neumann eigenfunction φ_2 of D, we say that the hot spots conjecture holds for φ_2.

According to Kawohl ([49]), Conjecture 1.5.1 holds for balls, annuli and rectangles in $\mathbb{R}^d$. Burdzy and Werner constructed a counterexample to Conjecture 1.5.1, in which the maximum of φ_2 is attained only in the interior of D and the minimum at the boundary of D. More recently, Bass and Burdzy constructed a stronger counterexample, in which both the maximum and the minimum of φ_2 are attained only in the interior of D. In both counterexamples the domain D was not convex.

Recent advances in the hot spots problem identified classes of domains for which Conjecture 1.5.1 holds. It is know that Conjecture 1.5.1 holds for bounded convex domains with two orthogonal axes of symmetry([7], [48]), or just one axis of symmetry and additional hypothesis on the domain ([7]). The question whether Conjecture 1.5.1 holds for bounded convex domains in $\mathbb{R}^d$ is still open.

1.5.1 Main results

The main result in this section is Theorem 1.5.6, which shows that the hot spots conjecture holds for antisymmetric second Neumann eigenfunctions of smooth convex domains having a line of symmetry.

Let $D \subset \mathbb{R}^2$ be a convex $C^{1,\alpha}$ domain $(0 < \alpha < 1)$ having a line of symmetry. Without loss of generality we will assume that D is symmetric with respect to the horizontal axis.

Set $D^+ = D \cap \{(x, y) : y > 0\}$, $\Gamma^+ = \partial D^+ \cap \partial D$, $\Gamma_0 = \overline{D} \cap \{(x, y) : y = 0\}$. For $w_0 \in \overline{D^+}$, denote by $\tau^{w_0} = \inf\{s : \operatorname{Im} W_s = 0\}$, where W_s is a reflecting Brownian motion in D starting at w_0.

Remark 1.5.2. *Let f_1 be a conformal map of U^+ onto D^+, such that the parts of the boundaries of U^+ and D^+ lying on the horizontal axis correspond to each other. By the symmetry principle, f_1 extends to a conformal map of U onto D. Since $\Gamma_0 = f_1([-1, 1])$, it follows that Γ_0 is a hyperbolic line in $\overline{D}$.*

The key to the Theorem 1.5.4 is the following lemma, showing a monotonicity property of τ^{w_0} (as a function of w_0) on the family of hyperbolic lines in D intersecting the horizontal axis:

Lemma 1.5.3. *Given $\widetilde{w}_0 \in \overline{D}$, $w_1 \in D \cap \Gamma_0$ and $w_0 \in [w_1, \widetilde{w}_0]$, there exist filtrations $(\mathcal{F}_t)$, $(\widetilde{\mathcal{F}}_t)$ and $(\mathcal{F}_t)$, respectively $(\widetilde{\mathcal{F}}_t)$-adapted reflecting Brownian motions W_t, $\widetilde{W}_t$ in D, starting at w_0, respectively $\widetilde{w}_0$, such that if τ^{w_0}, $\tau^{\widetilde{w}_0}$ are the hitting times to the horizontal axis of W_t, respectively $\widetilde{W}_t$, then with probability one we have:*

$$\tau^{w_0} \leq \tau^{\widetilde{w}_0}.$$

Proof. The proof is trivial if $w_0 = \widetilde{w}_0$ or $w_0 = w_1$, so we can assume that $w_0, \widetilde{w}_0$ and w_1 are distinct.

Let $(W_t, \widetilde{W}_t)$ be a scaling coupling of reflecting Brownian motions in D starting at $(w_0, \widetilde{w}_0)$, with parameter w_1, that is a scaling coupling obtained as the image under a conformal map $f : U \to D$ with $f(0) = w_1$ of the scaling coupling $(Z_t, \tilde{Z}_t)$ in U. The filtrations $(\mathcal{F}_t)$, $(\widetilde{\mathcal{F}}_t)$ are the corresponding "time changes" of the filtrations $(\mathcal{F}^Z_t)$, respectively $(\mathcal{F}^{\widetilde{Z}}_t)$, as indicated in Proposition 1.4.1.

Remark 1.5.2 above shows that Γ_0 is a hyperbolic line in U, hence $f^{-1}(\Gamma_0)$ is a hyperbolic line in U. Since $0 = f^{-1}(w_1) \in f^{-1}(\Gamma_0)$, it follows that $f^{-1}(\Gamma_0)$ is in fact a diameter of U. By Remark 1.4.7, without loss of generality we can assume that $f^{-1}(\Gamma_0) = [-1, 1]$.

If $s > 0$ is such that $\operatorname{Im} \widetilde{W}_s = 0$, by construction of the coupling we have $\operatorname{Im} f(\frac{Z_{\beta_s}}{M_{\beta_s}}) = 0$. Because under f the parts of the boundaries of U^+ and D^+ lying on the horizontal axis correspond to each other, $\operatorname{Im} \frac{Z_{\beta_s}}{M_{\beta_s}} = 0$. Since M_{β_s} is real, it follows that $\operatorname{Im} Z_{\beta_s} = 0$.

Since D is convex, by Proposition 1.4.9 it follows that $\beta_s \le \alpha_s$. Since α_s is increasing (and a bijection) on $[0, \infty)$, there exists $s' \le s$ such that $\alpha_{s'} = \beta_s$.

It follows that $\operatorname{Im} Z_{\alpha_{s'}} = \operatorname{Im} Z_{\beta_s} = 0$, hence $W_{s'} = f(Z_{\alpha_{s'}})$ is on the horizontal axis.

We have shown that if $\operatorname{Im} \widetilde{W}_s = 0$, then there exists $s' \le s$ such that $\operatorname{Im} W_{s'} = 0$, which implies that $\tau^{w_0} \le \tau^{\widetilde{w}_0}$ a.s., as needed. □

We can now prove a first version of our main result:

Theorem 1.5.4. *Let D be a convex $C^{1,\alpha}$ domain ($0 < \alpha < 1$) which is symmetric with respect to the horizontal axis. If φ is a second Neumann eigenfunction for D which is antisymmetric with respect to the horizontal axis, then φ is monotone on the family of hyperbolic lines in D which intersect the horizontal axis.*

In particular, φ must attain its maximum and minimum over $\overline{D}$ on the boundary of D.

Proof. By the assumption, φ must be identical zero on the horizontal axis, and therefore the nodal line for φ is Γ_0 (the part of the horizontal axis contained in $\overline{D}$). It follows that D^+ and D^- are the nodal domains of φ.

Since φ has constant sign on each nodal domain, without loss of generality we will assume that φ is positive on D^+. Since φ is antisymmetric with respect to the horizontal axis, it suffices to prove the monotonicity of φ in D^+ along the indicated family of curves.

Consider an arbitrarily fixed hyperbolic line in D, which intersects the horizontal axis, and denote by w_1 the point of intersection. If $w_0, \widetilde{w}_0 \in \overline{D^+}$ are arbitarily chosen points lying on this hyperbolic line, such that $w_0 \in [w_1, \widetilde{w}_0]$, we will show that $\varphi(w_0) \le \varphi(\widetilde{w}_0)$.

Since the restriction of a second Neumann eigenfunction for D to one of its nodal domains has constant sign, by preliminary remarks it follows that it is a first mixed Neuman-Dirichlet eigenfunction for the corresponding nodal domain. Therefore, the restriction of φ to $\overline{D^+}$ is a first mixed Dirichlet-Neumann eigenfunction for D^+, with Neumann conditions on Γ^+ and Dirichlet conditions on Γ_0.

It can be shown ([63], pp. 20) that the transition density $p_{D^+}(t, x, y)$ of reflecting Brownian motion in D^+, killed on hitting the horizontal axis, has an eigenfunction expansion in terms of the mixed Dirichlet-Neumann eigenfunctions for D^+, with Dirichlet boundary conditions on Γ_0 and Neumann conditions on Γ^+. More precisely, it can be shown that:

$$p_{D^+}(t, x, y) = \sum_{i \ge 1} e^{-\mu_i t} \varphi_i(x) \varphi_i(y),$$

where $0 < \mu_1 < \mu_2 \le \ldots$ are the mixed Dirichlet-Neumann eigenvalues for D^+ repeated according to the multiplicity, and $\{\varphi_i\}_{i \ge 1}$ is an orthonormal sequence of eigenfunctions corresponding to the eigenvalues $\{\mu_i\}_{i \ge 1}$. Moreover, the convergence is uniform and absolute on $\overline{D^+}$.

Note that since μ_1 is simple, the corresponding eigenspace is 1-dimensional, and therefore $\varphi_1 = c\varphi$, for some nonzero constant c. Also note that since $\mu_1 < \mu_i$, for all $i \ge 2$, we can write:

$$\begin{aligned} p_{D^+}(t, x, y) &= e^{-\mu_1 t} \varphi_1(x) \varphi_1(y) + R(t, x, y) \\ &= c^2 e^{-\mu_1 t} \varphi(x) \varphi(y) + R(t, x, y), \end{aligned} \tag{1.5.3}$$

where $\lim_{t \to \infty} e^{\mu_1 t} R(t, x, y) = 0$, uniformly in $x, y \in \overline{D^+}$.

Consider the function $u : (0, \infty) \times D \to \mathbb{R}$ given by $u(t, x) = E[1; \tau^x > t]$, where τ^x is the lifetime of reflecting Brownian motion in D^+ starting at x, killed on hitting Γ_0. Integrating the eigenfunction expansion (1.5.3), we obtain:

$$\begin{aligned} u(t,x) &= \int_{D^+} p_{D^+}(t,x,y)dy \qquad (1.5.4) \\ &= c^2 e^{-\mu_1 t}\varphi(x)\int_{D^+}\varphi(y)dy + \int_{D^+} R(t,x,y)dy = \\ &= ae^{-\mu_1 t}\varphi(x) + R_1(t,x), \end{aligned}$$

where, by assumption, $a = c^2 \int_{D^+} \varphi(y)dy > 0$ and $R_1(t,x)$ approaches zero faster than $e^{-\mu_1 t}$ as $t \to \infty$.

Since $u(t,x) = P(\tau^x > t)$, and using the monotonicity property in Lemma 1.5.3, it follows that for any $t \geq 0$ we have:

$$\begin{aligned} u(t,w_0) &= P(\tau^{w_0} > t) \\ &\leq P(\tau^{\widetilde{w}_0} > t) \\ &= u(t,\widetilde{w}_0). \end{aligned}$$

Using the analytic representation (1.5.4) of $u(t,\cdot)$ we obtain therefore:

$$a\varphi(w_0) + e^{\mu_1 t}R_1(t,w_0) \leq a\varphi(\widetilde{w}_0) + e^{\mu_1 t}R_1(t,\widetilde{w}_0),$$

for all $t \geq 0$. Letting $t \to \infty$, it follows $a\varphi(w_0) \leq a\varphi(\widetilde{w}_0)$, and since $a > 0$, we obtain $\varphi(w_0) \leq \varphi(\widetilde{w}_0)$, as needed. □

The above theorem leaves open the question whether φ can also attain its maximum/minimum over $\overline{D}$ inside the domain. We will show that under the hypotheses of the previous theorem this cannot happen; more generally, we will show that φ is strictly monotone on the family of hyperbolic lines which intersect nontrivially the axis of symmetry of D. We have:

Theorem 1.5.5. *Under the hypotheses of the previous theorem, for any $a \neq 0$, the intersection between a level set $\varphi^{-1}(\{a\}) = \{x \in \overline{D} : \varphi(x) = a\}$ of φ and a hyperbolic line of D which intersects the horizontal axis, consists of at most one point.*

Proof. Assume there exists distinct points w_1, w_2 such that $\varphi(w_1) = \varphi(w_2) \neq 0$ and the hyperbolic line w_1w_2 intersects the horizontal axis, say at w_0. Without loss of generality we can assume $\varphi(w_1) = \varphi(w_2) > 0$, and therefore $w_1, w_2 \in \overline{D^+}$. Also, we may assume that $w_1 \in [w_0, w_2]$.

Consider the points $z_1, z_2 \in \Gamma_0$ such that $z_1 < w_0 < z_2$.

We will show that $[w_1, z_3] \cap [z_2, w_2] \neq \emptyset$, where z_3 is the endpoint of the hyperbolic line z_1w_1 lying on $\partial D^+ - \Gamma_0$.

Note that by Proposition 1.4.3, the intersection of any two hyperbolic lines consists of at most one point.

Since $[w_1, z_3] \cap w_1w_2 = \{w_1\}$, it follows that the hyperbolic segment $[w_1, z_3]$ is contained in the right (hyperbolic) half plane determined by the hyperbolic line w_1w_2 (see Figure 1.1). If $[w_1, z_3] \cap [z_2, w_2] = \emptyset$, then $[w_1, z_3]$ is also contained in the left (hyperbolic) half plane determined by the hyperbolic line z_2w_2. It follows that $z_3 \in [w_1, z_3]$ must be contained in their intersection; however, since $z_3 \in \overline{D^+} - \Gamma_0$, this is possible only if $z_3 = w_2$, contradicting $[w_1, z_3] \cap [z_2, w_2] = \emptyset$.

Therefore we must have $[w_1, z_3] \cap [z_2, w_2] \neq \emptyset$, and we denote by w_3 the point of intersection.

By applying the previous theorem to the hyperbolic lines w_1w_2, w_1w_3 and respectively w_3w_2, we obtain:

$$\begin{aligned} \varphi(w_1) &\leq \varphi(z) \leq \varphi(w_2) \\ \varphi(w_1) &\leq \varphi(z') \leq \varphi(w_3) \\ \varphi(w_3) &\leq \varphi(z'') \leq \varphi(w_2), \end{aligned}$$

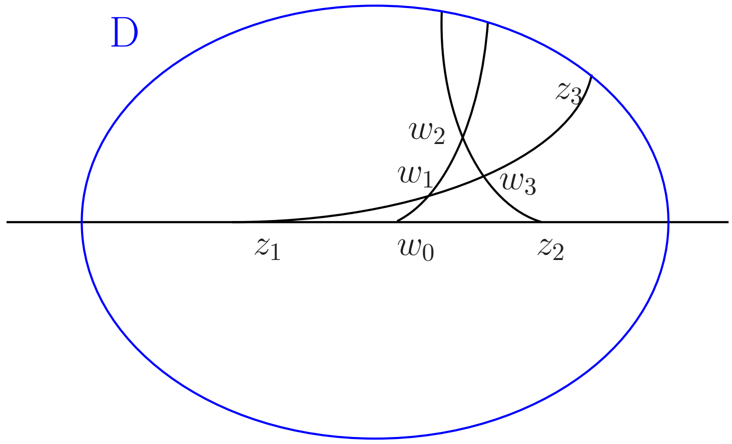

Figure 1.1: Hyperbolic lines in D.

for all $z \in [w_1, w_2]$, $z' \in [w_1, w_3]$, $z'' \in [w_3, w_2]$. Since by hypothesis $\varphi(w_1) = \varphi(w_2)$, we obtain that $\varphi(z) = \varphi(w_1) = \varphi(w_2) = \varphi(w_3)$, for all $z \in .[w_1, w_2] \cup [w_1, w_3] \cup [w_3, w_2]$.

Choosing now an arbitrarily fixed $z_2 \in [w_1, w_3]$ and applying again the previous theorem to the hyperbolic line zw_2 (which intersects the horizontal axis between w_0 and z_2), we obtain

$$\varphi(w_2) = \varphi(z) \leq \varphi(z') \leq \varphi(w_2),$$

for all $z' \in [z, w_2]$. Therefore we have that $\varphi(z') = \varphi(w_2)$, for all $z' \in [z, w_2]$ where $z \in [w_1, w_3]$.

It follows that φ is constant on the interior of the (hyperbolic) triangle with vertices w_1, w_2 and w_3. Since the points w_1, w_2, w_3 do not lie on the same hyperbolic line, the interior of this triangle is not empty.

However, φ, being a nonconstant real analytic function, cannot be constant on a nonempty open set. The contradiction shows that φ is injective on every hyperbolic line (of course, except Γ_0) intersecting the horizontal axis, thus proving the claim. □

Using the above theorem, we can strengthen the result in Theorem 1.5.4, and we state:

Theorem 1.5.6. *Let D be a convex $C^{1,\alpha}$ domain ($0 < \alpha < 1$) which is symmetric with respect to the horizontal axis. If φ is a second Neumann eigenfunction for D which is antisymmetric with respect to the horizontal axis, then φ is **strictly** monotone on the family of hyperbolic lines in D which intersect nontrivially the horizontal axis.*

In particular, φ must attain its maximum and minimum over $\overline{D}$ solely at the boundary of D.

As an immediate consequence, we have:

Corollary 1.5.7. *Let D be a convex $C^{1,\alpha}$ domain ($0 < \alpha < 1$) which is symmetric with respect to the horizontal axis. If φ is a second Neumann eigenfunction for D which is antisymmetric with respect to the horizontal axis, then the hot spots conjecture holds for φ.*

As another consequence of Theorem 1.5.6, we obtain the following properties of the level sets of antisymmetric second Neumann eigenfunctions of convex $C^{1,\alpha}$ domains with a line of symmetry:

Corollary 1.5.8. *Let D be a convex $C^{1,\alpha}$ domain ($0 < \alpha < 1$) which is symmetric with respect to the horizontal axis. If φ is a second Neumann eigenfunction for D which is antisymmetric with respect to the horizontal axis, then the level curves of φ are unions of disjoint simple curves in $\overline{D}$ (possibly reduced to a single point). Moreover, no level curve of φ can terminate in D.*

1.5.2 Comparisons with known results

The hot spots conjecture is known to be true for relatively small classes of domains (parallelepipeds, balls and annuli in $\mathbb{R}^d$, obtuse triangles). According to the knowledge of the author, the only papers in the literature which contain the proof of the hot spots conjecture for general classes of domains are [7] and [48]. We will refer to these papers for a comparison of our results.

For a bounded planar domain D, we will denote the diameter of D by d_D, the length of the projection of D on the vertical axis by w_D, the length of the projection of D on the horizontal axis by l_D, and will refer them as the diameter, the width and respectively the length of D.

In [7], Bañuelos and Burdzy used probabilistic techniques (based on synchronous and mirror couplings of reflecting Brownian motions in polygonal domains) to prove the following:

Theorem 1.5.9. *Suppose that a convex polygonal domain D is symmetric with respect to the horizontal axis S and the ratio d_D/w_D is greater than 1.54. Let x and y be the intersection points of S with ∂D. Make at least one of the following two assumptions:*

a) D has another line of symmetry S_1 which is perpendicular on S;

b) For every $r > 0$, the intersection of the circle $\partial B(x, r)$ with D is either empty or is a connected arc, and the same holds for $\partial B(y, r)$.

Then hot spots conjecture, holds for D.

More recently, Jerison and Nadirashvili used deformation of the domain techniques (see [48]) in order to refine the above result in the case of domains having two orthogonal axes of symmetry. They showed the following:

Theorem 1.5.10. *Let Ω be a bounded convex domain in the plane that is symmetric with respect to both coordinate axes. Let u be any Neumann eigenfunction with lowest nonzero eigenvalue. Then, except in the case of a rectangle, u achieves its maximum over $\overline{\Omega}$ on the boundary at exactly one point, and likewise for its minimum. Furthermore, if $x^0 \in \partial\Omega$ and $-x^0 \in \partial\Omega$ denote the places where u achieves its maximum and minimum, then u is monotone along the two arcs of the boundary from $-x^0$ to x^0. Let ν be any outer normal to $\partial\Omega$ at x^0, that is $\nu \cdot (x - x^0) < 0$ for all $x \in \Omega$. Then $\nu \cdot \nabla u(x) > 0$ for all $x \in \Omega$.*

As indicated in [48], the uniqueness of the location of the extrema of u may fail for domains not having two orthogonal axis of symmetry (for example an equilateral triangle has a second Neumann eigenfunction which attains its maximum at two of its vertices and the minimum at the third vertex). We make the remark that in the present paper we are concerned with the hot spots conjecture in the form presented in Conjecture 1.5.1, and we will not attempt to prove that the extrema are attained at a single point (this may fail for general domains, as showed in the example above).

Theorem 1.5.9 provides additional hypotheses under which the hot spots conjecture holds, for the case of convex domains having just one axis of symmetry (additional hypothesis b)).

We will show that our main result in Theorem 1.5.6 gives the hot spots conjecture for smooth convex domains with two orthogonal axes of symmetry (without any restriction on their diameter to width ratio), thus obtaining the cited results (Theorem 1.5.9 with the additional hypothesis a) and Theorem 1.5.10).

Also, we will show that we can apply Theorem 1.5.6 to the case of domains having just one axis of symmetry and satisfying an additional restriction on their diameter to width ratio, obtaining a result similar to Theorem 1.5.9 (with additional hypothesis b)), but which is complimentary to it.

We will consider first the case of convex domains having two orthogonal axes of symmetry. We have:

Theorem 1.5.11. *If D is a convex $C^{1,\alpha}$ domain ($0 < \alpha < 1$), symmetric with respect to both coordinate axes, then hot spots conjecture holds for D.*

Proof. It is known (see [60]) that the eigenspace corresponding to the second Neumann eigenvalue of a simply connected planar domain is at most 2-dimensional.

We will consider first the case when the eigenspace corresponding to the second Neumann eigenvalue for D is 1-dimensional. We will show that in this case we can find a second Neumann eigenfunction $\widetilde{\varphi}$ for D which is antisymmetric with respect to one of the coordinate axes. From this and using Theorem 1.5.6 the result follows.

Let φ be a second Neumann eigenfunction for D. If φ is antisymmetric with respect to one of the coordinate axes, we are done. Otherwise we consider the functions $\varphi_1(x,y) = \varphi(x,y) - \varphi(x,-y)$ and $\varphi_2(x,y) = \varphi(x,y) - \varphi(-x,y)$, and note that they are antisymmetric with respect to the horizontal, respectively vertical axis and that both of them are eigenfunctions (possibly identically zero) corresponding to the second eigenvalue of D.

Note that φ_1 and φ_2 cannot be both identically zero. This is so for otherwise φ would be symmetric with respect to both coordinate axes. However, this cannot happen since the nodal line of φ cannot form a closed loop and since φ has exactly two nodal domains.

It follows that we can always find a (not identically zero) second eigenfunction $\widetilde{\varphi}$ for D which is antisymmetric with respect to one of the coordinate axes. It follows that Theorem 1.5.6 applies to $\widetilde{\varphi}$, and therefore the hot spots conjecture holds for $\widetilde{\varphi}$. Since the eigenspace corresponding to the second Neumann eigenvalue for D is 1-dimensional, it follows that the hot spots conjecture holds for D, ending the proof in this case.

We will consider now the case when the eigenspace corresponding to the second Neumann eigenvalue is 2-dimensional. We will show that we can find two independent second Neumann eigenfunctions $\widetilde{\varphi}_1$ and $\widetilde{\varphi}_2$ which are antisymmetric with respect to horizontal, respectively vertical axis.

Consider φ_1 and φ_2 two linearly independent second Neumann eigenfunctions for D.

Note that if one of the eigenfunctions is symmetric with respect to one of the axes, then it has to be antisymmetric with respect to the other axis; conversely, if one of the eigenfunctions is antisymmetric with respect to one of the axes, then it has to be symmetric with respect to the other axis. To see this, if for example φ_1 is antisymmetric with respect to horizontal axis, note that $\overline{\varphi}_1(x,y) = \widetilde{\varphi}_1(x,y) - \widetilde{\varphi}_1(-x,y)$ is a second Neumann eigenfunction for D, antisymmetric with respect to both coordinate axes; unless $\overline{\varphi}_1$ is identically zero, $\overline{\varphi}_1$ would have at least four nodal domains, which is impossible by Courant Nodal domain theorem. Hence $\overline{\varphi}_1$ is identically zero, or equivalent φ_1 is symmetric with respect to the horizontal axis. Similar reasoning applies to the other cases.

It follows that φ_1 and φ_2 cannot be symmetric with respect to the same symmetry axis, for they are independent. Also, neither φ_1, nor φ_2 cannot be symmetric with respect to both symmetry axes.

Without loss of generality we can thus assume that φ_1 is not symmetric with respect to horizontal, and that φ_2 is not symmetric with respect to vertical axis. We can therefore choose $\widetilde{\varphi}_1(x,y) = \varphi_1(x,y) - \varphi_1(x,-y)$ and $\widetilde{\varphi}_2(x,y) = \varphi_2(x,y) - \varphi_2(-x,y)$ and note that they are independent, not identically zero second Neumann eigenfunctions for D, antisymmetric with respect to horizontal, respectively vertical axis. Moreover, from the previous part of the proof it follows that $\widetilde{\varphi}_1$ is symmetric with respect to the vertical axis, and $\widetilde{\varphi}_2$ is symmetric with respect to the horizontal axis.

It follows that Theorem 1.5.6 applies to $\widetilde{\varphi}_1$ and $\widetilde{\varphi}_2$, and therefore the hot spots conjecture holds for them. Moreover, $\widetilde{\varphi}_1$ is strictly monotone on the family of hyperbolic lines intersecting the horizontal axis and $\widetilde{\varphi}_2$ is strictly monotone on the family of hyperbolic lines intersecting the vertical axis. In particular, both $\widetilde{\varphi}_1$ and $\widetilde{\varphi}_2$ are strictly monotone on all hyperbolic lines passing through the origin.

Consider now an arbitrarily chosen second Neumann eigenfunction φ for D. Since $\widetilde{\varphi}_1$ and $\widetilde{\varphi}_2$ are independent and the eigenspace corresponding to the second Neumann eigenvalue is 2-dimensional, we can find constants a, b such that $\varphi = a\widetilde{\varphi}_1 + b\widetilde{\varphi}_2$.

Consider a point $z_0 = (x_0, y_0) \in \overline{D}$ where φ attains its maximum over $\overline{D}$. We will show that we must have $z_0 \in \partial D$. If $a\widetilde{\varphi}_1(z_0) < 0$, since $\widetilde{\varphi}_1$ is antisymmetric with respect to the horizontal axis and $\widetilde{\varphi}_2$ is symmetric with respect to the horizontal axis, we obtain:

$$\begin{aligned}\varphi(z_0) &= a\widetilde{\varphi}_1(x_0,y_0)+b\widetilde{\varphi}_2(x_0,y_0)\\ &< a\widetilde{\varphi}_1(x_0,-y_0)+b\widetilde{\varphi}_2(x_0,-y_0)\\ &= \varphi(x_0,-y_0),\end{aligned}$$

a contradiction. It follows that we must have $a\widetilde{\varphi}_1(z_0)\geq 0$, and similarly $b\widetilde{\varphi}_2(z_0)\geq 0$. Since both $\widetilde{\varphi}_1$ and $\widetilde{\varphi}_2$ are strictly monotone on the hyperbolic line through 0 and z_0, it follows that φ is strictly increasing on the hyperbolic half line with the endpoint 0 and passing through z_0. Since φ attains its maximum at z_0, it follows that $z_0\in\partial D$. Similar reasoning shows that if φ attains its minimum at a point, then the point must belong to the boundary ∂D, and therefore the hot spots conjecture holds for φ.

Since φ was arbitrarily chosen, it follows that the hot spots conjecture holds for D. □

In [7] it is shown that if D is a convex domain symmetric with respect to the horizontal axis, and the ratio d_D/w_D is larger than 1.54, then the eigenspace corresponding to the second eigenvalue is 1-dimensional, and there is no second Neumann eigenfunction for D which is antisymmetric with respect to the horizontal axis, hence and our result in Theorem 1.5.6 does not apply. Rephrasing this, we can say that if a domain D having the horizontal axis as line of symmetry is "long enough", then the second Neumann eigenfunctions for D have to be symmetric with respect to the horizontal axis. One might expect that if the domain is "wide enough", then the second Neumann eigenfunctions have to be antisymmetric with respect to the horizontal axis. This is true, and we state:

Proposition 1.5.12. *Suppose that D is a convex $C^{1,\alpha}$ domain ($0<\alpha<1$) which is symmetric with respect to the horizontal axis, and the diameter to length ratio d_D/l_D is larger than $\frac{4j_0}{\pi}\approx 3.06$. Then the eigenspace corresponding to the second Neumann eigenvalue for D is 1-dimensional and it is given by a function which is antisymmetric with respect to the horizontal axis.*

Proof. The first part follows from Proposition 2.4i) in [7]. To prove the second part of the statement, we will use a modification of an argument found in [7].

Assume there exists a second Neumann eigenfunction φ for D which is symmetric with respect to the horizontal axis. Let D_1 be one of the nodal domains of φ, with the property that the horizontal component of the inner pointing normal at the common boundary of D_1 and D always points in the same direction (i.e. either left or right). The existence of D_1 with the above property follows from the convexity of the domain and the fact that the nodal domains of φ are symmetric with respect to the horizontal axis.

We will estimate the first mixed Dirichlet-Neumann eigenvalue for D_1 (denoted λ_1), which is the same as the second eigenvalue for D (denoted μ_2).

For this, consider a reflecting Brownian motion (X_t,Y_t) in D_1, killed upon hitting the nodal line of φ. Note that the horizontal component X_t is a Brownian motion plus (or minus) a nondecreasing process (we are using here the fact that the horizontal component of the inner pointing normal to the common boundary of D_1 and D is always pointing in the same direction).

By a comparison argument, it can be shown that the distribution of the lifetime of X_t is majorized by the distribution of the lifetime of a one dimensional Brownian motion on a interval of length l_D, killed at one end, and reflected at the other. This latter is majorized by $ce^{-\lambda t}$, where λ is the first Dirichlet-Neumann eigenvalue for the given interval, i.e. $\lambda=\frac{\pi^2}{8l_D^2}$ (for it is the second Neumann eigenvalue of an interval of double length, i.e. $\frac{\pi^2}{2(2l_D)^2}$). It follows that we have $\mu_2=\lambda_1\geq\lambda=\frac{\pi^2}{8l_D^2}$.

It is also known (see [7]) that for convex domains, we have $\mu_2\leq\frac{2j_0^2}{l_D^2}$, thus we must have $\frac{\pi^2}{8l_D^2}\leq\frac{2j_0^2}{d_D^2}$, or equivalent $\frac{d_D}{l_D}\leq\frac{4j_0}{\pi}$.

It follows that if this inequality doesn't hold (i.e. for sufficiently "wide" domains, symmetric with respect to the horizontal axis), there are no second Neumann eigenfunctions for D which are symmetric with respect to the horizontal axis.

To conclude the proof, consider an arbitrarily fixed second Neumann eigenfunction φ_1 for D, and define $\widetilde{\varphi}(x,y) = \varphi_1(x,y) - \varphi_1(x,-y)$. From the previous part of the proof it follows that $\widetilde{\varphi}$ is not identically zero, and hence it is a second Neumann eigenfunction for D, antisymmetric with respect to the horizontal axis. □

Using the above proposition and Corollary 1.5.7 we obtain the proof of the hot spots conjecture for a new class of domains, as follows:

Corollary 1.5.13. *Suppose that D is a $C^{1,\alpha}$ convex domain ($0 < \alpha < 1$), symmetric with respect to the horizontal axis and the diameter to width ratio d_D/l_D is larger than $\frac{4j_0}{\pi} \approx 3.06$. Then the hot spots conjecture holds for D.*

1.5.3 Further developments

We conclude with three remarks.

First, note that because of the ratio diameter to width restriction in Theorem 1.5.9, essentially the results of Burdzy and Bañuelos apply to convex domains with one axis of symmetry, for which the eigenspace corresponding to the second Neumann eigenvalue is 1-dimensional, being given by an eigenfunction which is symmetric with respect to the axis of symmetry. Our result in Corollary 1.5.13 is complimentary to this, giving the proof of the hot spots conjecture for (smooth) convex domains with one axis of symmetry, for which the eigenspace corresponding to the second Neumann eigenvalue is 1-dimensional, being given by an eigenfunction which is antisymmetric with respect to the axis of symmetry.

In our view, these two results together should give a resolution to the hot spots conjecture in the case of convex domains with a line of symmetry (and no restrictions on their diameter to width), but we were not able to implement it.

Secondly, we will discuss how the hypotheses of our main result in Theorem 1.5.6 can be weakened. There are two main ingredients in the proof: the symmetry and the convexity.

Even though the symmetry of the domain and the antisymmetry of the Neumann eigenfunction hypotheses are needed in the proof, we can carry the proof with weaker assumptions. A carefull examination of the proof shows that we can replace these hypotheses as follows:

Theorem 1.5.14. *Let D be a convex $C^{1,\alpha}$ $(0 < \alpha < 1)$ domain and let φ be a second Neumann eigenfunction for D. If the nodal line of φ is a hyperbolic line in D, then φ is strictly monotone on the family of hyperbolic lines in D which intersect nontrivially the nodal line of φ.*

In particular the hot spots conjecture holds for φ.

The hypothesis on the nodal line φ in the above theorem can still be weakened, by requiring instead that the nodal domains of φ are *hyperbolically starlike in D (*i.e. starlike with respect to the hyperbolic lines in D). Geometrically this means that the nodal line of φ is "completely visible" from w along the hyperbolic lines in D. We state:

Theorem 1.5.15. *Let D be a convex $C^{1,\alpha}$ ($0 < \alpha < 1$) domain and let φ be a second Neumann eigenfunction for D, with nodal domains $D^+ = \{x \in D : \varphi(x) > 0\}$ and $D^- = \{x \in D : \varphi(x) < 0\}$. Assume there exists a point $w \in D^-$ such that D^- is hyperbolically starlike with respect to w. Then φ is monotone in D^+ along the family of hyperbolic lines in D which pass through w. In particular φ must attain its maximum over $\overline{D}$ at the boundary of D. Similar statement holds for D^-.*

The convexity of the domain is a key element in our construction of scaling coupling of reflecting Brownian motions (see Proposition 1.4.9), needed in order to prove the hot spots conjecture, and therefore we cannot dispense of it. However, the scaling coupling can be defined in certain annuli-like domains, and it has the same properties outlined in our construction for convex $C^{1,\alpha}$ domains $(0 < \alpha < 1)$. This in turn leads to a proof of the hot spots conjecture for certain type of second Neumann eigenfunctions of doubly connected domains, where almost nothing is known in the literature. We have:

Theorem 1.5.16. *Let $D = \{f(z) : a < |z| < 1\}$, where $f : U \to \mathbb{C}$ is a conformal map of U onto a convex $C^{1,\alpha}$ domain ($0 < \alpha < 1$) and $a \in (0, 1)$ is arbitrarily fixed. If φ is a second Neumann eigenfunction for D, for which the nodal line is a part of a hyperbolic line l in $f(U)$, then φ is monotone in D on the family of hyperbolic lines in $f(U)$ passing through $f(0)$. In particular φ attains its maximum and minimum over $\overline{D}$ at the boundary of D.*

1.6 A Hot-Spots property for the Brownian motion with killing and reflection

In this section we investigate a "hot–spots" property for the survival time probability of Brownian motion with killing and reflection in planar convex domains whose boundary consists of two curves, one of which is an arc of a circle, intersecting at acute angles. In turn, this leads to the "hot–spots" property for the mixed Dirichlet–Neumann eigenvalue problem in the domain with Neumann conditions on one of the curves and Dirichlet conditions on the other.

The Hot Spots conjecture introduced in Section 1.5 can be formulated in terms of a mixed Dirichlet–Neumann eigenvalue problem as discussed in [9] and [64]. In this section we explore this mixed boundary value problem further and in particular we extend the results in [64] and [9].

Briefly, the connection is the following. Assume that D is a planar convex domain for which the Laplacian with Neumann boundary conditions has discrete spectrum (see Section 1.5 for the details about eigenfunctions and eigenvalues). Under various conditions on D (see for example [13]) it can be shown that the second Neumann eigenvalue λ_2 for the Laplacian on D is simple (note that by Courant's Nodal Line Theorem, the multiplicity of λ_2 is at most 2). If φ_2 is any Neumann eigenfunction corresponding to λ_2, and γ is its corresponding nodal line, and $D_{1,2}$ are the corresponding nodal domains, then the restriction of φ_2 to D_1 (or D_2) is an eigenfunction corresponding to the smallest eigenvalue μ_1 for the Laplacian in D_1 with Dirichlet boundary conditions on γ and Neumann boundary conditions on $\partial D_1 \backslash \gamma$. In this formulation, the Hot Spots Conjecture 1.5.1 is equivalent to the assertion that the restrictions of φ_2 to D_1 and D_2 attain their extrema on, and only on, ∂D_1, respectively ∂D_2.

The results in [64] and [9] can be stated in terms of the above mixed Dirichlet-Neumann boundary value problem as follows. Suppose that D is planar convex domain whose boundary consists of the curve γ_1 and the line segment γ_2. Let μ_1 be the lowest eigenvalue for the Laplacian in D with Neumann boundary conditions on γ_1 and Dirichlet boundary conditions on γ_2. Let $\psi_1 : \overline{D} \to [0, \infty)$ be the ground state eigenfunction (unique up to a multiplicative constant) corresponding to μ_1. Then ψ_1 attains its maximum on, and only on, γ_1. In fact, the results in [64], [9] prove more. If B_t is a reflecting Brownian motion in D starting at $z \in \overline{D}$ and killed on γ_2, and τ denote its lifetime (the first time B_t hits γ_2), then, for an arbitrarily fixed $t > 0$, the function $u(z) = P^z\{\tau > t\}$ attains its maximum, as a function of $z \in \overline{D}$, on, and only on, γ_1. Furthermore, both functions $u(z)$ and $\psi_1(z)$ are strictly increasing as z moves toward the boundary γ_1 of D along hyperbolic line segments (see [64] and [9] for the precise definitions of hyperbolic line segments and for the details of how the result for u implies the result for ψ_1.) The following question, first raised in [9], naturally arises:

Question 1.6.1. *Given a bounded simply connected planar domain whose boundary consists of two smooth curves, what conditions must one impose on these two curves in order for the ground state eigenfunction of the mixed boundary value problem (Dirichlet conditions on one curve and Neumann on the other) to attain its maximum on the boundary and only on the boundary?*

In this section we prove the following theorem which extends the results in [64] and [9] by replacing the hypothesis that γ_2 is a line segment by the hypothesis that γ_2 is an arc of a circle.

Theorem 1.6.2. *Suppose D is a bounded convex planar domain whose boundary consists of two curves $\gamma_1 = (\gamma_1(t))_{t\in[0,1]}$ and $\gamma_2 = (\gamma_2(t))_{t\in[0,1]}$, one of which is an arc of a circle. Further, suppose that the angle between the curves γ_1 and γ_2 is less than or equal to $\frac{\pi}{2}$, that is, the angle formed by the two half-tangents at $\gamma_1(0) = \gamma_2(0)$ and $\gamma_1(1) = \gamma_2(1)$ is less than or equal to $\frac{\pi}{2}$.*

If B_t is a reflecting Brownian motion in D starting at z and killed on γ_2, and if τ_D denote its lifetime, then for each $t > 0$ arbitrarily fixed, the function $u(z) = P^z\{\tau_D > t\}$, $z \in \overline{D}$, attains it maximum on, and only on γ_1.

As a corollary, we derive the following.

Corollary 1.6.3 ("Hot–spots" for the mixed boundary value problem)**.** *Let D be as in Theorem 1.6.2, and let ψ_1 be a first mixed Dirichlet-Neumann eigenfunction for the Laplacian in D, with Neumann boundary conditions on γ_1 and Dirichlet boundary conditions on γ_2. Then the function $\psi_1(z)$, $z \in \overline{D}$, attains its maximum on, and only on γ_1.*

Proof. Follows from Theorem 1.6.2 exactly as in [9] (see [10] for the details). □

Remark 1.6.4. *Moreover, in the proofs of the above results we show that the functions $u(z)$ and $\psi_1(z)$ are in fact strictly increasing along certain families of curves in D: Euclidean radii contained in D in the case when γ_1 is an arc of a circle, and hyperbolic line segments in D (see Definition 1.4.2), in the case when γ_2 is an arc of a circle.*

The proof of Theorem 1.6.2 is different depending on which one of the curves γ_1 or γ_2 is an arc of a circle. In the case when γ_2 is an arc of a circle, the proof rests on several preliminary results, which we present below.

Proposition 1.6.5. *Let D be as in Theorem 1.6.2 and suppose that γ_2 is an arc of a circle $C = \partial B(z_0, R)$. Let D_s be the domain which is symmetric to the domain D with respect to the circle C, that is*

$$D_s = \{z_0 + \frac{R^2}{\overline{z} - \overline{z_0}} : z \in D\}.$$

Then $D^* = D \cup \gamma_2 \cup D_s$ is a convex domain.

Proof. For a complex number z we will use Rez and Imz to denote the real, respectively the imaginary part of the complex number $z \in \mathbb{C}$.

Without loss of generality we can assume that $C = \partial B(0,1)$ is the circle centered at the origin of radius 1 and that $\gamma_1(0)$ and $\gamma_1(1)$ are symmetric with respect to the vertical axis, that is $\mathrm{Im}\gamma_1(0) = \mathrm{Im}\gamma_2(1)$. Further, we may assume that γ_2 contains the point $-i$.

We will first show that $\mathrm{Im}\gamma_1(0) \le 0$. To see this, note that since the domain D is convex, it lies below its half-tangent at the point $\gamma_1(0)$, and by the angle restriction this half-line lies below the line passing through $\gamma_1(0)$ and the origin. If $\mathrm{Im}\gamma_1(0) > 0$ then also $\mathrm{Im}\gamma_1(1) = \mathrm{Im}\gamma_1(0) > 0$, and therefore the point $\gamma_1(1) \in \partial D$ does not lie below (or on) the line determined by $\gamma_1(0)$ and 0, a contradiction. We must therefore have $\mathrm{Im}\gamma_1(0) = \mathrm{Im}\gamma_1(1) \le 0$.

If $\mathrm{Im}\gamma_1(0) = \mathrm{Im}\gamma_1(1) = 0$, by the angle restriction at these points, together with the fact that D is a convex domain (and hence γ_1 is a concave down curve), it follows that the curve γ_1 is in this case the line segment $[-1, 1]$, and therefore $D = \{z \in \mathbb{C} : \mathrm{Im}z < 0, |z| < 1\}$. The proof is trivial in this case since $D_s = \{z \in \mathbb{C} : \mathrm{Im}z < 0, |z| > 1\}$, and therefore $D^* = D \cup \gamma_2 \cup D_s = \{z \in \mathbb{C} : \mathrm{Im}z < 0\}$ which is a convex domain.

A similar argument shows that if $0 \in \gamma_1 \subset \partial D$, then the curve γ_1 consists of the union of the two line segments from $\gamma_1(0)$ to 0, respectively from 0 to $\gamma_1(1)$, hence $D$is a sector of the unit disk. It follows that $D^* = D \cup \gamma_2 \cup D_s = \{z \in \mathbb{C} - \{0\} : \arg\gamma_1(0) < \arg z < \arg\gamma_1(1)\}$, which is again a convex set. We can therefore assume that $\mathrm{Im}\gamma_1(0) = \mathrm{Im}\gamma_1(1) < 0$ and $0 \notin D \cup \partial D$. It follows that the domain D is contained in the circular sector $\{z \in \mathbb{C} - \{0\} : |z| < 1, \arg\gamma_1(0) < \arg z < \arg\gamma_1(1)\}$, and therefore $D^* = D \cup \gamma_2 \cup D_s$ is contained in $\{z \in \mathbb{C} - \{0\} : \arg\gamma_1(0) < \arg z < \arg\gamma_1(1)\}$. It follows that for any points $w_1, w_2 \in D^* = D \cup \gamma_2 \cup D_s$, the line segment $[w_1, w_2]$ may intersect the circle C only on the arc γ_2 (and not on $C - \gamma_2$). Since D is convex domain, it follows that $D^* = D \cup \gamma_2 \cup D_s$ is a convex domain if and only if

$$w_1 \in D_s, w_2 \in \gamma_2 \cup D_s \text{ s.t. } [w_1, w_2] \cap \gamma_2 \in \{\emptyset, \{w_2\}\} \Rightarrow [w_1, w_2] \subset D^*, \tag{1.6.1}$$

where $[w_1, w_2]$ denotes the line segment with endpoints w_1 and w_2.

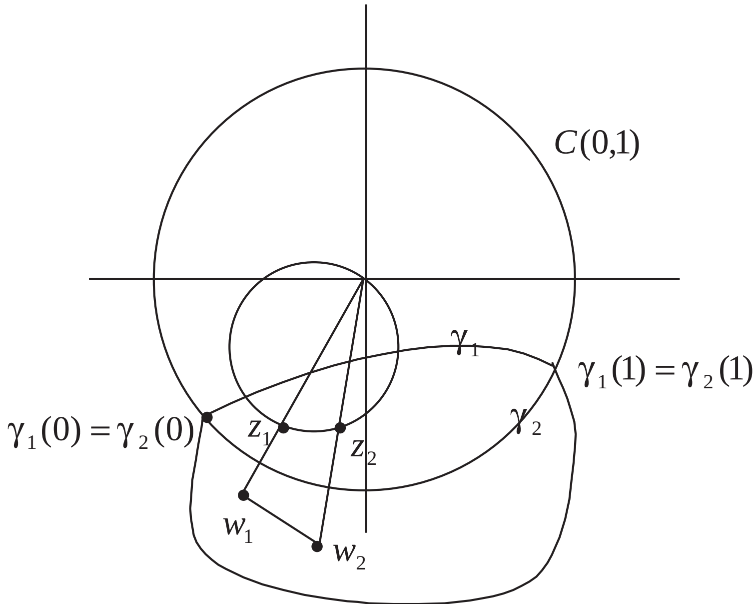

Figure 1.2: The set $D^* = D \cup \gamma_2 \cup D_s$.

Since the set is symmetric to a line with respect to C is a circle passing through the origin, by letting z_1, z_2 be the symmetric points of w_1, respectively w_2 with respect to C, (1.6.1) can be rewritten equivalently as

$$z_1 \in D, z_2 \in \gamma_2 \cup D \text{ s.t. } \widehat{z_1 z_2} \cap \gamma_2 \in \{\emptyset, \{z_2\}\} \Rightarrow \widehat{z_1 z_2} \subset \gamma_2 \cup D, \tag{1.6.2}$$

where $\widehat{z_1 z_2}$ denotes the arc of the circle $C(0, z_1, z_2)$ passing through z_1, z_2 and 0, between (and including) z_1 and z_2, and not containing 0. If the points z_1, z_2 and 0 are collinear, the arc $\widehat{z_1 z_2}$ becomes the line segment $[z_1, z_2]$.

To show the claim, we will prove (1.6.2). Let $z_1 \in D$, $z_2 \in \gamma_2 \cup D$ such that $\widehat{z_1 z_2} \cap \gamma_2 \in \{\emptyset, \{z_2\}\}$. If the points $0, z_1$ and z_2 are collinear, $\widehat{z_1 z_2} = [z_1, z_2] \subset \gamma_2 \cup D$, so we may assume that $0, z_1$ and z_2 are not collinear.

Assume first that the circle $C(0, z_1, z_2)$ does not intersect C. Since γ_1 bounds the convex domain D, the intersection $\gamma_1 \cap C(0, z_1, z_2)$ consists of exactly two points u_1 and u_2(see Figure 1.2). It follows that the intersection between D and $C(0, z_1, z_2)$ is the arc $\widehat{u_1 u_2}$, and therefore we have $\widehat{z_1 z_2} \subset \widehat{u_1 u_2} \subset D$ in this case.

If the circle $C(0, z_1, z_2)$ intersects C, the intersection $C(0, z_1, z_2) \cap D$ is either one or two (connected) arcs c_1 and c_2. Note that z_1 and z_2 must lie on the same connected arc c_i ($i = 1$ or $i = 2$), for otherwise the intersection $\widehat{z_1 z_2} \cap \gamma_2$ would consist of two distinct points (the two endpoints of c_1 and c_2 lying on γ_2). If $z_1, z_2 \in c_1$, since c_1 is a connected arc lying in D, we have $\widehat{z_1 z_2} \subset c_1 \cup \gamma_2 \subset D \cup \gamma_2$ and the claim follows. This completes the proof of the Proposition. □

As a corollary, we obtain the following.

Corollary 1.6.6. *Let D be as in Theorem 1.6.2 and suppose that γ_2 is an arc of a circle. Let $U = \{z \in \mathbb{C} : |z| < 1\}$ be the unit disk and $U^+ = \{z \in U : \mathrm{Im} z > 0\}$ be the upper half-disk. Let $f : \overline{U^+} \to \overline{D}$ be a conformal map such that $f[-1, 1] = \gamma_2$. Then f extends to a conformal map from U onto the convex domain D^*.*

Proof. Follows by using the Schwarz reflection principle and the previous lemma. See [10] for the details. □

In order to prove Theorem 1.6.2 in the case when γ_2 is an arc of a circle we also need the following theorem, which may be of independent interest.

Theorem 1.6.7. *Let $U_d = \{\zeta \in \mathbb{R}^d : \|\zeta\| < 1\}$ be the unit ball in $\mathbb{R}^d$, $d \geq 2$, and let $U_d^+ = \{\zeta = (\zeta_1, ..., \zeta_d) \in U_d : \zeta_d > 0\}$ be the upper hemisphere in $\mathbb{R}^d$.*

Suppose that $V : \overline{U_d^+} \to (0, \infty)$ is a continuous potential for which $r^2 V(r\zeta)$ is a nondecreasing function of $r \in (0, \frac{1}{\|\zeta\|})$ for any $\zeta \in U_d^+$ arbitrarily fixed. That is, suppose that

$$r_1^2 V(r_1\zeta) \leq r_2^2 V(r_2\zeta), \tag{1.6.3}$$

for all $\zeta \in U_d^+$, $0 < r_1 < r_2 < \frac{1}{\|\zeta\|}$.

Let B_t be a reflecting Brownian motion in U_d^+ killed on the hyperplane $H = \{\zeta = (\zeta_1, ..., \zeta_d) \in \mathbb{R}^d : \zeta_d = 0\}$, and let $\tau_{U_d^+}$ denote its lifetime. Then for any arbitrarily fixed $t > 0$ and $\zeta \in U_d^+$, $P^{r\zeta}\left(\int_0^{\tau_{U_d^+}} V(B_s)ds > t\right)$ is a non-decreasing function of $r \in (0, \frac{1}{\|\zeta\|})$, that is

$$P^{r_1\zeta}\left(\int_0^{\tau_{U_d^+}} V(B_s)ds > t\right) \leq P^{r_2\zeta}\left(\int_0^{\tau_{U_d^+}} V(B_s)ds > t\right), \tag{1.6.4}$$

for all $t > 0$, $\zeta \in U_d^+$ and $0 < r_1 < r_2 < \frac{1}{\|\zeta\|}$.

Moreover, if the inequality in (1.6.3) is a strict inequality, so is the one in (1.6.4).

Proof. Fix $t > 0$, $\zeta \in U_d^+$ and $0 < r_1 < r_2 < \frac{1}{\|\zeta\|}$.

Consider a scaling coupling of reflecting Brownian motions $(B_t, \widetilde{B}_t)$ in the unit ball U_d starting at $(r_1\zeta, r_2\zeta)$ (see Section 1.3). More precisely, let B_t be reflecting Brownian motion in U_d starting at $r_1\zeta \in U_d$, with its natural filtration $(\mathcal{F}_t)$, and consider

$$\widetilde{B}_t = \frac{1}{M_{\alpha_t}} B_{\alpha_t}, \quad t \geq 0, \tag{1.6.5}$$

where

$$M_t = \frac{r_1}{r_2} \vee \sup_{s \leq t} \|B_s\|, \tag{1.6.6}$$

$$A_t = \int_0^t \frac{1}{M_s^2} ds, \tag{1.6.7}$$

and

$$\alpha_t = \inf\{s > 0 : A_s \geq t\}. \tag{1.6.8}$$

Theorem 1.3.1 and Remark 1.3.2 show that $\widetilde{B}_t$ is an $(\mathcal{F}_{\alpha_t})$-adapted reflecting Brownian in U_d.

Letting $\tau_{U_d^+}$, $\widetilde{\tau}_{U_d^+}$ denote the killing times of B_t, respectively $\widetilde{B}_t$, on the hyperplane H, we have almost surely $\tau_{U_d^+} = \alpha_{\widetilde{\tau}_{U_d^+}}$, and therefore we obtain

$$\begin{aligned} \int_0^{\tau_{U_d^+}} V(B_s)ds &= \int_0^{\alpha_{\widetilde{\tau}_{U_d^+}}} V(B_s)ds \\ &= \int_0^{\widetilde{\tau}_{U_d^+}} V(B_{\alpha_u})d\alpha_u \\ &= \int_0^{\widetilde{\tau}_{U_d^+}} V(B_{\alpha_u})M_{\alpha_u}^2 du \\ &\leq \int_0^{\widetilde{\tau}_{U_d^+}} V(\frac{1}{M_{\alpha_u}} B_{\alpha_u})du \\ &= \int_0^{\widetilde{\tau}_{U_d^+}} V(\widetilde{B}_u)du. \end{aligned} \tag{1.6.9}$$

The inequality above follows from the assumption that $r^2 V(r\zeta)$ is a non-decreasing function of r for $\zeta \in U_d^+$ arbitrarily fixed:

$$V(B_{\alpha_u}) = 1^2 V(1 B_{\alpha_u}) \leq \frac{1}{M_{\alpha_u}^2} V\left(\frac{1}{M_{\alpha_u}} B_{\alpha_u}\right),$$

since by (1.6.6) we have $M_{\alpha_u} \leq 1$ for all $u \geq 0$.

By the construction above, $(B_t, \widetilde{B}_t)$ is a pair of reflecting Brownian motions in U_d starting at $(r_1\zeta, r_2\zeta)$, and the inequality (1.6.9) shows that in particular we have

$$P^{r_1\zeta}\left\{\int_0^{\tau_{U_d^+}} V(B_s)ds > t\right\} \leq P^{r_2\zeta}\left\{\int_0^{\widetilde{\tau}_{U_d^+}} V(\widetilde{B}_s)ds > t\right\},$$

which proves the first part of the Theorem 1.6.7.

To prove the strict increasing part of the theorem, we will use the following support lemma for the d-dimensional Brownian motion (see [80], page 374).

Lemma 1.6.8. *Given an d-dimensional Brownian motion B_t starting at x and a continuous function $f : [0, 1] \to \mathbb{R}^d$ with $f(0) = x$ and $\varepsilon > 0$, we have*

$$P^x\left(\sup_{t\leq 1} \|B_t - f(t)\| < \varepsilon\right) > 0.$$

Assume now that we have strict inequality in (1.6.3). By the continuity of the potential $V : \overline{U_d^+} \to (0, \infty)$ and the strict monotonicity of $r^2V(r\zeta)$ for $0 < r < \frac{1}{\|\zeta\|}$, we have

$$\int_0^1 V((1-u)r_1\zeta)du < \int_0^1 \left(\frac{r_2}{r_1}\right)^2 V\left(\frac{r_2}{r_1}(1-u)r_1\zeta\right)du,$$

and therefore we can choose $T > 0$ such that

$$T\int_0^1 V((1-u)r_1\zeta)du < t < T\int_0^1 \left(\frac{r_2}{r_1}\right)^2 V\left(\frac{r_2}{r_1}(1-u)r_1\zeta\right)du,$$

and we may further choose $\varepsilon > 0$ and $\delta > 0$ small enough so that

$$\frac{T}{1+\delta}\int_0^{1+\frac{\varepsilon}{r_1}} V((1-u)r_1\zeta)du < t < \frac{T}{1+\delta}\int_0^{1-\frac{\varepsilon}{r_1}} \left(\frac{r_2}{r_1}\right)^2 V\left(\frac{r_2}{r_1}(1-u)r_1\zeta\right)du. \tag{1.6.10}$$

Consider the function $f : \mathbb{R} \to \mathbb{R}^d$ defined by

$$f(s) = \left(1 - \frac{(1+\delta)}{T}s\right)r_1\zeta.$$

With the change of variable $u = \frac{1+\delta}{T}s$, the double inequality in (1.6.10) can be rewritten as

$$\int_0^{\frac{1+\frac{\varepsilon}{r_1}}{1+\delta}T} V(f(s))ds < t < \int_0^{\frac{1-\frac{\varepsilon}{r_1}}{1+\delta}T} \left(\frac{r_2}{r_1}\right)^2 V\left(\frac{r_2}{r_1}f(s)\right)ds.$$

By eventually choosing a smaller $\varepsilon > 0$, and by the uniform continuity of V on $\overline{U^+}$, we also have

$$\int_0^{\frac{1+\frac{\varepsilon}{r_1}}{1+\delta}T} V(b(s))ds < t < \int_0^{\frac{1-\frac{\varepsilon}{r_1}}{1+\delta}T} \left(\frac{r_2}{r_1}\right)^2 V\left(\frac{r_2}{r_1}b(s)\right)ds, \tag{1.6.11}$$

for any continuous functions $b : [0, \frac{T}{1+\delta}] \to \mathbb{R}^n$ such that

$$\sup_{s\leq\frac{T}{1+\delta}} \|b(s) - f(s)\| < \varepsilon.$$

Let B_t and $\tilde{B}_t$ be the reflecting Brownian motions in $U_d$starting at $r_1\zeta$, respectively $r_2\zeta$, as constructed above. By Lemma 1.6.8, B_t lies in the ε-tube about $f(t)$ for $0 < t < T$ with positive probability, that is,

$$P\left(\sup_{s\leq T} |B_s - f(s)| < \varepsilon\right) > 0.$$

We may assume that $\varepsilon > 0$ is chosen small enough so that this tube does not intersect ∂U, and therefore on a set Q of positive probability, the coupled Brownian motion $\widetilde{B}_s$ does not reach ∂U_d, hence the process M_s is constant on this set.

Thus, on the set Q we have

$$\begin{aligned} M_s &= \frac{r_1}{r_2}, && (1.6.12)\\ A_s &= \int_0^s \frac{1}{M_u^2} du = \left(\frac{r_2}{r_1}\right)^2 s, && (1.6.13)\\ \alpha_s &= A_s^{-1} = \left(\frac{r_1}{r_2}\right)^2 s, && (1.6.14)\end{aligned}$$

and $\widetilde{\tau}_{U_d^+} = A_{\tau_{U_d^+}} = \left(\frac{r_2}{r_1}\right)^2 \tau_{U_d^+}$. Therefore on Q we have

$$\begin{aligned} \int_0^{\widetilde{\tau}_{U_d^+}} V(\widetilde{B}_s)ds &= \int_0^{\left(\frac{r_2}{r_1}\right)^2 \tau_{U_d^+}} V\left(\frac{1}{M_{\alpha_s}} B_{\alpha_s}\right) ds && (1.6.15)\\ &= \int_0^{\tau_{U_d^+}} \left(\frac{r_2}{r_1}\right)^2 V\left(\frac{r_2}{r_1} B_u\right) du\\ &> \int_0^{\tau_{U_d^+}} V(B_s)ds. \end{aligned}$$

Also, by the construction of the set Q we have $\frac{1-\frac{\varepsilon}{r_1}}{1+\delta}T < \tau_{U_d^+} < \frac{1+\frac{\varepsilon}{r_1}}{1+\delta}T$ on Q, and combining with (1.6.11) and (1.6.15), we obtain the strict inequality

$$\begin{aligned} \int_0^{\tau_{U_d^+}} V(B_s)ds &\leq \int_0^{T\frac{1+\frac{\varepsilon}{r_1}}{1+\delta}} V(B_s)ds < t && (1.6.16)\\ &< \int_0^{T\frac{1-\frac{\varepsilon}{r_1}}{1+\delta}} \left(\frac{r_2}{r_1}\right)^2 V\left(\frac{r_2}{r_1} B_s\right) ds\\ &\leq \int_0^{\tau_{U_d^+}} \left(\frac{r_2}{r_1}\right)^2 V\left(\frac{r_2}{r_1} B_s\right) ds\\ &= \int_0^{\widetilde{\tau}_{U_d^+}} V(\widetilde{B}_s)ds, \end{aligned}$$

almost surely on Q.

Therefore we have:

$$\begin{aligned} P^{r_1\zeta}\left\{\int_0^{\tau_{U_d^+}} V(B_s)ds > t\right\} &= P^{r_1\zeta}\left\{\int_0^{\tau_{U_d^+}} V(B_s)ds > t, Q\right\} \\ &+ P^{r_1\zeta}\left\{\int_0^{\tau_{U_d^+}} V(B_s)ds > t, Q^c\right\} \\ &= 0 + P^{r_1\zeta}\left\{\int_0^{\tau_{U_d^+}} V(B_s)ds > t, Q^c\right\} \\ &\leq P^{r_2\zeta}\left\{\int_0^{\widetilde{\tau}_{U_d^+}} V(\widetilde{B}_s)ds > t, Q^c\right\} \\ &< P^{r_2\zeta}\left\{Q\right\} + P^{r_2\zeta}\left\{\int_0^{\widetilde{\tau}_{U_d^+}} V(\widetilde{B}_s)ds > t, Q^c\right\} \\ &= P^{r_2\zeta}\left\{\int_0^{\widetilde{\tau}_{U_d^+}} V(\widetilde{B}_s)ds > t, Q\right\} \\ &+ P^{r_2\zeta}\left\{\int_0^{\widetilde{\tau}_{U_d^+}} V(\widetilde{B}_s)ds > t, Q^c\right\} \\ &= P^{r_2\zeta}\left\{\int_0^{\widetilde{\tau}_{U_d^+}} V(\widetilde{B}_s)ds > t\right\}, \end{aligned}$$

which proves the strict inequality in (1.6.4) in the case when the $r^2V(r\zeta)$ is a strictly increasing function of r, ending the proof of Theorem 1.6.7. □

With this preparation we can now proceed with the proof of Theorem 1.6.2.

Proof of Theorem 1.6.2. We distinguish the following cases.

Case 1. *γ_2 is an arc of a circle.*

Let f a the conformal mapping given by Corollary 1.6.6, and let B_t be a reflecting Brownian motion in U^+ killed on hitting $[-1,1]$, and denote its lifetime by τ_{U^+}.

By Proposition 1.4.8, the potential $V : U^+ \to \mathbb{R}$ defined by $V(z) = |f'(z)|^2$ satisfies the hypothesis of Theorem 1.6.7, and therefore we have

$$P^{z_1}\left\{\int_0^{\tau_{U^+}} |f'(B_s)|^2 ds > t\right\} \leq P^{z_2}\left\{\int_0^{\tau_{U^+}} |f'(B_s)|^2 ds > t\right\}, \tag{1.6.17}$$

for all $t > 0$ and $z_1 = r_1e^{i\theta},\ z_2 = r_2e^{i\theta}$ with $0 < r_1 < r_2 < 1$ and $0 < \theta < \pi$. By Lévy's conformal invariance of the Brownian motion, this is exactly the same as

$$P^{f(z_1)}\left\{\tau_D > t\right\} \leq P^{f(z_2)}\left\{\tau_D > t\right\}, \tag{1.6.18}$$

where τ_D is as in Theorem 1.6.2. From this it follows that the function $u(z) = P^z\{\tau_D > t\}$ is nondecreasing as z moves toward γ_1 along the curve $\gamma_\theta = f\{re^{i\theta} : 0 < r < 1\}$, for any $\theta \in (0,\pi)$ arbitrarily fixed. This together with the real analyticity of the function $u(z)$ implies that $u(z)$ is in is in fact strictly increasing along the family of curves $\{\gamma_\theta : 0 < \theta < \pi\}$, which completes the proof of Theorem 1.6.2 when γ_2 is an arc of a circle.

Case 2. *γ_1 is an arc of a circle.*

Without loss of generality we may assume that γ_1 is an arc of the unit circle centered at the origin. An argument similar to the one in Proposition 1.6.5 shows that $0 \notin D$, and if $0 \in \partial D$ then the domain D is a sector of the unit disk. It either case, the origin belongs to $U \backslash D$.

We claim that $U\backslash D$ is starlike with respect to the origin. If $0 \in \partial D$, the set D is a sector of the unit disk U and the claim follows. We can assume therefore that $0 \notin \overline{D}$.By the angle restriction in the hypothesis of our theorem, together with the convexity of the domain, it follows that D is contained in a sector of the unit disk U, which without loss of generality may be assumed to be symmetric with respect to the imaginary axis. That is, $D \subset \{z \in U : \alpha < \arg z < \pi - \alpha\}$, where $\alpha = \min\{\arg \gamma_1(0), \arg \gamma_1(1)\} \in (0, \frac{\pi}{2})$. Let $z \in U\backslash D$ and $t \in [0,1]$ be arbitrarily fixed. If $\arg z \notin (\alpha, \pi - \alpha)$ then $tz \in U\backslash\{z \in U : \alpha < \arg z < \pi - \alpha\} \subset U\backslash D$. Thus $tz \in U\backslash D$ in this case. If $\arg z \in (\alpha, \pi - \alpha)$ and $tz \notin U\backslash D$, then, since$\frac{1}{|z|}z \in \gamma_1 \subset \overline{D}$, we obtain by the convexity of D that the line segment with endpoints tz and $\frac{1}{|z|}z$ is contained in D, and in particular it follows that $z \in D$, a contradiction. In both cases we obtained that $tz \in U\backslash D$, which proves that $U\backslash D$ is starlike with respect to the origin.

We now follow the proof of Theorem 1.6.7 in the case $d = 2$. For arbitrarily fixed $t > 0$ and $r_1 e^{i\theta}, r_2 e^{i\theta} \in D$ with$r_1 < r_2$, let $(B_t, \widetilde{B}_t)$ be a scaling coupling of reflecting Brownian motions in the unit disk U starting at$(r_1 e^{i\theta}, r_2 e^{i\theta})$, as in the case of Theorem 1.6.2. We note that that if for $s > 0$ we have $\frac{1}{M_s} B_s \in \gamma_2 \subset U\backslash D$, then by the starlikeness of the set $U\backslash D$ also $B_s \in U\backslash D$. That is,

$$\frac{1}{M_s} B_s \notin D \Rightarrow B_{s'} \notin D \text{ for some } 0 < s' \leq s. \tag{1.6.19}$$

Recalling that $\widetilde{B}_s = \frac{1}{M_{\alpha_s}} B_{\alpha_s}$ and that $\alpha_s \leq s$ for all $s > 0$, we can rewrite (1.6.19) as follows

$$\widetilde{B}_s \notin D \Rightarrow B_{s'} \notin D \text{ for some } 0 < s' \leq \alpha_s \leq s, \tag{1.6.20}$$

which in turn is equivalent to

$$\tau_{\gamma_2} \leq \alpha_{\widetilde{\tau}_{\gamma_2}} \leq \widetilde{\tau}_{\gamma_2}, \tag{1.6.21}$$

where τ_{γ_2} and $\widetilde{\tau}_{\gamma_2}$ denote the killing times of B_t, respectively $\widetilde{B}_t$, on the curve γ_2. From this, it follows that we have

$$P^{r_1 e^{i\theta}} \left\{\tau_{\gamma_2} > t\right\} \leq P^{r_2 e^{i\theta}} \left\{\widetilde{\tau}_{\gamma_2} > t\right\}, \tag{1.6.22}$$

and thus the function $u(z) = P^z \left\{\tau_D > t\right\}$ is nondecreasing on the part of the radii $r_\theta = \{re^{i\theta}, 0 < r < 1\}$ which is contained in the domain D. As before, this together with the real analyticity of the function u shows that it is in fact strictly increasing, completing the proof of the theorem. □

We end with some other remarks related to Theorem 1.6.7. Consider the Schrödinger operator $\frac{1}{2}\Delta u - Vu$ in $U_d^+$with Dirichlet boundary conditions on the part of ∂U_d^+ lying in the hyperplane $H = \{(\zeta_1, ..., \zeta_d) \in \mathbb{R}^d : \zeta_n = 0\}$, and Neumann boundary conditions on the "top" portion of the sphere. If we let $P_t^V(\xi, \zeta)$, $\xi, \zeta \in U_d^+$ be the heat kernel for this problem, then

$$u(\xi) = E^\xi \left\{e^{-\int_0^t V(B_s)ds} ; \tau_{U_d^+} > t\right\} = \int_{U_d^+} P_t^V(\xi, \zeta) d\zeta.$$

1.7 Open Problems

Even though there is a rich literature containing many positive results on the validity of Hot Spots Conjecture (see [49], [7], [26], [13], [25], [48], [63], [3], [64], [10], [20], [42], [43], [58], [86], [77], [59], [78], etc), at the present moment the conjecture is still open in its full generality.

It is known that the Hot Spots conjecture holds for balls, parallelepipeds and annuli in $\mathbb{R}^n$ ([49]), smooth convex domains with two orthogonal axes of symmetry (by the results in [64], [7], or [48]), with just one axis of symmetry and an additional hypothesis (if the second Neumann eigenfunction is antisymmetric one can use the results in [64], and the results in [7] in the eigenfunction is symmetric), for nearly circular domains ([58]), for a certain class of doubly connected domains ([20]), etc.

Strikingly, one of the simplest cases for which the Hot Spots conjecture is still open is the case of acute triangles (except for particular acute triangles, such as equilateral or isosceles triangles).

It is also interesting to note that the conjecture holds for obtuse triangles by the results in [7]), but this method fails for acute triangles. This indicates that one needs new tools in approaching this famous conjecture, which drew the attention of many famous mathematicians over the last 30 years since it was formulated, such as J. Rauch, D. Jerisson, N. Nadirashvili, W. Werner, K, Burdzy, R. Bass, R. Bañuelos, and others.

Chapter 2

Mirror coupling of Reflecting Brownian motions

The notion of *mirror coupling* was introduced by W. S. Kendall in [50] in the case of Brownian motions on a complete Riemannian manifold with nonnegative Ricci curvature, and was considered in [85] in the case of reflected processes. In [25], and more recently in [4] and [5], K. Burdzy et al. gave a detailed construction of the mirror coupling of reflecting Brownian motions in a smooth domain in $\mathbb{R}^n$ ($n \geq 2$), and used it in order to derive various properties related to Neumann eigenvalues and eigenfunctions of the Laplacian on D.

Using a detailed analysis of the mirror coupling of reflecting Brownian motions in the case of the unit ball in $\mathbb{R}^n$, in the present chapter we settle a conjecture of R. Laugesen and C. Morpurgo which asserts that the diagonal of the Neumann heat kernel of the unit ball $\mathbb{U} \subset \mathbb{R}^n$ is a strictly increasing radial function, and we prove some other inequalities for the Neumann heat kernel in the ball which are of independent interest.

Next, we present an extension of the mirror coupling, recently obtained by the author in [65], in the case when the two reflecting Brownian motions live in different smooth domains $D_1, D_2 \subset \mathbb{R}^n$ satisfying an additional assumption (this condition is satisfied in particular if D_1, D_2 have non-tangential boundaries and $D_1 \cap D_2$ is a convex domain). As applications of this construction, we derive a unifying proof of the two main results concerning the validity of Chavel's conjecture on the domain monotonicity of the Neumann heat kernel, due to I. Chavel ([29]), respectively W. S. Kendall ([51]), and a new proof of Chavel's conjecture for domains satisfying the ball condition, such that the inner domain is star-shaped with respect to the center of the ball.

The results in this chapter are based on [65], [66] and [67].

2.1 Introduction

The technique of coupling of reflecting Brownian motions is a useful tool, used by several authors in connection to the study of the Neumann heat kernel of the corresponding domain (see [5], [7], [21], [27], [51], [50], [85], [64], etc).

In a series of paper, Krzysztof Burdzy et al. ([4], [5], [7], [21], [25]) introduced the *mirror coupling* of reflecting Brownian motions in a smooth domain $D \subset \mathbb{R}^d$ and used it in order to derive properties of eigenvalues and eigenfunctions of the Neumann Laplaceian on D.

In this chapter we present a detailed analysis of the mirror coupling of reflecting Brownian motions in the unit ball in $\mathbb{R}^n$ ($n \geq 2$), and we show that in this case the hyperplane of symmetry between the two reflecting Brownian motions (the *mirror* of the coupling) moves towards the origin. This allows us to obtain a double inequality for the Neumann heat kernel of the unit ball, and as a corollary we conclude with a short proof of Laugesen-Morpurgo conjecture.

We also show that the mirror coupling can be extended to the case when the two reflecting Brownian motions live in different domains $D_1, D_2 \subset \mathbb{R}^d$.

The main difficulty in the extending the construction of the mirror coupling comes from the fact that the stochastic differential equation(s) describing the mirror coupling has a singularity at the times when coupling occurs. In the case $D_1 = D_2 = D$ considered by Burdzy et al. this problem is not a major problem (although the technical details are quite involved, see [5]), since after the coupling time the processes move together. In the case $D_1 \neq D_2$ however, this is a major problem: after the processes have coupled, it is possible for them to decouple (for example in the case when the processes are coupled and they hit the boundary of one of the domains).

It is worth mentioning that the method used for proving the existence of the solution is new, and it relies on the additional hypothesis that the smaller domain D_2 (or more generally $D_1 \cap D_2$) is a convex domain. This hypothesis allows us to construct an explicit set of solutions in a sequence of approximating polygonal domains for D_2, which converge to the desired solution.

As applications of the extended mirror coupling, we derive a unifying proof of the two most important results on the challenging Chavel's conjecture on the domain monotonicity of the Neumann heat kernel ([29], [51]), and a new proof of Chavel's conjecture for domains satisfying the ball condition, such that the inner domain is star-shaped with respect to the center of the ball. This is also a possible new line of approach for Chavel's conjecture (note that by the results in [12], Chavel's conjecture does not hold in its full generality, but the additional hypotheses under which this conjecture holds are not known at the present moment).

The structure of the chapter is as follows: in Section 2.2 we briefly describe the construction of Burdzy et al. of the mirror coupling in a smooth bounded domain $D \subset \mathbb{R}^d$ and we establish the notation.

In Section 2.3, we begin with a detailed analysis of mirror coupling of reflecting Brownian motions in the unit ball, which shows that the hyperplane of symmetry between the two reflecting Brownian motions (the *mirror* of the coupling) moves towards the origin (Lemma 2.3.16). From this we obtain a comparison result for the transition probabilities of reflecting Brownian motion in the unit ball (Theorem 2.3.17), which is the key for our proof of the Laugesen-Morpurgo conjecture. Using this result, we obtain a double inequality for Neumann heat kernel of the unit ball (the double inequality in Theorem 6.2.4), and as a corollary we conclude with a short proof of Laugesen-Morpurgo conjecture (Theorem 2.3.20).

In Section 2.4, in Theorem 2.4.1, we give the main result which shows that the mirror coupling can be extended to the case when $\overline{D_2} \subset D_1$ are smooth bounded domains in $\mathbb{R}^d$ and D_2 is convex (some extensions of the theorem are presented in Section 2.6).

Before proceeding with the proof of theorem, in Remark 2.4.4 we show that the proof can be reduced to the case when $D_1 = \mathbb{R}^d$. Next, in Section 2.4.1, we show that in the case $D_2 = (0, \infty) \subset D_1 = R$ the solution is essentially given by Tanaka's formula (Remark 2.4.5), and then we give the proof of the main theorem in the 1-dimensional case (Proposition 2.4.6).

In Section 2.4.2, we first prove the existence of the mirror coupling in the case when D_2 is a half-space in $\mathbb{R}^d$ and $D_1 = \mathbb{R}^d$ (Lemma 2.4.8), and then we use this result in order to prove the existence of the mirror coupling in the case when D_2 is a polygonal domain in $\mathbb{R}^d$ and $D_1 = \mathbb{R}^d$ (Theorem 2.4.9). In Proposition 2.4.10 we present some of the properties of the mirror coupling in the particular case when D_2 is a convex polygonal domain and $D_1 = \mathbb{R}^d$, which are essential for the construction of the general mirror coupling.

In Section 2.5 we give the proof of the main Theorem 2.4.1. The idea of the proof is to construct a sequence (Y_t^n, X_t) of mirror couplings in $\left(D_n, \mathbb{R}^d\right)$, where $D_n \nearrow D_2$ is a sequence of convex polygonal domains in $\mathbb{R}^d$. Then, using the properties of the mirror coupling in convex polygonal domains (Proposition 2.4.10), we show that the sequence Y_t^n converges to a process Y_t, which gives the desired solution to the problem.

The last section (Section 2.6) is devoted to the applications and the extensions of the mirror coupling constructed in Theorem 2.4.1.

First, in Theorem 2.6.3 we use the mirror coupling in order to give a simple, unifying proof of the results of I. Chavel and W. S. Kendall on the domain monotonicity of the Neumann heat kernel (Chavel's Conjecture 2.6.1). The proof is probabilistic in spirit, relying on the geometric properties of the mirror coupling.

Next, in Theorem 2.6.4 we show that Chavel's conjecture also holds in the more general case when one can interpose a ball between the two domains, and the inner domain is star-shaped with respect to the center of the ball (instead of being convex). The analytic proof given here is parallel to the geometric proof of the previous theorem, and it can also serve as an alternate proof of it.

Without giving all the technical details, we discuss the extension of the mirror coupling to the case of smooth bounded domains $D_{1,2} \subset \mathbb{R}^d$ with non-tangential boundaries, such that $D_1 \cap D_2$ is a convex domain.

The section concludes with a discussion of the non-uniqueness of the mirror coupling. The lack of uniqueness is due to the fact that after coupling the processes may decouple, not only on the boundary of the domain, but also when they are inside the domain.

The two basic solutions give rise to the *sticky*, respectively *non-sticky* mirror coupling, and there is a whole range of intermediate possibilities. The stickiness refers to the fact that after coupling the processes "stick" to each other as long as possible ("sticky" mirror coupling, constructed in Theorem 2.4.1), or they can immediately split apart after coupling ("non-sticky" mirror coupling), the general case (*weak/mild* mirror coupling) being a mixture of these two basic behaviors.

We developed the extension of the mirror coupling having in mind the application to Chavel's conjecture, for which the sticky mirror coupling is the "right" tool, but perhaps the other mirror couplings (the non-sticky and the mild mirror couplings) might prove useful in other applications.
To be re-read - perhaps shortened

2.2 Mirror couplings of reflecting Brownian motions

We denote by $\mathbb{U} = \left\{z \in \mathbb{R}^d : \|z\| < 1\right\}$ the open unit ball in $\mathbb{R}^d$ ($d \geq 1$) and by $\nu(z) = -z$, $z \in \partial\mathbb{U}$, the inward unit vector field on the boundary of $\mathbb{U}$.

Given a hyperplane $\mathcal{H} \subset \mathbb{R}^d$, we say that the points $x, y \in \mathbb{R}^d$ are separated by $\mathcal{H}$ if x and y lie in different components of $\mathbb{R}^d - \mathcal{H}$, and we say that they are not separated by $\mathcal{H}$ otherwise.

Reflecting Brownian motion in a smooth domain $D \subset \mathbb{R}^d$ can be defined as a solution of the stochastic differential equation

$$X_t = X_0 + B_t + \int_0^t \nu_D(X_s)dL_s, \qquad t \geq 0, \tag{2.2.1}$$

where B_t is a d-dimensional Brownian motion, ν_D is the inward unit normal vector field on ∂D and L_t^X is the boundary local time of X_t (the continuous non-decreasing process which increases only when $X_t \in \partial D$).

Formally we have:

Definition 2.2.1. *X_t is a reflecting Brownian motion in D starting at $x_0 \in \overline{D}$ if it satisfies (2.2.1), where:*

(a) B_t is a d-dimensional Brownian motion started at 0,

(b) L_t is a continuous nondecreasing process which increases only when $X_t \in \partial D$,

(c) X_t is $(\mathcal{F}_t^B)$-adapted, and almost surely $X_0 = x_0$ and $X_t \in \overline{D}$ for all $t \geq 0$.

Remark 2.2.2. *For pathwise existence and uniqueness of reflecting Brownian motion in the sense of the above definition see for example [14].*

In [4], the authors introduced the *mirror coupling* of reflecting Brownian motion in a smooth domain $D \subset \mathbb{R}^d$ (piecewise C^2 domain in $\mathbb{R}^2$ with a finite number of convex corners or a C^2 domain in $\mathbb{R}^d$, $d \geq 3$).

The idea of the mirror coupling is that the two processes $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ behave like ordinary Brownian motions (symmetric with respect to a hyperplane, called the *mirror* of the coupling) when both of them are inside the domain D. When one of the processes hits the boundary, the mirror $\mathcal{M}_t$ gets a minimal push towards the inward unit normal at the corresponding point at the boundary, needed in order to keep both processes in $\overline{D}$. Considering the coupling time $\tau = \inf\{t > 0 : X_t = Y_t\}$, the mirror coupling evolves as described above for $t \leq \tau$, and we let $X_t = Y_t$ for $t \geq \tau$ (the two processes move together after the coupling time). For definiteness, for $t \geq \tau$ we define the mirror $\mathcal{M}_t$ as the hyperplane parallel to $\mathcal{M}_\tau$ passing through $X_t = Y_t$.

The formal construction of the mirror coupling is the following. Consider the system of stochastic differential equations:

$$X_t = x + W_t + \int_0^t \nu_D (X_s)\, dL_s^X \tag{2.2.2}$$

$$Y_t = y + Z_t + \int_0^t \nu_D (X_s)\, dL_s^Y \tag{2.2.3}$$

$$Z_t = W_t - 2\int_0^t \frac{Y_s - X_s}{||Y_s - X_s||^2} (Y_s - X_s) \cdot dW_s \tag{2.2.4}$$

for $t < \xi$, where $\xi = \inf \{s > 0 : X_s = Y_s\}$ is the coupling time of the processes, after which the processes X and Y evolve together, i.e. $X_t = Y_t$ and $Z_t = W_t + Z_\xi - W_\xi$ for $t \geq \xi$.

Following [4], and considering the Skorokhod map $\Gamma : C\left([0,\infty) : \mathbb{R}^d\right) \to C\left([0,\infty) : \overline{D}\right)$, we have $X = \Gamma(x + W)$, $Y = \Gamma(y + Z)$, and therefore the above system is equivalent to

$$Z_t = \int_0^{t\wedge\xi} G\left(\Gamma(y+Z)_s - \Gamma(x+W)_s\right) dW_s + 1_{t\geq\xi}\left(W_t - W_\xi\right), \tag{2.2.5}$$

where $\xi = \inf\{s > 0 : \Gamma(x+W)_s = \Gamma(y+Z)_s\}$. In [4] the authors proved the pathwise uniqueness and the strong existence of the process Z_t in (2.2.5) (given the Brownian motion W_t).

In the above $G : \mathbb{R}^d \to \mathcal{M}_{d\times d}$ denotes the function defined by

$$G(z) = \begin{cases} H\left(\frac{z}{\|z\|}\right), & \text{if } z \neq 0 \\ 0, & \text{if } z = 0 \end{cases}, \tag{2.2.6}$$

where for a unitary vector $m \in \mathbb{R}^d$, $H(m)$ represents the linear transformation given by the $d \times d$ matrix

$$H(m) = I - 2m\, m', \tag{2.2.7}$$

that is

$$H(m)\, v = v - 2(m \cdot v)\, m \tag{2.2.8}$$

is the mirror image of $v \in \mathbb{R}^d$ with respect to the hyperplane through the origin perpendicular to m (m' denotes the transpose of the vector m, vectors being considered as column vectors).

The pair $(X_t, Y_t)_{t\geq 0}$ constructed above is called a *mirror coupling* of reflecting Brownian motions in D starting at $(x, y) \in \overline{D} \times \overline{D}$.

Remark 2.2.3. *The relation (2.2.4) can be written in the equivalent form*

$$dZ_t = G(X_t - Y_t)\, dW_t,$$

which shows that for $t < \xi$ the increments of Z_t are mirror images of the increments of W_t with respect to the hyperplane $\mathcal{M}_t$ of symmetry between X_t and Y_t, justifying the name of mirror coupling*.*

In the particular case of the unit ball $D = \mathbb{U} \subset \mathbb{R}^d$, for arbitrarily fixed points $x, y \in \overline{\mathbb{U}}$, the mirror coupling of reflecting Brownian motions in the unit ball $\mathbb{U} \subset \mathbb{R}^d$ starting at (x, y) is the pair $(X_t, Y_t)_{t\geq 0}$ of stochastic processes defined by

$$\begin{cases} X_t = x + W_t + \int_0^t \nu(X_s)\, dL_s^X \\ Y_t = y + Z_t + \int_0^t \nu(Y_s)\, dL_s^Y \end{cases}, \tag{2.2.9}$$

where W_t is a d-dimensional Brownian motion starting at $W_0 = 0$, Z_t is the mirror image of the Brownian motion W_t with respect to the hyperplane $\mathcal{M}_t$ of symmetry between X_t and Y_t, that is

$$Z_t = W_t - 2\int_0^t \frac{X_s - Y_s}{\|X_s - Y_s\|^2}\left(X_s - Y_s\right)\cdot dW_s, \tag{2.2.10}$$

and L_t^X and L_t^Y denote the boundary local times of the reflecting Brownian motions X_t and respectively Y_t.

The processes X_t and Y_t evolve according to (2.2.9) above for $t \le \tau$, where τ is the coupling time

$$\tau = \inf\left\{t > 0 : X_t = Y_t\right\} \in \mathbb{R} \cup \{\infty\},$$

and they evolve together after the coupling time (i.e. $X_t = Y_t$ for $t \ge \tau$).

2.3 Laugesen-Morpurgo Conjecture

The Laugesen-Morpurgo conjecture appeared, as we learned from Rodrigo Bañuelos, in connection with their work on conformal extremals of the Riemann zeta function of eigenvalues (see [53]). The conjecture states the diagonal element of the Neumann heat kernel of the Laplacian in the unit disk $U = \left\{x \in \mathbb{R}^2 : |x| < 1\right\}$ in $\mathbb{R}^2$ is a radially increasing function, that is

$$p_U\left(t,x,x\right) < p_U\left(t,y,y\right), \qquad t \ge 0, \tag{2.3.1}$$

for all $x, y \in U$ with $0 \le |x| < |y| \le 1$, where $p_U\left(t,x,y\right)$ denotes the heat kernel for the Laplacian with Neumann boundary conditions (or, equivalently, the transition density for the Brownian motion with normal reflection on the boundary) in the unit disk U. The conjecture extends naturally to the Neumann heat kernel of the Laplacian in the unit ball $\mathbb{U} = \left\{x \in \mathbb{R}^d : ||x|| < 1\right\}$ in $\mathbb{R}^d$, $d \ge 1$, as follows.

Conjecture 2.3.1 (Laugesen-Morpurgo Conjecture). *Let $p_{\mathbb{U}}(t,x,y)$ denote the heat kernel for the Laplacian with Neumann boundary conditions on the unit ball $\mathbb{U} = \left\{z \in \mathbb{R}^d : \|z\| < 1\right\}$ in $\mathbb{R}^d$ ($d \ge 1$).*

For any $t > 0$ we have

$$p_{\mathbb{U}}(t,x,x) < p_{\mathbb{U}}(t,y,y), \tag{2.3.2}$$

for all $x, y \in \overline{\mathbb{U}}$ with $\|x\| < \|y\|$.

The probabilistic interpretation of the conjecture is that a reflecting Brownian motion starting closer to the boundary is more likely to return to its starting position (after t units of time), than a reflecting Brownian motion starting further away from the boundary (after the same t units of time).

The physical interpretation is that introducing an atom of heat in a circular room with thermally insulated boundary, the closer this point to the boundary, the warmer we feel at this point, after any fixed number of units of time.

Despite the seemingly simple nature of this conjecture and the fact that it seems to have been well known since 1994, until 2011 (when we settled this conjecture in [66]) only some partial results were known (see [8], [62], [67] and [68]).

In 2009 Bañuelos et al. ([8]) proved the following result related to the Laugesen-Morpurgo conjecture:

Theorem 2.3.2. *The diagonal element $p_{\mathbb{U}}^B(t,x,x)$ of the transition probabilities for the d-dimensional Bessel processes on $(0;1]$, reflected at 1, is an increasing function of $x \in (0,1]$ for $d > 2$ and this is false for $d = 2$.*

Remark 2.3.3. *Since the norm of a d-dimensional Brownian motion is a Bessel process of order d, the above result is equivalent to the monotonicity with respect to $r \in (0,1)$ (for any $t > 0$ arbitrarily fixed) of the integral mean*

$$r^{d-1} \int_{\partial \mathbb{U}} p_{\mathbb{U}} \left(t, re_1, ru \right) d\sigma(u),$$

where $e_1 = (1, 0, \ldots, 0) \in \mathbb{R}^d$ and σ is the normalized surface measure on $\partial \mathbb{U}$.

The fact that the Laugesen-Morpurgo conjecture is true in the 1-dimensional case is known (see for example [8], Remark 5.4 for an analytic proof, or [68] for a probabilistic proof). In [67] we obtained a discrete version of the Laugesen-Morpurgo conjecture, and as a corollary we derived a new probabilistic proof of the Laugesen-Morpurgo conjecture in the 1-dimensional case, which we will present next.

2.3.1 A discrete version of the conjecture

In this section we will prove a discrete 1-dimensional version of the Laugesen-Morpurgo conjecture, as follows: if X_n is a simple random walk on $\{-s, \ldots, s\}$ with reflecting barriers at $\pm s$, then for any $n \in \mathbb{N}$ arbitrarily fixed, $P^{\,i} \left(X_n = i \right)$ is a strictly increasing function of $|i|$, that is:

$$P^i \left(X_n = i \right) \leq P^j \left(X_n = j \right), \tag{2.3.3}$$

for any $i, j \in \{-s+1, \ldots, s-1\}$ with $|i| < |j|$ and any $n \in \mathbb{N}$.

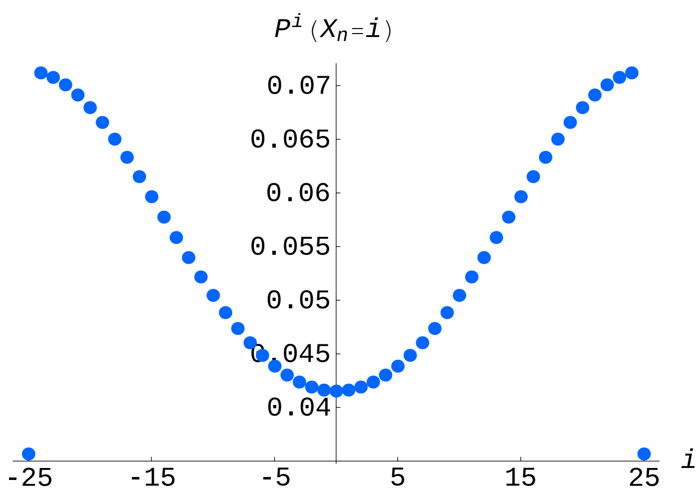

Figure 2.1: The graph of the probabilities $P^i \left(X_n = i \right)$, $i = -s, \ldots, s$ for $s = 25$ and $n = 500$.

It is interesting to note that the inequality (2.3.3) does not hold for $j = s$, as it can be seen from the Figure 2.3.1 above. Also note that when n is odd, (2.3.3) is trivial, since in this case $P^i \left(X_n = i \right) = 0$ for any $i \in \{-s, \ldots, s\}$.

Let $S = \{-s, -s+1, ..., s-1, s\}$, where $s \in \mathbb{N} - \{0\}$. Define new states $s^+ = s^- = s$, $(-s)^+ = (-s)^- = -s$ and let $i^\pm$, $i \in \{-s+1, \ldots, s-1\}$ be distinct, such that

$$S^+ \cap S^- := \left\{ i^+ \mid i \in S \right\} \cap \left\{ i^- \mid i \in S \right\} = \{-s, s\}.$$

Setting $S^\pm = S^+ \cup S^-$ and $S_i = \{i^+, i^-\}$, $i \in S$,

$$S^\pm = S_{-s} \cup S_{-s+1} \cup \ldots \cup S_{s-1} \cup S_s \tag{2.3.4}$$

is a decomposition of $S^\pm$ in disjoint sets.

By a finite cyclic random walk on $S^{\pm}$ (or simply a random walk on $S^{\pm}$), we understand a random walk $(X_n^{\pm})_{n\in\mathbb{N}}$ with state space $S^{\pm}$ and transition matrix $P^{\pm} = (P_{ij}^{\pm})_{i,j\in S^{\pm}}$ given by

$$P^{\pm}_{i^+,(i+1)^+} = P^{\pm}_{i^+,(i-1)^+} = P^{\pm}_{i^-,(i+1)^-} = P^{\pm}_{i^-,(i-1)^-} = \frac{1}{2}, \tag{2.3.5}$$

for $i \in \{-s+1, \ldots, s-1\}$, and

$$P^{\pm}_{-s,(-s+1)^+} = P^{\pm}_{-s,(-s+1)^-} = P^{\pm}_{s,(s-1)^+} = P^{\pm}_{s,(s-1)^-} = \frac{1}{2}. \tag{2.3.6}$$

Given the random walk $(X_n^{\pm})_{n\in\mathbb{N}}$ on $S^{\pm}$, we define a new sequence of random variable $(X_n)_{n\in\mathbb{N}}$ with state space S by setting

$$X_n = i \text{ if and only if } X_n^{\pm} \in \left\{i^+, i^-\right\}, \tag{2.3.7}$$

where $i \in S$ and $n \in \mathbb{N}$.

Remark 2.3.4. *It can be shown (see for example [41], p. 166) that $(X_n^{\pm})_{n\in\mathbb{N}}$ is groupable with respect to the partition (2.3.4) and that $(X_n)_{n\in\mathbb{N}}$ is the corresponding grouped Markov chain, with transition probability matrix P given by*

$$P_{i,i-1} = P_{i,i+1} = \frac{1}{2}, \tag{2.3.8}$$

for $i \in \{-s+1, \ldots, s-1\}$, and

$$P_{-s,-s+1} = P_{s,s-1} = 1. \tag{2.3.9}$$

Remark 2.3.5. *Defining the projection function $pr : S^{\pm} \to S$ by*

$$pr(i^+) = pr\left(i^-\right) = i,$$

for $i \in \{-s+1, \ldots, s-1\}$, and

$$pr(-s) = -s \text{ and } pr\,(s) = s,$$

it can be seen that

$$X_n = pr(X_n^{\pm}), \qquad n \in \mathbb{N}.$$

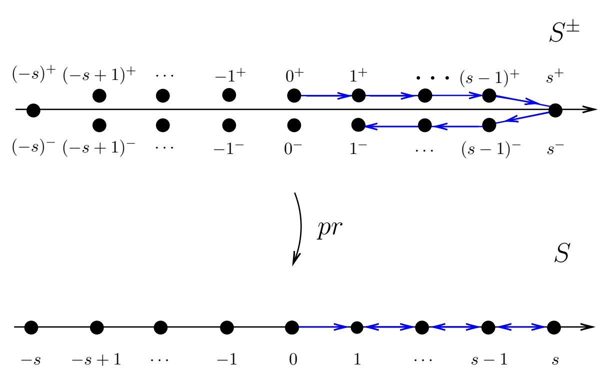

Figure 2.2: The projection of a random walk on $S^{\pm}$ onto a reflecting random walk on S.

Remark 2.3.6. *From (2.3.8) and (2.3.9), it can be seen that $(X_n)_{n\in\mathbb{N}}$ is a random walk on $S = \{-s, \ldots, s\}$, with reflecting barriers at $-s$ and s. We will refer to $(X_n)_{n\in\mathbb{N}}$ as the reflecting random walk on S corresponding to the random walk $(X_n^{\pm})_{n\in\mathbb{N}}$.*

For an arbitrary fixed starting point $X_0^{\pm} = x \in S^{\pm}$, we denote by $\mathbf{P}^x$ the probability measure associated with the random walk $(X_n^{\pm})_{n\in\mathbb{N}}$ and by $\mathbf{P}^{pr(x)}$ the probability measure associated with the corresponding reflecting random walk $(X_n)_{n\in\mathbb{N}}$.

The relationship between the transition probabilities of a random walk on $S^{\pm}$ and those of the corresponding reflecting random walk on S is given by the following.

Proposition 2.3.7. *For any $i \in S - \{-s, s\}$ and $n \in \mathbb{N}$ we have*

$$\mathbf{P}^i(X_n = i) = \mathbf{P}^{i^+}(X_n^{\pm} = i^+) + \mathbf{P}^{i^+}(X_n^{\pm} = i^-),$$

where $(X_n^{\pm})_{n\in\mathbb{N}}$ is a random walk on $S^{\pm}$ and $(X_n)_{n\in\mathbb{N}}$ is the corresponding reflecting random walk on S.

Proof. See [67]. □

Remark 2.3.8. *Alternately, letting $U_{4s} = \{\exp(\frac{ik\pi}{2s}) : k \in \{0, 1, ..., 4s-1\}\}$ denote the vertices of a regular polygon with $4s$ sides and defining the bijection $f : S^{\pm} \to U_{4s}$ by*

$$f(k^+) = \exp\left((s-k)\frac{i\pi}{2s}\right), \qquad k \in S,$$

and

$$f(k^-) = \exp\left((3s+k)\frac{i\pi}{2s}\right), \qquad k \in S,$$

we can view a random walk on $S^{\pm} = \{-s, \dots, s\}$ as a rotationally invariant random walk on the vertices of the polygon U_{4s} (see Figure 2.3).

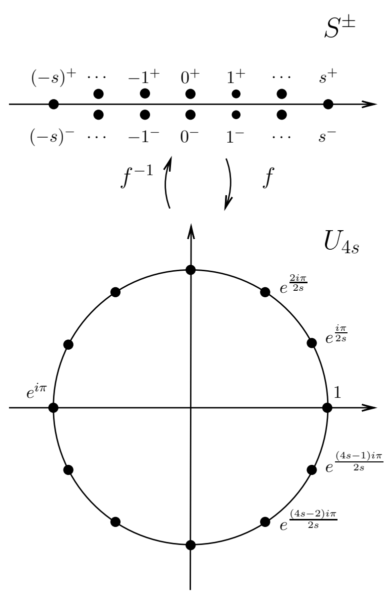

Figure 2.3: The bijective correspondence between random walks on $S^{\pm}$ and U_{4s}.

We introduce two couplings of random walks on $S^{\pm}$: the *translation coupling* and the *mirror coupling*.

A *translation coupling* of random walks on $S^{\pm}$ with starting points $(x, y) \in S^{\pm} \times S^{\pm}$ (without loss of generality we may assume that $pr(x) < pr(y)$) is a pair $(X_n^{\pm}, Y_n^{\pm})_{n\in\mathbb{N}}$, where $(X_n^{\pm})_{n\in\mathbb{N}}$ is a random walk on $S^{\pm}$ with starting point $x \in S^{\pm}$, and $(Y_n^{\pm})_{n\in\mathbb{N}}$ is the random walk on $S^{\pm}$ with starting point $y \in S^{\pm}$ defined by

$$Y_n^{\pm} = tr_{pr(y)-pr(x)}(X_n^{\pm}), \qquad n \in \mathbb{N},$$

where for $a \in \mathbb{Z}$, $tr_a : S^{\pm} \to S^{\pm}$ is the translation function on $S^{\pm}$ defined by

$$tr_a(t) = f^{-1}\left(f(t)\exp\left(-i\frac{a\pi}{2s}\right)\right), \qquad t \in S^{\pm}, \tag{2.3.10}$$

and f is the bijection defined in Remark 2.3.8 (see Figure 2.4).

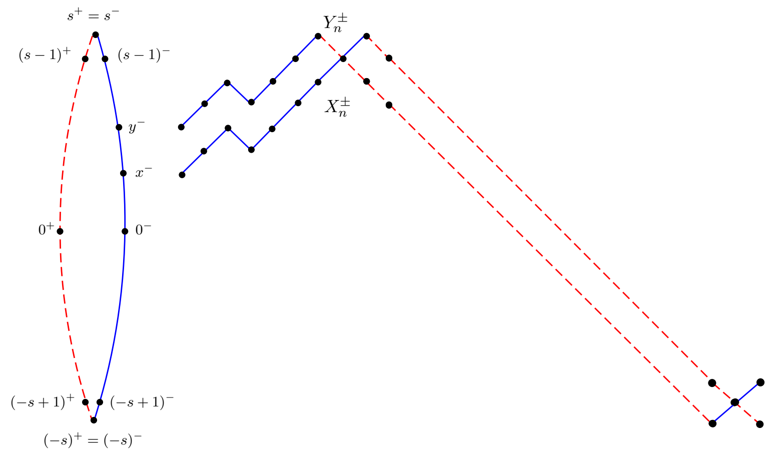

Figure 2.4: Sample paths of random walks on $S^{\pm}$ coupled by translation.

A *mirror coupling* of random walks on $S^{\pm}$ with starting points $(x, y) \in S^- \times S^-$ chosen such that $pr(x)+pr(y)$ is an even number (without loss of generality we may assume that $pr(x) < pr(y)$) is the pair $(X_n^{\pm}, Y_n^{\pm})_{n\in\mathbb{N}}$, where $(X_n^{\pm})_{n\in\mathbb{N}}$ is a random walk on $S^{\pm}$ with starting point $x \in S^{\pm}$, and $(Y_n^{\pm})_{n\in\mathbb{N}}$ is the random walk on $S^{\pm}$ with starting point $y \in S^{\pm}$ defined by

$$Y_n^{\pm} = \begin{cases} sym_{\left(\frac{pr(x)+pr(y)}{2}\right)^-}(X_n^{\pm}) & , \quad n \le \tau \\ X_n^{\pm} \quad , & \quad n > \tau \end{cases},$$

where for $a \in S^{\pm}$, $sym_a : S^{\pm} \to S^{\pm}$ is the symmetry function on $S^{\pm}$ with respect to the line passing through the origin and the point $f(a)$, defined by

$$sym_a(t) = f^{-1}\left(\overline{f(t)} f^2(a)\right), \qquad t \in S^{\pm}, \tag{2.3.11}$$

τ is the coupling time defined by

$$\tau = \inf\left\{n \ge 0 : X_n^{\pm} = \left(\frac{pr(x)+pr(y)}{2}\right)^- \text{ or } X_n^{\pm} = \left(-\frac{pr(x)+pr(y)}{2}\right)^+\right\},$$

and f is the bijection defined in Remark 2.3.8 (see Figure 2.5).

As an application of the translation coupling, we obtain the following.

Proposition 2.3.9. *For any $i, j \in S$ with $0 \le i \le j$ and any $n \in \mathbb{N}$ we have*

$$\mathbf{P}^{i^+}(X_n^{\pm} = i^+) + \mathbf{P}^{i^+}(X_n^{\pm} = i^-) = \mathbf{P}^{j^+}(X_n^{\pm} = j^+) + \mathbf{P}^{j^+}(X_n^{\pm} = (2i-j)^-), \tag{2.3.12}$$

where $X_n^{\pm}$ is a random walk on $S^{\pm}$.

Proof. See [67]. □

As an application of the mirror coupling we obtain the following.

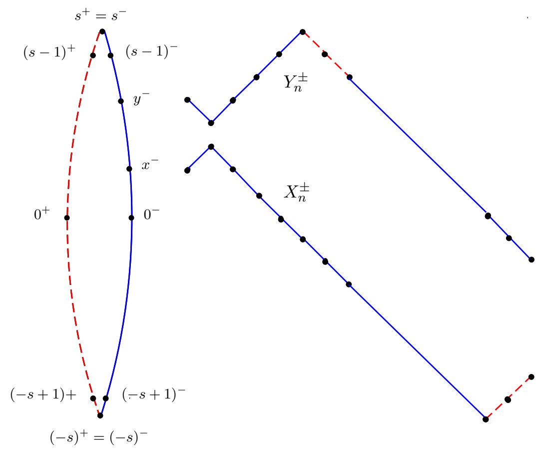

Figure 2.5: Sample paths of random walks on $S^\pm$ coupled by mirror coupling.

Proposition 2.3.10. *For any $i, j \in S$ with $0 \le i < j$ such that $i + j$ is an even number and any $n \in \mathbb{N}$, we have*

$$\mathbf{P}^{i^-}(X_n^\pm = j^+) < \mathbf{P}^{j^-}(X_n^\pm = j^+),$$

where $X_n^\pm$ is a random walk on $S^\pm$.

Proof. See [67]. □

With this preparation, we are now ready to prove the main result of this section, as follows.

Theorem 2.3.11. *For any $s \in \mathbb{N} - \{0\}$ and $n \in \mathbb{N}$, $P^i(X_n = i)$ is a strictly increasing function of $i \in \{0, \ldots, s-1\}$, that is, for any $i, j \in \{0, \ldots, s-1\}$, with $i < j$ and any $n \in \mathbb{N}$, we have*

$$\mathbf{P}^i(X_n = i) < \mathbf{P}^j(X_n = j),$$

where X_n is a reflecting random walk on S with reflecting barriers at $\pm s$.

Proof. Let $X_n^\pm$ be a random walk on $S^\pm$ and let $X_n = pr(X_n^\pm)$. By Proposition 2.3.7, for $i \in S - \{-s, s\}$ and $n \in \mathbb{N}$ we have

$$\mathbf{P}^i(X_n = i) = \mathbf{P}^{i^+}(X_n^\pm = i^+) + \mathbf{P}^{i^+}(X_n^\pm = i^-).$$

Consider first the case $i \in \{1, ..., s-2\}$. Using Proposition 2.3.9 with j replaced by $i+1$, we have

$$\begin{aligned}
\mathbf{P}^i(X_n = i) &= \mathbf{P}^{i^+}(X_n^\pm = i^+) + \mathbf{P}^{i^+}(X_n^\pm = i^-) \\
&= \mathbf{P}^{(i+1)^+}(X_n^\pm = (i+1)^+) + \mathbf{P}^{(i+1)^+}(X_n^\pm = (i-1)^-) \\
&< \mathbf{P}^{(i+1)^+}(X_n^\pm = (i+1)^+) + \mathbf{P}^{(i+1)^+}(X_n^\pm = (i+1)^-),
\end{aligned}$$

for any $n \in \mathbb{N}$, where the last inequality follows from the symmetry of the transition matrix of a random walk on $S^\pm$ and Proposition 2.3.10 with j replaced by $i+1$ and i replaced by $i-1$. Using again Proposition 2.3.7 we obtain

$$\mathbf{P}^i(X_n = i) < \mathbf{P}^{i+1}(X_n = i+1),$$

for any $n \in \mathbb{N}$ and $i \in \{1, ..., s-2\}$.

To conclude the proof, we need to show that the previous inequality also holds for $i = 0$. In this case, using an argument similar to the one in the proofs of Proposition 2.3.9 and Proposition 2.3.10 it can be shown that we have

$$\begin{aligned} \mathbf{P}^{0^+}(X_n^{\pm} = 0^+) + \mathbf{P}^{0^+}(X_n^{\pm} = 0^-) &= \mathbf{P}^{1^+}(X_n^{\pm} = 1^+) + \mathbf{P}^{1^+}(X_n^{\pm} = (-1)^-) \\ &< \mathbf{P}^{1^+}(X_n^{\pm} = 1^+) + \mathbf{P}^{1^+}(X_n^{\pm} = 1^-), \end{aligned}$$

which by Proposition 2.3.7 shows that

$$\mathbf{P}^0(X_n = 0) < \mathbf{P}^1(X_n = 1),$$

for any $n \in \mathbb{N}$, concluding the proof. □

Using similar techniques, we can generalize the result in Theorem 2.3.11 as follows.

Theorem 2.3.12. *For any $s, n \in \mathbb{N} - \{0\}$, $i, j \in \{0, \ldots, s-1\}$ with $i < j$ and any $k \in S$ with $|i+k| < s$, $|j+k| < s$ and $k > -\min\{i+j, 2i-j+s\}$, we have*

$$\mathbf{P}^i(X_n = i+k) < \mathbf{P}^j(X_n = j+k),$$

where X_n is a reflecting random walk on $S = \{-s, \ldots, s\}$ with reflecting barriers at $\pm s$.

Proof. See [67]. □

From the previous theorem we obtain the following.

Corollary 2.3.13. *For any $s, n \in \mathbb{N} - \{0\}$, $i, j \in \{0, \ldots, s-1\}$ with $i < j$ and any $k \in \{1, \ldots, s\}$ with $k \le \min\{s-j, i+j\}$, we have*

$$\mathbf{P}^i(|X_n - i| < k) < \mathbf{P}^j(|X_n - j| < k),$$

where X_n is a reflecting random walk on $S = \{-s, \ldots, s\}$ with reflecting barriers at $\pm s$.

Proof. Follows from the previous theorem by summing over k. □

Using the above corollary and the fact that the reflecting Brownian motion can be approximated by random walks (see for example [24]), we obtain the following.

Corollary 2.3.14. *For any $x, y, \varepsilon \in (0,1)$ with $x < y$ and $\varepsilon < \min\{1-y, x+y\}$, we have*

$$\mathbf{P}^x\left(B_t \in (x-\varepsilon, x+\varepsilon)\right) \le \mathbf{P}^y\left(B_t \in (y-\varepsilon, y+\varepsilon)\right), \qquad t > 0,$$

where B_t is a 1-dimensional reflecting Brownian motion on $[-1,1]$.

Proof. In the case when both x and y are dyadic rationals in $[0,1]$, the proof follows from the previous corollary by using the fact that a reflecting Brownian motion B_t on $[-1,1]$ starting at $B_0 = \frac{k}{2^l}$ ($k, l \in \mathbb{Z}$, $l > 0$) can be approximated by a reflecting random walk $(X_m^n)_{m \ge 0}$, more precisely it can be shown (see for example [24]) that we have

$$X^n_{[2^{2n}t]} \underset{n\to\infty}{\to} B_t, \qquad t \ge 0,$$

where $(X_m^n)_{m \ge 0}$ is a (simple) reflecting random walk on

$$\left\{-1, -\frac{2^n-1}{2^n}, -\frac{2^n-2}{2^n}, \ldots, \frac{2^n-2}{2^n}, \frac{2^n-1}{2^n}, 1\right\}$$

with reflecting barriers at ± 1 and starting at $X_0^n = \frac{k}{2^l}$.

The general case follows by approximating x and y by dyadic rationals and using the previous part of the proof. □

As a corollary, we obtain the proof of the Laugesen-Morpurgo conjecture in the 1-dimensional case, as follows.

Corollary 2.3.15. *For any $x, y \in (0,1)$ with $x < y$, we have*

$$p\left(t,x,x\right) < p\left(t,y,y\right), \qquad t>0,$$

where $p(t,x,y)$ denotes the transition probabilities of the 1-dimensional reflecting Brownian motion on $[-1,1]$.

Proof. The fact that $p\left(t,x,x\right)$ is increasing in $x \in (0,1)$ follows from the previous corollary, using the fact that by the continuity of $p\left(t,x,y\right)$ in the second variable we have

$$p\left(t,x,x\right) = \lim_{\varepsilon\searrow 0} \frac{1}{2\varepsilon}\int_{x-\varepsilon}^{x+\varepsilon} p\left(t,x,y\right)dy = \lim_{\varepsilon\searrow 0}\frac{1}{2\varepsilon}P^x\left(|B_t - x| < \varepsilon\right).$$

To show that $p\left(t,x,x\right)$ is in fact strictly increasing, note that since $p\left(t,x,x\right)$ is a real analytic function of $x \in (-1,1)$ for any $t>0$ arbitrarily fixed (it is the diagonal of the heat kernel of an operator with real analytic coefficients), it cannot be constant on a nonempty open subset of $[-1,1]$ unless it is identically constant on the entire interval $[-1,1]$.

It can be shown (see for example [8]) that for any $t>0$ arbitrarily fixed we have

$$p\left(t,x,x\right) + \widetilde{p}\left(t,x,x\right) = c, \qquad x \in (-1,1),$$

where c is a constant depending on $t>0$ and $\widetilde{p}\left(t,x,y\right)$ denotes the transition density of Brownian motion on $(-1,1)$ killed on hitting the boundary of the interval.

If $p\left(t,x,x\right)$ were constant in $x \in (-1,1)$ for an arbitrarily fixed $t>0$, then $\widetilde{p}\left(t,x,x\right)$ would also be constant in $x\in(-1,1)$. However, this leads to a contradiction, since

$$\lim_{x\nearrow 1}\widetilde{p}\left(t,x,x\right) = 0 < \widetilde{p}\left(t,0,0\right).$$

This, together with the fact that $p\left(t,x,x\right)$ is increasing in $x\in(0,1)$ for any $t>0$ arbitrarily fixed, shows that $p\left(t,x,x\right)$ is in fact strictly increasing in $x \in (0,1)$, concluding the proof. □

2.3.2 The resolution of the conjecture

Our proof of Laugesen-Morpurgo conjecture relies on a certain property of the mirror coupling of reflecting Brownian motions in the unit disk and a representation of the Neumann heat kernel as an occupation time density of reflecting Brownian motion. We begin with a presentation of these results.

The key for proving the Laugesen-Morpurgo conjecture (Conjecture 2.3.1) is the double inequality (2.3.22) in Theorem 6.2.4, which in turn relies on proving the following inequality:

$$p_{\mathbb{U}}(t,y,z) \le p_{\mathbb{U}}(t,x,z), \qquad t>0, \tag{2.3.13}$$

for all $x,y,z\in\mathbb{U}$ satisfying $\|x-z\|\le\|y-z\|$ and $\|y\|\le\|x\|$.

Consider a mirror coupling X_t, Y_t of reflecting Brownian motions in $\mathbb{U}$ given by (2.2.9) – (2.2.10), with starting points $X_0 = x, Y_0 = y \in \overline{\mathbb{U}}$.

For $t<\tau=\inf\{t>0 : X_t = Y_t\}$, the mirror $\mathcal{M}_t$ of the coupling (the hyperplane of symmetry between X_t and Y_t) is given by

$$\mathcal{M}_t = \left\{ z\in\mathbb{R}^n : \left(z - \frac{X_t+Y_t}{2}\right)\cdot\left(X_t - Y_t\right) = 0\right\}. \tag{2.3.14}$$

The idea for proving the inequality (2.3.13) is that the mirror $\mathcal{M}_t$ moves towards the origin, in the sense of Lemma 2.3.16 below. This property is a rigorous version of Example 4.5 in [25], used by the authors to prove the efficiency of the mirror coupling in the case of the unit disk.

Lemma 2.3.16. *Let X_t, Y_t be a mirror coupling of reflecting Brownian motions in $\mathbb{U}$ with starting points $X_0 = x$, $Y_0 = y \in \overline{\mathbb{U}}$, and let $\tau = \inf\{t > 0 : X_t = Y_t\}$ be the coupling time and $\tau_1 = \inf\{t > 0 : 0 \in \mathcal{M}_t\}$.*

For all times $t < \tau \wedge \tau_1$, the mirror $\mathcal{M}_t$ moves towards the origin, in such a way that if a point $P \in \mathbb{U}$ and the origin are separated by $\mathcal{M}_{t_1}$ for some $t_1 \in [0, \tau \wedge \tau_1)$, then the point P and the origin are separated by $\mathcal{M}_{t_2}$ for all $t_2 \in [t_1, \tau \wedge \tau_1)$ (see Figure 2.6).

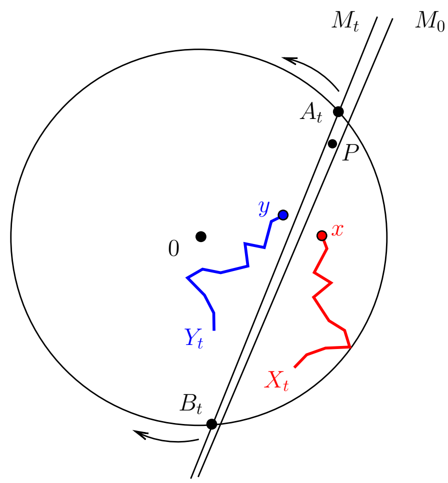

Figure 2.6: Mirror coupling of reflecting Brownian motions in the unit disk (the case $d = 2$).

Proof. If $\|x\| = \|y\|$, then $\tau_1 = 0$ and there is nothing to prove in this case (the mirror $\mathcal{M}_0$ passes through the origin). Without loss of generality we may therefore assume that $\|x\| > \|y\|$.

Setting

$$\begin{cases} U_t = X_t - Y_t \\ V_t = X_t + Y_t \end{cases}, \qquad t \geq 0, \tag{2.3.15}$$

from the definition (2.2.9) – (2.2.10) of the mirror coupling we obtain:

$$\begin{cases} U_t^i = x^i - y^i + W_t^i - Z_t^i - \int_0^t X_s^i dL_s^X + \int_0^t Y_s^i dL_s^Y \\ V_t^i = x^i + y^i + W_t^i + Z_t^i - \int_0^t X_s^i dL_s^X - \int_0^t Y_s^i dL_s^Y \end{cases}, \qquad i = 1, \ldots, d,$$

for all $t \leq \tau$, where the superscript i indicates the i^{th} cartesian coordinate of the given point.

Using the definition (2.2.10) of Z_t, we have

$$\begin{cases} U_t^i = x^i - y^i + 2\int_0^t \frac{U_s^i}{\|U_s\|^2} U_s \cdot dW_s - \int_0^t X_s^i dL_s^X + \int_0^t Y_s^i dL_s^Y \\ V_t^i = x^i + y^i + 2W_t^i - 2\int_0^t \frac{U_s^i}{\|U_s\|^2} U_s \cdot dW_s - \int_0^t X_s^i dL_s^X - \int_0^t Y_s^i dL_s^Y \end{cases}, \tag{2.3.16}$$

for all $i = 1, \ldots, d$ and $t < \tau$, and therefore we obtain the following formulae for the quadratic variation of the processes U and V:

$$\begin{cases} \langle U^i, U^j \rangle_t = 4\int_0^t \frac{U_s^i U_s^j}{\|U_s\|^2} ds \\ \langle V^i, V^j \rangle_t = 4\int_0^t \delta_{ij} - \frac{U_s^i U_s^j}{\|U_s\|^2} ds \\ \langle U^i, V^j \rangle_t = 0 \end{cases}, \qquad i, j, = 1, \ldots, d. \tag{2.3.17}$$

Note that since $\|X_0\| = \|x\| > \|y\| = \|Y_0\|$, it follows that for all $t < \tau \wedge \tau_1$ we have

$$U_t \cdot V_t = (X_t - Y_t) \cdot (X_t + Y_t) = \|X_t\|^2 - \|Y_t\|^2 > 0, \tag{2.3.18}$$

and therefore for $t < \tau \wedge \tau_1$ we may define the process A_t by

$$A_t = \frac{2}{U_t \cdot V_t} U_t. \tag{2.3.19}$$

We will first show that for $t < \tau \wedge \tau_1$ the components of the process A_t are processes of bounded variation, satisfying

$$dA_t^i = \frac{2}{U_t \cdot V_t} \left(A_t^i - \frac{U_t^i + V_t^i}{2} \right) dL_t^X, \qquad i = 1, \ldots, d. \tag{2.3.20}$$

Applying the Itô formula to the C^2 function $f(u, v) = \frac{u^i}{u \cdot v}$ and to the processes U_t and V_t, we have:

$$\begin{aligned}
\frac{1}{2} dA_t^i &= d\left(\frac{U_t^i}{U_t \cdot V_t} \right) \\
&= \frac{1}{(U_t \cdot V_t)^2} \sum_{j=1}^{d} \left(\left(\delta_{ij}\, U_t \cdot V_t - U_t^i V_t^j \right) dU_t^j - U_t^i U_t^j dV_t^j \right) \\
&\quad + \frac{1}{2(U_t \cdot V_t)^3} \sum_{j,k=1}^{d} \left(2U_t^i V_t^j V_t^k - \delta_{ij} V_t^k U_t \cdot V_t - \delta_{ik} V_t^j U_t \cdot V_t \right) d\langle U^j, U^k \rangle_t \\
&\quad + \frac{1}{2(U_t \cdot V_t)^3} \sum_{j,k=1}^{d} \left(2U_t^i U_t^j U_t^k \right) d\langle V^j, V^k \rangle_t \\
&\quad + \frac{1}{(U_t \cdot V_t)^3} \sum_{j,k=1}^{d} \left(2U_t^i U_t^k V_t^j - \delta_{ij} U_t^k U_t \cdot V_t - \delta_{jk} U_t^i U_t \cdot V_t \right) d\langle U^j, V^k \rangle_t.
\end{aligned}$$

Using the relations in (2.3.16) it can be seen that the martingale part in the above expression reduces to zero, and combining with (2.3.17) we obtain

$$\begin{aligned}
\frac{1}{2} dA_t^i &= \frac{1}{(U_t \cdot V_t)^2} \sum_{j=1}^{d} \left(\left(\delta_{ij}\, U_t \cdot V_t - U_t^i V_t^j \right) \left(-X_t^j dL_t^X + Y_t^j dL_t^Y \right) \right) \\
&\quad - \frac{1}{(U_t \cdot V_t)^2} \sum_{j=1}^{d} \left(U_t^i U_t^j \left(-X_t^j dL_t^X - Y_t^j dL_t^Y \right) \right) \\
&\quad + \frac{1}{2(U_t \cdot V_t)^3} \sum_{j,k=1}^{d} \left(2U_t^i V_t^j V_t^k - \delta_{ij} V_t^k U_t \cdot V_t - \delta_{ik} V_t^j U_t \cdot V_t \right) 4 \frac{U_t^j U_t^k}{\|U_t\|^2} dt \\
&\quad + \frac{1}{2(U_t \cdot V_t)^3} \sum_{j,k=1}^{d} \left(2U_t^i U_t^j U_t^k \right) 4 \left(\delta_{jk} - \frac{U_t^j U_t^k}{\|U_t\|^2} \right) dt \\
&= \frac{1}{(U_t \cdot V_t)^2} \sum_{j=1}^{d} \left(U_t^i U_t^j + U_t^i V_t^j - \delta_{ij}\, U_t \cdot V_t \right) X_t^j dL_t^X \\
&\quad + \frac{1}{(U_t \cdot V_t)^2} \sum_{j=1}^{d} \left(U_t^i U_t^j - U_t^i V_t^j + \delta_{ij}\, U_t \cdot V_t \right) Y_t^j dL_t^Y.
\end{aligned}$$

Using the fact that $L_t^Y \equiv 0$ on the time interval $[0, \tau \wedge \tau_1)$ (the process Y_t cannot reach the boundary $\partial \mathbb{U}$ before either coupling first with X_t or before the first time when $\|X_t\| = \|Y_t\| = 1$, that is before $0 \in \mathcal{M}_t$), and that L_t^X increases only when $X_t \in \partial\mathbb{U}$, that is only when $\|X_t\| = \|\frac{U_t+V_t}{2}\| = 1$, we obtain:

$$\begin{aligned}
\frac{1}{2}dA_t^i &= \frac{1}{(U_t\cdot V_t)^2}\sum_{j=1}^d \left(U_t^iU_t^j + U_t^iV_t^j - \delta_{ij}\,U_t\cdot V_t\right)X_t^j dL_t^X \\
&= \frac{U_t^i}{2(U_t\cdot V_t)^2}\sum_{j=1}^d \left(U_t^j + V_t^j\right)^2 dL_t^X - \frac{U_t^i+V_t^i}{2U_t\cdot V_t}dL_t^X \\
&= \frac{U_t^i}{2(U_t\cdot V_t)^2}\|U_t+V_t\|^2 dL_t^X - \frac{U_t^i+V_t^i}{2U_t\cdot V_t}dL_t^X \\
&= \frac{2U_t^i}{(U_t\cdot V_t)^2}dL_t^X - \frac{U_t^i+V_t^i}{2U_t\cdot V_t}dL_t^X \\
&= \frac{1}{U_t\cdot V_t}\left(A_t^i - \frac{U_t^i+V_t^i}{2}\right)dL_t^X,
\end{aligned}$$

thus proving the claim (2.3.20).

To prove the claim of the lemma, assume by contradiction that there exists a point $P \in \mathbb{U}$ and times $0 < t_1 < t_2 < \tau\wedge\tau_1$ such that the point P and the origin are separated by $\mathcal{M}_{t_1}$, but are not separated by $\mathcal{M}_{t_2}$. By eventually changing the point P, without loss of generality we may assume that $P \notin \mathcal{M}_{t_2}$, and using (2.3.14) and (2.3.15) we obtain:

$$P\cdot U_{t_2} - \frac{1}{2}U_{t_2}\cdot V_{t_2} < 0 < P\cdot U_{t_1} - \frac{1}{2}U_{t_1}\cdot V_{t_1},$$

or equivalently (recall the definition (2.3.19) of the process A_t and that $U_t\cdot V_t > 0$ for $t \in [0, \tau\wedge\tau_1)$)

$$P\cdot A_{t_2} < 1 < P\cdot A_{t_1}.$$

Setting $t_0 = \inf\{t > t_1 : P\cdot A_t < 1\} \in (t_1, t_2)$ and using (2.3.20), we obtain:

$$\begin{aligned}
P\cdot A_{t_0} &= P\cdot A_{t_1} + \int_{t_1}^{t_0} P\cdot dA_t \\
&= P\cdot A_{t_1} + \int_{t_1}^{t_0}\frac{2}{U_t\cdot V_t}\left(P\cdot A_t - \frac{1}{2}P\cdot(U_t+V_t)\right)dL_t^X \\
&\geq P\cdot A_{t_1} \\
&> 1,
\end{aligned}$$

since $P\cdot A_t \geq 1$ for $t \in [t_1, t_0]$ and

$$\left|\frac{1}{2}P\cdot(U_t+V_t)\right| = |P\cdot X_t| \leq \|P\|\,\|X_t\| \leq 1.$$

By the continuity of the process A_t and the choice of t_0 we must also have $P\cdot A_{t_0} = 1$, contradiction which concludes the proof of the lemma. □

From the previous lemma we obtain the following:

Theorem 2.3.17. *For any points $x, y \in \overline{\mathbb{U}}$ with $\|y\| \leq \|x\|$ and any $z \in \overline{\mathbb{U}}$ such that $\|x - z\| \leq \|y - z\|$, we have:*

$$P^y\left(\|Y_t - z\| < \varepsilon\right) \leq P^x\left(\|X_t - z\| < \varepsilon\right), \tag{2.3.21}$$

for any $t \geq 0$ and $\varepsilon \in (0, \min\{\|z\|, 1 - \|z\|\})$, where X_t and Y_t are reflecting Brownian motions in $\mathbb{U}$ starting at x, respectively y, and P^x, P^y denote the corresponding probability measures.

Proof. Without loss of generality we may assume that x and y are distinct points.

Let X_t, Y_t be a mirror coupling of reflecting Brownian motions in $\mathbb{U}$ with starting points $X_0 = x$ and $Y_0 = y$, and let τ be the coupling time and $\tau_1 = \inf\{t > 0 : 0 \in \mathcal{M}_t\}$.

If $\mathcal{M}_t$ separates X_t and z for some $t < \tau \wedge \tau_1$, there exists a point $P \in \mathbb{U}$ such that the origin and the point P are separated by $\mathcal{M}_0$, but are not separated by $\mathcal{M}_t$, contradicting Lemma 2.3.16. It follows that the mirror $\mathcal{M}_t$ does not separate the points X_t and z for all $t < \tau \wedge \tau_1$, and therefore the distance from X_t to z is not greater than the distance from Y_t to z in this case.

Since for $t \geq \tau \wedge \tau_1$, either the processes X_t and Y_t are symmetric with respect to the (fixed) hyperplane $\mathcal{M}_{\tau\wedge\tau_1}$ passing through the origin (for $t \in (\tau \wedge \tau_1, \tau)$), or they have coupled (for $t \in (\tau, \infty)$), it follows that the distance from X_t to z is also not greater than the distance from Y_t to z.

In all cases we obtained that the distance from X_t to z is not greater than the distance from Y_t to z, and the claim follows. $\square$

Denoting by $p_{\mathbb{U}}(t, x, y)$ the heat kernel for the Laplacian with Neumann boundary conditions on the unit ball $\mathbb{U} \subset \mathbb{R}^d$ (or equivalently, the transition density of reflecting Brownian motion in $\mathbb{U}$), we can now prove the following double inequality:

Theorem 2.3.18. *For any $x \in \mathbb{U} - \{0\}$, $r \in (0, \min\{\|x\|, 1 - \|x\|\})$ and $t > 0$ we have:*

$$\int_{\partial\mathbb{U}} p_{\mathbb{U}}\left(t, x + ru, x\right) d\sigma(u) \leq p_{\mathbb{U}}(t, x + r\tfrac{x}{\|x\|}, x) \leq p_{\mathbb{U}}(t, x + r\tfrac{x}{\|x\|}, x + r\tfrac{x}{\|x\|}), \tag{2.3.22}$$

where σ is the normalized surface measure on $\partial\mathbb{U}$.

Proof. Using the continuity of the transition density $p_{\mathbb{U}}(t, x, y)$ of reflecting Brownian motion in the space variable, it follows $p_{\mathbb{U}}(t, x, y)$ can be written as

$$p_{\mathbb{U}}(t, x, y) = \lim_{\varepsilon \searrow 0} \frac{1}{c_d \varepsilon^d} \int_{\|y-z\|<\varepsilon} p_{\mathbb{U}}(t, x, z) dz = \lim_{\varepsilon \searrow 0} \frac{1}{c_d \varepsilon^d} P^x\left(\|W_t - y\| < \varepsilon\right), \tag{2.3.23}$$

where W_t is a reflecting Brownian motion in the unit ball $\mathbb{U} \subset \mathbb{R}^d$ starting at $W_0 = x$, P^x denotes the corresponding probability measure and $c_d = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)}$ is the volume of the unit ball $\mathbb{U} \subset \mathbb{R}^d$.

For $u \in \mathbb{U}$ fixed, it is easy to see that the hypotheses of Theorem 2.3.17 are verified if we replace x, y and z respectively by $x + r\frac{x}{\|x\|}$, $x + ru$ and x. From this theorem, and combining with the above representation, we obtain

$$p_{\mathbb{U}}\left(t, x + ru, x\right) \leq p_{\mathbb{U}}(t, x + r\tfrac{x}{\|x\|}, x), \qquad u \in \mathbb{U}, \tag{2.3.24}$$

and integrating with respect to $u \in \mathbb{U}$ we obtain the left inequality in (2.3.22).

The right inequality in (2.3.22) can be proved similarly, replacing x, y and z in Theorem 2.3.17 respectively by $x + r\frac{x}{\|x\|}$, x and $x + r\frac{x}{\|x\|}$, and using the symmetry of $p_{\mathbb{U}}(t, x, y)$ in $x, y \in \mathbb{U}$. $\square$

Remark 2.3.19. *The inequality (2.3.24) obtained in the previous proof might be of independent interest, and can be interpreted as an extremal property of reflecting Brownian motion in the unit ball $\mathbb{U}$, as follows:*

$$\max_{y \in \mathbb{U} : \|y-x\|=r} p_{\mathbb{U}}\left(t, x, y\right) = p_{\mathbb{U}}(t, x, x + r\tfrac{x}{\|x\|}),$$

that is, among all reflecting Brownian motions in the unit ball $\mathbb{U}$ with starting points on the sphere $\{y \in \mathbb{R}^n : \|y - x\| = r\}$, the reflecting Brownian motion starting closest to the boundary of $\mathbb{U}$ (i.e. at the point $x + r\frac{x}{\|x\|}$) is most likely to return to (a neighborhood of) x. This extremal property of reflecting Brownian motion is the key of our proof of the Laugesen-Morpurgo conjecture.

As a corollary of the above theorem, we obtain the following resolution of the Laugesen-Morpurgo conjecture:

Theorem 2.3.20. *Let $p_{\mathbb{U}}(t,x,y)$ denote the heat kernel for the Laplacian with Neumann boundary conditions on the unit ball $\mathbb{U}=\left\{z\in\mathbb{R}^d:\|z\|<1\right\}$ in $\mathbb{R}^d$ $(d\geq 1)$.*

For any $t>0$ we have

$$p_{\mathbb{U}}(t,x,x)<p_{\mathbb{U}}(t,y,y), \tag{2.3.25}$$

for all $x,y\in\overline{\mathbb{U}}$ with $\|x\|<\|y\|$.

Proof. First note that for $t>0$ fixed, by the radial symmetry of the problem it follows that $p_{\mathbb{U}}(t,x,x)$ is a function of $\|x\|\in[0,1]$.

For an arbitrarily fixed $x\in\mathbb{U}-\{0\}$, from Theorem 6.2.4 we obtain

$$\begin{aligned}
& p_{\mathbb{U}}(t,x+r\tfrac{x}{\|x\|},x+r\tfrac{x}{\|x\|})-p_U\left(t,x,x\right)\\
\geq\ & \int_{\partial\mathbb{U}}p_{\mathbb{U}}\left(t,x+ru,x\right)d\sigma(u)-p_{\mathbb{U}}\left(t,x,x\right)\\
=\ & \int_{\partial\mathbb{U}}\left(p_{\mathbb{U}}\left(t,x+ru,x\right)-p_U\left(t,x,x\right)\right)d\sigma(u),
\end{aligned}$$

for any $r\in(0,\min\{\|x\|,1-\|x\|\})$. Dividing by r and passing to the limit with $r\searrow 0$, we obtain:

$$\begin{aligned}
\frac{d}{d\|x\|}p_{\mathbb{U}}\left(t,x,x\right) &= \lim_{r\searrow 0}\frac{p_{\mathbb{U}}(t,x+r\frac{x}{\|x\|},x+r\frac{x}{\|x\|})-p_U\left(t,x,x\right)}{r}\\
&\geq \lim_{r\searrow 0}\int_{\partial\mathbb{U}}\frac{p_{\mathbb{U}}\left(t,x+ru,x\right)-p_{\mathbb{U}}\left(t,x,x\right)}{r}d\sigma(u).
\end{aligned}$$

By bounded convergence theorem ($p_{\mathbb{U}}\left(t,\cdot,x\right)$ is a C^2 function in the second variable, hence $\nabla p_U\left(t,\cdot,x\right)$ is bounded in a neighborhood of x), we obtain

$$\frac{d}{d\|x\|}p_{\mathbb{U}}\left(t,x,x\right)\geq\int_{\partial\mathbb{U}}\nabla p_U\left(t,x,x\right)\cdot u\,d\sigma(u)=0,$$

where we denoted by $\nabla p_{\mathbb{U}}$ the gradient of $\nabla p_{\mathbb{U}}\left(t,\cdot,x\right)$ in the second variable.

Since $x\in\mathbb{U}-\{0\}$ was arbitrarily fixed, we have

$$\frac{d}{d\|x\|}p_{\mathbb{U}}(t,x,x)\geq 0,\qquad x\in(0,1),$$

which shows that $p_{\mathbb{U}}\left(t,x,x\right)$ is a non-decreasing function of $\|x\|\in(0,1)$, and by continuity this also holds for $\|x\|\in[0,1]$.

Since $p_{\mathbb{U}}\left(t,x,x\right)$ is the diagonal of a heat kernel of an operator with real analytic coefficients, $p_{\mathbb{U}}\left(t,x,x\right)$ is a real analytic function. If $p_{\mathbb{U}}\left(t,x,x\right)$ were constant on a non-empty open subset of $\overline{\mathbb{U}}$, then it would be identically constant in $\overline{\mathbb{U}}$, which is impossible (it can be shown that $p_{\mathbb{U}}(t,0,0)<p_{\mathbb{U}}(t,1,1)$ for any $t>0$). This, together with the fact that $p_{\mathbb{U}}\left(t,x,x\right)$ is a non-decreasing radial function shows that $p_{\mathbb{U}}\left(t,x,x\right)$ is in fact a strictly increasing radial function for any $t>0$, concluding the proof. □

We conclude with the remark that the Laugesen-Morpurgo conjecture implies the famous Hot Spots conjecture of J. Rauch (see for example [7], [48], [64]) in the case of the unit ball $\mathbb{U}\subset\mathbb{R}^d$, and that extending the Laugesen-Morpurgo conjecture to more general domains would also give a resolution of the Hot Spots conjecture for the corresponding domains (the Hot Spots conjecture is only partially solved at the present moment).

2.4 Extension of the mirror coupling

In [65] we showed that the mirror coupling introduced in Section 2.2 above can be extended to the case when the two reflecting Brownian motion have different state spaces, that is when X_t is a reflecting Brownian motion in a domain D_1 and Y_t is a reflecting Brownian motion in a domain D_2. Although the construction can be carried out in a more general setup (see the concluding remarks in Section 2.6), in the present section we will consider the case when one of the domains is strictly contained in the other.

The main result is the following:

Theorem 2.4.1. *Let $D_{1,2} \subset \mathbb{R}^d$ be smooth bounded domains (piecewise C^2-smooth boundary with convex corners in $\mathbb{R}^2$, or C^2-smooth boundary in $\mathbb{R}^d$, $d \geq 3$ will suffice) with $\overline{D_2} \subset D_1$ and D_2 convex domain, and let $x \in \overline{D}_1$ and $y \in \overline{D}_2$ be arbitrarily fixed points.*

Given a d-dimensional Brownian motion $(W_t)_{t\geq 0}$ starting at 0 on a probability space $(\Omega, \mathcal{F}, P)$, there exists a strong solution of the following system of stochastic differential equations

$$X_t = x + W_t + \int_0^t \nu_{D_1}(X_s)\, dL_s^X \tag{2.4.1}$$

$$Y_t = y + Z_t + \int_0^t \nu_{D_2}(Y_s)\, dL_s^Y \tag{2.4.2}$$

$$Z_t = \int_0^t G(Y_s - X_s)\, dW_s \tag{2.4.3}$$

or equivalent

$$Z_t = \int_0^t G\left(\widetilde{\Gamma}(y+Z)_s - \Gamma(x+W)_s\right) dW_s, \tag{2.4.4}$$

where Γ and $\widetilde{\Gamma}$ denote the corresponding Skorokhod maps which define the reflecting Brownian motion $X = \Gamma(x+W)$ in D_1, respectively $Y = \widetilde{\Gamma}(y+Z)$ in D_2, and $G : \mathbb{R}^d \to \mathcal{M}_{d\times d}$ denotes the following modification of the function G defined in the previous section:

$$G(z) = \begin{cases} H\left(\frac{z}{\|z\|}\right), & \text{if } z \neq 0 \\ I, & \text{if } z = 0 \end{cases}. \tag{2.4.5}$$

Remark 2.4.2. *As it will follow from the proof of the theorem, with the choice of G above, the solution of the equation (2.4.4) in the case $D_1 = D_2 = D$ is the same as the solution of the equation (2.2.5) considered by the authors in [4] (as also pointed out by the authors, the choice of $G(0)$ is irrelevant in this case).*

Therefore, the above theorem is a natural generalization of the mirror coupling to the case when the two processes live in different spaces. We will refer to a solution (X_t, Y_t) given by the above theorem as a mirror coupling *of reflecting Brownian motions in (D_1, D_2) starting from $(x, y) \in \overline{D_1} \times \overline{D_2}$, with driving Brownian motion W_t.*

*As indicated in Section 2.6, the solution of (2.4.4) is not pathwise unique, due to the fact that the stochastic differential equation has a singularity at the times when coupling occurs. The general mirror coupling can be thought as depending on a parameter which is a measure of the stickiness of the coupling: once the processes X_t and Y_t have coupled, they can either move together until one of them hits the boundary (*sticky mirror coupling *- this is in fact the solution constructed in the above theorem), or they can immediately split apart after coupling (*non-sticky mirror coupling*), and there is a whole range of intermediate possibilities (see the discussion at the end of Section 2.6).*

As an application, in Section 2.6 we will use the former mirror coupling to give a unifying proof of Chavel's conjecture on the domain monotonicity of the Neumann heat kernel for domains $D_{1,2}$ satisfying the ball condition, although the other possible choices for the mirror coupling might prove useful in other contexts.

Before carrying out the proof, we begin with some preliminary remarks which will allow us to reduce the proof of the above theorem to the case $D_1 = \mathbb{R}^d$.

Remark 2.4.3. *The main difference from the case when $D_1 = D_2 = D$ considered by the authors in [4] is that after the coupling time ξ the processes X_t and Y_t may decouple. For example, if $t \geq \xi$ is a time when $X_t = Y_t \in \partial D_2$, the process Y_t (reflecting Brownian motion in D_2) receives a push in the direction of the inward unit normal to the boundary of D_2, while the process X_t behaves like a free Brownian motion near this point (we assumed that D_2 is strictly contained in D_1), and therefore the processes X and Y will drift apart, that is they will* decouple. *Also, as shown in Section 2.6, because the function G has a discontinuity at the origin, it is possible that the solutions decouple even when they are inside the domain D_2. This shows that without additional assumptions, the mirror coupling is not uniquely determined (there is no pathwise uniqueness of (2.4.4)).*

Remark 2.4.4. *To fix ideas, for an arbitrarily fixed $\varepsilon > 0$ chosen small enough such that $\varepsilon <$ $\operatorname{dist}(\partial D_1, \partial D_2)$, we consider the sequence $(\xi_n)_{n\geq 1}$ of* coupling times *and the sequence $(\tau_n)_{n\geq 0}$ of times when the processes are ε-decoupled (ε-*decoupling times*, or simply* decoupling times *by an abuse of language) defined inductively by*

$$\begin{aligned}
\xi_n &= \inf\{t > \tau_{n-1} : X_t = Y_t\}, && n \geq 1,\\
\tau_n &= \inf\{t > \xi_n : \|X_t - Y_t\| > \varepsilon\}, && n \geq 1,
\end{aligned}$$

where $\tau_0 = 0$ and $\xi_1 = \xi$ is the first coupling time.

To construct the general mirror coupling (that is, to prove the existence of a solution to (2.4.1) – (2.4.3) above, or equivalent to (2.4.4)), we proceed as follows.

First note that on the time interval $[0, \xi]$, the arguments used in the proof of Theorem 2 in [4] (pathwise uniqueness and the existence of a strong solution Z of (2.4.4)) do not rely on the fact that $D_1 = D_2$, hence the same arguments can be used to prove the existence of a strong solution of (2.4.4) on the time interval $[0, \xi_1] = [0, \xi]$. Indeed, given W_t, (2.4.1) has a strong solution which is pathwise unique (the reflecting Brownian motion X_t in D_1), and therefore the proof of pathwise uniqueness and the existence of a strong solution of (2.4.4) is the same as in [4] considering $D = D_2$. Also note that as also pointed out by the authors, the value $G(0)$ is irrelevant in their proof, since the problem is constructing the processes until they meet, that is for $Y_t - X_t \neq 0$, for which their definition of G is the same as in (2.4.5).

We obtain therefore the existence of a strong solution Z_t to (2.4.4) on the time interval $[0, \xi_1]$. By this we understand that the process Z verifies (2.4.4) for all $t \leq \xi_1$ and Z_t is $\mathcal{F}_t$ measurable for $t \leq \xi_1$, where $(\mathcal{F}_t)_{t\geq 0}$ denotes the corresponding filtration of the driving Brownian motion W_t.

For an arbitrarily fixed $T > 0$, if $\xi_1 < T$, we can extend Z to a solution of (2.4.4) on the time interval $[0, T]$ as follows. Consider $\xi_1^T = \xi_1 \wedge T$, and note that if Z solves (2.4.4), then

$$\begin{aligned}
Z_{\xi_1^T + t} - Z_{\xi_1^T} &= \int_{\xi_1^T}^{\xi_1^T + t} G\left(\widetilde{\Gamma}(y + Z)_s - \Gamma(x + W)_s\right) dW_s\\
&= \int_0^t G\left(\widetilde{\Gamma}(y + Z)_{\xi_1^T + s} - \Gamma(x + W)_{\xi_1^T + s}\right) dW_{\xi_1^T + s}.
\end{aligned}$$

By the uniqueness results on the Skorokhod map (in the deterministic sense), we have

$$\widetilde{\Gamma}(y + Z)_{\xi_1^T + s} = \widetilde{\Gamma}\left(\widetilde{\Gamma}(y + Z)_{\xi_1^T} - Z_{\xi_1^T} + Z_{\xi_1^T + \cdot}\right)_s$$

and

$$\Gamma(x + W)_{\xi_1^T + s} = \Gamma\left(\Gamma(x + W)_{\xi_1^T} - W_{\xi_1^T} + W_{\xi_1^T + \cdot}\right)_s$$

for $s \geq 0$.

It is known that $\widetilde{W}_s = W_{\xi_1^T + s} - W_{\xi_1^T}$ is a Brownian motion starting at the origin, with corresponding filtration $\widetilde{\mathcal{F}}_s = \sigma\left(B_{\xi_1^T + u} - B_{\xi_1^T} : 0 \leq u \leq s\right)$ independent of $\mathcal{F}_{\xi_1^T}$.

Setting $\widetilde{Z}_t = Z_{\xi_1^T+t} - Z_{\xi_1^T}$ and combining the above equations we obtain

$$\widetilde{Z}_t = \int_0^t G\left(\widetilde{\Gamma}\left(\widetilde{\Gamma}\,(y+Z)_{\xi_1^T} + \widetilde{Z}\right)_s - \Gamma\left(\Gamma\,(x+W)_{\xi_1^T} + \widetilde{W}\right)_s\right) d\widetilde{W}_s, \tag{2.4.6}$$

which is the same as the equation (2.4.4) for $\widetilde{Z}$, with the initial points x, y of the coupling replaced by $Y_{\xi_1^T} = \widetilde{\Gamma}(y+Z)_{\xi_1^T}$, respectively $X_{\xi_1^T} = \Gamma(x+W)_{\xi_1^T}$, and the Brownian motion W replaced by $\widetilde{W}$.

If we assume the existence of a strong solution $\widetilde{Z}_t$ of (2.4.6) until the first ε-decoupling time, by patching Z and $\widetilde{Z}$ we obtain that

$$Z_t 1_{t\le\xi_1^T} + \widetilde{Z}_{t-\xi_1^T} 1_{\xi_1^T\le t\le\tau_1^T}$$

is a strong solution to (2.4.4) on the time interval $[0, \tau_1^T]$, where $\tau_1^T = \tau_1 \wedge T$.

If $\tau_1^T = T$, we are done. Otherwise, since at time τ_1^T the processes X and Y are ε units apart, we can apply again the results in [4] (with the Brownian motion $W_{\tau_1^T+t} - W_{\tau_1^T}$ instead of W_t, and the starting points of the coupling $X_{\tau_1^T}$ and $Y_{\tau_1^T}$ instead of x and y) in order to obtain a strong solution of (2.4.4) until the first coupling time. By patching we obtain the existence of a strong solution of (2.4.4) on the time interval $\left[0, \xi_2^T\right]$.

Proceeding inductively as indicated above, since only a finite number of coupling/decoupling times ξ_n and τ_n can occur in the time interval $[0, T]$, we can construct a strong solution Z to (2.4.4) on the time interval $[0, T]$ for any $T > 0$ (and therefore on $[0, \infty)$), provided we show the existence of strong solutions of equations of type (2.4.6) until the first ε-decoupling time.

In order to prove this claim, since $\widetilde{\Gamma}(y+Z)_{\xi_1^T}$ and $\Gamma\,(x+W)_{\xi_1^T}$ are $\mathcal{F}_{\xi_1^T}$ measurable and the σ-algebra $\mathcal{F}_{\xi_1^T}$ is independent of the filtration $\widetilde{\mathcal{F}} = (\widetilde{\mathcal{F}}_t)_{t\ge0}$ of the driving Brownian motion $\widetilde{W}_t$, it suffices to show that for any starting points $x = y \in \overline{D_2}$ of the mirror coupling, there exists a strong solution of (2.4.4) until the first ε-decoupling time τ_1. Since $\varepsilon < \operatorname{dist}(\partial D_1, \partial D_2)$, it follows that the process X_t cannot reach the boundary ∂D_1 before the first ε-decoupling time τ_1, and therefore we can consider that X_t is a free Brownian motion in $\mathbb{R}^d$, that is, we can reduce the proof of Theorem 2.4.1 to the case when $D_1 = \mathbb{R}^d$.

We will give the proof of the Theorem 2.4.1 first in the 1-dimensional case, then we will extend it to the case of polygonal domains in $\mathbb{R}^d$, and we will conclude with the proof in the general case.

2.4.1 The 1-dimensional case

From Remark 2.4.4 it follows that in order to construct the mirror coupling in the 1-dimensional case, it suffices to consider $D_1 = \mathbb{R}$ and $D_2 = (0, a)$, and to show that for an arbitrary choice $x \in [0, a]$ of the starting point of the mirror coupling, $\varepsilon \in (0, a)$ sufficiently small and $(W_t)_{t\ge0}$ a 1-dimensional Brownian motion starting at $W_0 = 0$, we can construct a strong solution on $[0, \tau_1]$ of the following system

$$X_t = x + W_t \tag{2.4.7}$$

$$Y_t = x + Z_t + L_t^Y \tag{2.4.8}$$

$$Z_t = \int_0^t G\,(Y_s - X_s)\, dW_s \tag{2.4.9}$$

where $\tau_1 = \inf\{s > 0 : |X_s - Y_s| > \varepsilon\}$ is the first ε-decoupling time and the function $G : \mathbb{R} \to \mathcal{M}_{1\times1} \equiv \mathbb{R}$ is given in this case by

$$G\,(x) = \begin{cases} -1, & \text{if } x \ne 0 \\ +1, & \text{if } x = 0 \end{cases}. \tag{2.4.10}$$

Remark 2.4.5. *Before proceeding with the proof, it is worth mentioning that the heart of the construction is Tanaka's formula. To see this, consider for the moment $a=\infty$, and note that Tanaka formula*

$$|x+W_t| = x + \int_0^t \operatorname{sgn}(x+W_s)\, dW_s + L_t^0(x+W)$$

gives a representation of the reflecting Brownian motion $|x+W_t|$ in which the increments of the martingale part of $|x+W_t|$ are the increments of W_t when $x+W_t \in [0,\infty)$, respectively the opposite (minus) of the increments of W_t in the other case ($L_t^0(x+W)$ denotes here the local time at 0 of $x+W_t$).

Since $x+W_t\in[0,\infty)$ is the same as $|x+W_t| = x+W_t$, from the definition of the function G it follows that the above can be written in the form

$$|x+W_t| = x + \int_0^t G\left(|x+W_s| - (x+W_s)\right) dW_s + L_t^{x+W},$$

which shows that a strong solution to (2.4.7) – (2.4.9) above (in the case $a=\infty$) is given explicitly by $X_t = x+W_t$, $Y_t = |x+W_t|$ and $Z_t = \int_0^t \operatorname{sgn}(x+W_s)\, dW_s$.

We have the following:

Proposition 2.4.6. *Given a 1-dimensional Brownian motion $(W_t)_{t\geq 0}$ starting at $W_0=0$, a strong solution on $[0,\tau_1]$ of the system (2.4.7) – (2.4.9) is given by*

$$\begin{cases} X_t = x+W_t \\ Y_t = |a - |x+W_t - a|| \\ Z_t = \int_0^t \operatorname{sgn}(W_s)\operatorname{sgn}(a-W_s)\, dW_s \end{cases},$$

where $\tau_1 = \inf\{s>0 : |X_s - Y_s| > \varepsilon\}$ and

$$\operatorname{sgn}(x) = \begin{cases} +1, & \text{if } x\geq 0 \\ -1, & \text{if } x<0 \end{cases}.$$

Proof. Since $\varepsilon < a$, it follows that for $t\leq\tau_1$ we have $X_t = x+W_t \in (-a,2a)$, and therefore

$$Y_t = |a - |x+W_t-a|| = \begin{cases} -(x+W_t), & x+W_t\in(-a,0) \\ x+W_t, & x+W_t\in[0,a] \\ 2a-x-W_t, & x+W_t\in(a,2a) \end{cases}. \tag{2.4.11}$$

Applying the Tanaka-Itô formula to the function $f(z) = |a-|z-a||$ and to the Brownian motion $X_t = x+W_t$, for $t\leq\tau_1$ we obtain

$$\begin{aligned} Y_t &= x + \int_0^t \operatorname{sgn}(x+W_s)\operatorname{sgn}(a-x-W_s)\, d(x+W_s) + L_t^0 - L_t^a \\ &= x + \int_0^t \operatorname{sgn}(x+W_s)\operatorname{sgn}(a-x-W_s)\, dW_s + \int_0^t \nu_{D_2}(Y_s)\, d\left(L_s^0 + L_s^a\right), \end{aligned}$$

where $L_t^0 = \sup_{s\leq t}(x+W_s)^-$ and $L_t^a = \sup_{s\leq t}(x+W_s-a)^+$ are the local times of $x+W_t$ at 0, respectively at a, and $\nu_{D_2}(0) = +1$, $\nu_{D_2}(a) = -1$.

From (2.4.11) and the definition (2.4.10) of the function G we obtain

$$\begin{aligned} \operatorname{sgn}(x+W_s)\operatorname{sgn}(a-x-W_s) &= \begin{cases} -1, & x+W_s\in(-a,0) \\ +1, & x+W_s\in[0,a] \\ -1, & x+W_s\in(a,2a) \end{cases} \\ &= \begin{cases} +1, & X_s = Y_s \\ -1, & X_s\neq Y_s \end{cases} \\ &= G(Y_s - X_s), \end{aligned}$$

and therefore the previous formula can be written equivalently

$$Y_t = x + Z_t + \int_0^t \nu_{D_2}(Y_s)\, dL_s^Y,$$

where

$$Z_t = \int_0^t G(Y_s - X_s)\, dW_s$$

and $L_t^Y = L_t^0 + L_t^a$ is a continuous nondecreasing process which increases only when $x + W_t \in \{0, a\}$, that is only when $Y_t \in \partial D_2$. □

2.4.2 The case of polygonal domains

In this section we will consider the case when $D_2 \subset D_1 \subset \mathbb{R}^d$ are polygonal domains (domains bounded by hyperplanes in $\mathbb{R}^d$). From Remark 2.4.4 it follows that we can consider $D_1 = \mathbb{R}^d$ and therefore it suffices to prove the existence of a strong solution of the following system

$$X_t = X_0 + W_t \tag{2.4.12}$$

$$Y_t = Y_0 + Z_t + \int_0^t \nu_{D_2}(Y_s)\, dL_s^Y \tag{2.4.13}$$

$$Z_t = \int_0^t G(Y_s - X_s)\, dW_s \tag{2.4.14}$$

or equivalently of the equation

$$Z_t = \int_0^t G\left(\widetilde{\Gamma}(Y_0 + Z)_s - X_0 - W_s\right) dW_s, \tag{2.4.15}$$

where W_t is a d-dimensional Brownian motion starting at $W_0 = 0$ and $X_0 = Y_0 \in \overline{D_2}$.

The construction relies on the following skew product representation of Brownian motion in spherical coordinates:

$$X_t = R_t \Theta_{\sigma_t}, \tag{2.4.16}$$

where $R_t = \|X_t\| \in \mathrm{BES}(d)$ is a Bessel process of order d and $\Theta_t \in \mathrm{BM}\left(S^{d-1}\right)$ is an independent Brownian motion on the unit sphere S^{d-1} in $\mathbb{R}^d$, run at speed

$$\sigma_t = \int_0^t \frac{1}{R_s^2} ds, \tag{2.4.17}$$

which depends only on R_t.

Remark 2.4.7. *One way to construct the Brownian motion $\Theta_t = \Theta_t^{d-1}$ on the unit sphere $S^{d-1} \subset \mathbb{R}^d$ is to proceed inductively on $d \geq 2$, using the following skew product representation of Brownian motion on the sphere $\Theta_t^{d-1} \in S^{d-1}$ (see [45]):*

$$\Theta_t^{d-1} = \left(\cos\theta_t^1, \sin\theta_t^1 \Theta_{\alpha_t}^{d-2}\right),$$

where $\theta^1 \in \mathrm{LEG}(d-1)$ is a Legendre process of order $d-1$ on $[0,\pi]$, and $\Theta_t^{d-2} \in S^{d-2}$ is an independent Brownian motion on S^{d-2}, run at speed

$$\alpha_t = \int_0^t \frac{1}{\sin^2\theta_s^1} ds.$$

Therefore, if $\theta_t^1, \ldots \theta_t^{d-1}$ are independent processes, with $\theta^i \in \mathrm{LEG}(d-i)$ on $[0,\pi]$ for $i = 1, \ldots, d-2$, and θ_t^{d-1} is a 1-dimensional Brownian (note that $\Theta_t^1 = \left(\cos\theta_t^1, \sin\theta_t^1\right) \in S^1$ is a Brownian motion on S^1), Brownian motion Θ_t^{d-1} on the unit sphere $S^{d-1} \subset \mathbb{R}^d$ is given by

$$\Theta_t^{d-1} = \left(\cos\theta_t^1, \sin\theta_t^1\cos\theta_t^2, \sin\theta_t^1\sin\theta_t^2\cos\theta_t^3, \ldots, \sin\theta_t^1 \cdot \ldots \cdot \sin\theta_t^{d-1}\sin\theta_t^{d-1}\right),$$

or by

$$\Theta_t^{d-1} = \left(\theta_t^1, \ldots, \theta_t^{d-2}, \theta_t^{d-1}\right) \tag{2.4.18}$$

in spherical coordinates.

To construct the solution of (2.4.12) – (2.4.14), we first consider the case when D_2 is a half-space $\mathcal{H}_d^+ = \left\{\left(z^1, \ldots, z^d\right) \in \mathbb{R}^d : z^d > 0\right\}$.

Given an angle $\varphi \in \mathbb{R}$, we introduce the rotation matrix $R(\varphi) \in \mathcal{M}_{d\times d}$ which leaves invariant the first $d-2$ coordinates and rotates clockwise by the angle α the remaining 2 coordinates, that is

$$R(\alpha) = \begin{pmatrix} 1 & & 0 & 0 & 0 \\ & \ddots & & \cdots & \cdots \\ 0 & & 1 & 0 & 0 \\ 0 & \cdots & 0 & \cos\varphi & -\sin\varphi \\ 0 & \cdots & 0 & \sin\varphi & \cos\varphi \end{pmatrix}. \tag{2.4.19}$$

We have the following:

Lemma 2.4.8. *Let $D_2 = \mathcal{H}_d^+ = \left\{\left(z^1, \ldots, z^d\right) \in \mathbb{R}^d : z^d > 0\right\}$ and assume that*

$$Y_0 = R(\varphi_0) X_0 \tag{2.4.20}$$

for some $\varphi_0 \in \mathbb{R}$.

Consider the reflecting Brownian motion $\widetilde{\theta}_t^{d-1}$ on $[0,\pi]$ with driving Brownian motion θ_t^{d-1}, where θ_t^{d-1} is the $(d-1)$ spherical coordinate of $G(Y_0 - X_0) X_t$, given by (2.4.16) – (2.4.18) above, that is:

$$\widetilde{\theta}_t^{d-1} = \theta_t^{d-1} + L_t^0\left(\widetilde{\theta}^{d-1}\right) - L_t^\pi\left(\widetilde{\theta}^{d-1}\right), \qquad t \geq 0,$$

and $L_t^0\left(\widetilde{\theta}^{d-1}\right)$, $L_t^\pi\left(\widetilde{\theta}^{d-1}\right)$ represent the local times of $\widetilde{\theta}^{d-1}$ at 0, respectively at π.

A strong solution of the system (2.4.12) – (2.4.14) is explicitly given by

$$Y_t = \begin{cases} R(\varphi_t) G(Y_0 - X_0) X_t, & t < \xi \\ |X_t|_d, & t \geq \xi \end{cases} \tag{2.4.21}$$

where $\xi = \inf\{t > 0 : X_t = Y_t\}$ is the coupling time, the rotation angle φ_t is given by

$$\varphi_t = L_t^0\left(\widetilde{\theta}^{d-1}\right) - L_t^\pi\left(\widetilde{\theta}^{d-1}\right), \qquad t \geq 0,$$

and for $z = \left(z^1, z^2 \ldots, z^d\right) \in \mathbb{R}^d$ we denoted by $|z|_d = \left(z^1, z^2, \ldots, \left|z^d\right|\right)$.

Proof. Recall that for $m \in \mathbb{R}^d - \{0\}$, $G(m)v$ denotes the mirror image of $v \in \mathbb{R}^d$ with respect to the hyperplane through the origin perpendicular to m.

By Itô formula, we have

$$Y_{t\wedge\xi} = Y_0 + \int_0^{t\wedge\xi} R(\varphi_s) G(Y_0 - X_0) dX_s + \int_0^{t\wedge\xi} R\left(\varphi_s + \frac{\pi}{2}\right) G(Y_0 - X_0) dL_s. \tag{2.4.22}$$

Note that the composition $R \circ G$ (a symmetry followed by a rotation) is a symmetry, and since $\|Y_t\| = \|X_t\|$ for all $t \geq 0$, it follows that X_t and Y_t are symmetric with respect to a hyperplane passing through the origin for all $t \leq \xi$. Therefore, from the definition (2.4.5) of the function G it follows that we have $Y_t = G(Y_t - X_t) X_t$ for all $t \leq \xi$.

Also note that when $L_s^0\left(\widetilde{\theta}^{d-1}\right)$ increases, $Y_s \in \partial D_2$ and we have

$$R\left(\varphi_s + \frac{\pi}{2}\right) G(Y_0 - X_0) X_s = R\left(\frac{\pi}{2}\right) Y_s = \nu_{D_2}(Y_s),$$

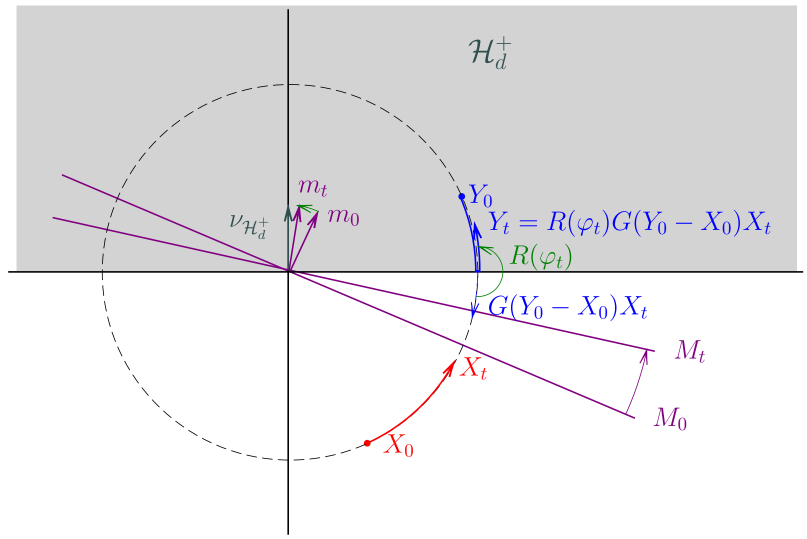

Figure 2.7: The mirror coupling of a free Brownian motion X_t and a reflecting Brownian motion Y_t in the half-space $\mathcal{H}_d^+$.

and if $L_s^\pi\left(\widetilde{\theta}^{d-1}\right)$ increases, $Y_s \in \partial D_2$ and we have

$$R\left(\varphi_s + \frac{\pi}{2}\right) G\left(Y_0 - X_0\right) X_s = R\left(\frac{\pi}{2}\right) Y_s = -\nu_{D_2}\left(Y_s\right).$$

It follows that the relation (2.4.22) can be written in the equivalent form

$$Y_{t\wedge\xi} = Y_0 + \int_0^{t\wedge\xi} G\left(Y_s - X_s\right) dX_s + \int_0^{t\wedge\xi} \nu_{D_2}\left(Y_s\right) dL_s^Y,$$

where $L_t^Y = L_t^0\left(\widetilde{\theta}^{d-1}\right) + L_t^\pi\left(\widetilde{\theta}^{d-1}\right)$ is a continuous non-decreasing process which increases only when $Y_t \in \partial D_2$, and therefore Y_t given by (2.4.21) is a strong solution of the system (2.4.12) – (2.4.14) for $t \leq \xi$.

For $t \geq \xi$, we have $Y_t = |X_t|_d = \left(X_t^1, X_t^2, \ldots, \left|X_t^d\right|\right)$, and proceeding similarly to the 1-dimensional case, by Tanaka formula we obtain:

$$\begin{aligned} Y_{t\vee\xi} &= Y_\xi + \int_\xi^{t\vee\xi} \left(1,\ldots,1,\operatorname{sgn}\left(X_s^d\right)\right) dX_s + \int_\xi^{t\vee\xi} \left(0,\ldots,0,1\right) L_t^0\left(X^d\right) \qquad (2.4.23) \\ &= Y_\xi + \int_\xi^{t\vee\xi} G\left(Y_s - X_s\right) dX_s + \int_\xi^{t\vee\xi} \nu_{D_2}\left(Y_s\right) L_t^Y, \end{aligned}$$

since in this case

$$\begin{aligned} G\left(Y_s - X_s\right) &= \begin{cases} \left(1,\ldots,1,+1\right), & X_s = Y_s \\ \left(1,\ldots,1,-1\right), & X_s \neq Y_s \end{cases} \\ &= \begin{cases} \left(1,\ldots,1,+1\right), & X_s^d \geq 0 \\ \left(1,\ldots,1,-1\right), & X_s^d < 0 \end{cases} \\ &= \left(1,\ldots,1,\operatorname{sgn}\left(X_s^d\right)\right). \end{aligned}$$

The process $L_t^Y = L_t^0\left(X^d\right)$ in (2.4.23) is a continuous non-decreasing process which increases only when $Y_t \in \partial D_2$ ($L_t^0\left(X^d\right)$ represents the local time at 0 of the last cartesian coordinate X^d of X), which shows that Y_t also solves (2.4.12) – (2.4.14) for $t \geq \xi$, and therefore Y_t is a strong solution of (2.4.12) – (2.4.14) for $t \geq 0$, concluding the proof. □

Consider now the case of a general polygonal domain $D_2 \subset \mathbb{R}^d$. We will show that a strong solution of the system (2.4.12) – (2.4.14) can be constructed from the previous lemma by choosing the appropriate coordinate system.

Consider the times $(\sigma_n)_{n\geq 0}$ at which the solution Y_t hits different bounding hyperplanes of ∂D_2, that is $\sigma_0 = \inf\{s \geq 0 : Y_s \in \partial D_2\}$ and inductively

$$\sigma_{n+1} = \inf\left\{ t \geq \sigma_n : \begin{array}{l} Y_t \in \partial\mathcal{D}_2 \text{ and } Y_t, Y_{\sigma_n} \text{ belong to different}^1 \\ \text{bounding hyperplanes of } \partial D_2 \end{array} \right\}, \ n \geq 0. \tag{2.4.24}$$

If $X_0 = Y_0 \in \partial D_2$ belong to a certain bounding hyperplane of D_2, we can chose the coordinate system so that this hyperplane is $\mathcal{H}_d = \left\{\left(z^1, \ldots, z^d\right) \in \mathbb{R}^d : z^d = 0\right\}$ and $D_2 \subset \mathcal{H}_d^+$, and we let $\mathcal{H}_d$ be any bounding hyperplane of D_2 otherwise.

By Lemma 2.4.8 it follows that on the time interval $[\sigma_0, \sigma_1)$, the strong solution of (2.4.12) – (2.4.14) is given explicitly by (2.4.21).

If $\sigma_1 < \infty$, we distinguish two cases: $X_{\sigma_1} = Y_{\sigma_1}$ and $X_{\sigma_1} \neq Y_{\sigma_1}$. Let $\mathcal{H}$ denote the bounding hyperplane of D which contains Y_{σ_1}, and let $\nu_{\mathcal{H}}$ denote the unit normal to $\mathcal{H}$ pointing inside D_2.

If $X_{\sigma_1} = Y_{\sigma_1} \in \mathcal{H}$, choosing again the coordinate system conveniently, we may assume that $\mathcal{H}$ is the hyperplane is $\mathcal{H}_d = \left\{\left(z^1, \ldots, z^d\right) \in \mathbb{R}^d : z^d = 0\right\}$, and on the time interval $[\sigma_1, \sigma_2)$ the coupling $(X_{\sigma_1+t}, Y_{\sigma_1+t})_{t\in[0,\sigma_2-\sigma_1)}$ is given again by Lemma 2.4.8.

If $X_{\sigma_1} \neq Y_{\sigma_1} \in \mathcal{H}$, in order to apply Lemma 2.4.8 we have to show that we can choose the coordinate system so that the condition (2.4.20) holds. If $Y_{\sigma_1} - X_{\sigma_1}$ is a vector perpendicular to $\mathcal{H}$, by choosing the coordinate system so that $\mathcal{H} = \mathcal{H}_d = \left\{\left(z^1, \ldots, z^d\right) \in \mathbb{R}^d : z^d = 0\right\}$, the problem reduces to the 1-dimensional case (the first $d-1$ coordinates of X and Y are the same), and it can be handled as in Proposition 2.4.6 by the Tanaka formula. The proof being similar, we omit it.

If $X_{\sigma_1} \neq Y_{\sigma_1} \in \mathcal{H}$ and $Y_{\sigma_1} - X_{\sigma_1}$ is not orthogonal to $\mathcal{H}$, consider $\widetilde{X}_{\sigma_1} = \mathrm{pr}_{\mathcal{H}} X_{\sigma_1}$ the projection of X_{σ_1} onto $\mathcal{H}$, and therefore $\widetilde{X}_{\sigma_1} \neq Y_{\sigma_1}$. The plane of symmetry of X_{σ_1} and Y_{σ_1} intersects the line determined by $\widetilde{X}_{\sigma_1}$ and Y_{σ_1} at a point, and we consider this point as the origin of the coordinate system (note that the intersection cannot be empty, for otherwise the vectors $Y_{\sigma_1} - X_{\sigma_1}$ and $Y_{\sigma_1} - \widetilde{X}_{\sigma_1}$ were parallel, which is impossible since then $Y_{\sigma_1} - X_{\sigma_1}, Y_{\sigma_1} - \widetilde{X}_{\sigma_1}$ and $Y_{\sigma_1} - \widetilde{X}_{\sigma_1}, X_{\sigma_1} - \widetilde{X}_{\sigma_1}$ were perpendicular pairs of vectors, contradicting $\widetilde{X}_{\sigma_1} \neq Y_{\sigma_1}$ – see Figure 2.8).

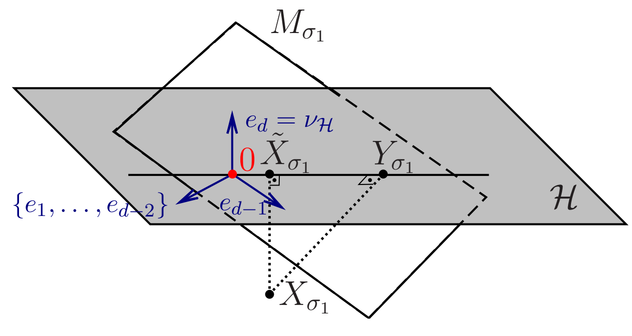

Figure 2.8: Construction of the appropriate coordinate system.

Choose an orthonormal basis $\{e_1, \ldots, e_d\}$ in $\mathbb{R}^d$ such that $e_d = \nu_{\mathcal{H}}$ is the normal vector to $\mathcal{H}$ pointing inside D_2, $e_{d-1} = \frac{1}{\|Y_{\sigma_1} - X_{\sigma_1}\|}(Y_{\sigma_1} - X_{\sigma_1})$ is a unit vector lying in the 2-dimensional plane determined by the origin and the vectors e_d and $Y_{\sigma_1} - X_{\sigma_1}$, and $\{e_1, \ldots, e_{d-2}\}$ is a completion of $\{e_{d-1}, e_d\}$ to an orthonormal basis in $\mathbb{R}^d$ (see Figure 2.8).

[1] Since 2-dimensional Brownian motion does not hit points a.s., the d-dimensional Brownian motion Y_t does not hit the edges of D_2 ($(d-2)$-dimensional hyperplanes in $\mathbb{R}^d$) a.s., thus there is no ambiguity in the definition.

Note that by the construction, the vectors $e_1, \ldots, e_{d-2}$ are orthogonal to the 2-dimensional hyperplane containing the origin and the points X_{σ_1} and Y_{σ_1}, and therefore X_{σ_1} and Y_{σ_1} have the same (zero) first $d-2$ coordinates; also, since X_{σ_1} and Y_{σ_1} are at the same distance from the origin, it follows that Y_{σ_1} can be obtained from X_{σ_1} by a rotation which leaves invariant the first $d-2$ coordinates, which shows that the condition (2.4.20) of Lemma 2.4.8 is satisfied.

Since by construction the bounding hyperplane $\mathcal{H}$ of D_2 at Y_{σ_1} is given by

$$\mathcal{H}_d = \left\{ \left(z^1, \ldots, z^d\right) \in \mathbb{R}^d : z^d = 0 \right\}$$

and $D_2 \subset \mathcal{H}_d^+ = \left\{ \left(z^1, \ldots, z^d\right) \in \mathbb{R}^d : z^d > 0 \right\}$, we can apply Lemma 2.4.8 to deduce that on the time interval $[\sigma_1, \sigma_2)$ a solution of (2.4.12) – (2.4.14) is given by $(X_{\sigma_1+t}, Y_{\sigma_1+t})_{t\in[0,\sigma_2-\sigma_1)}$.

Repeating the above argument we can construct inductively (in the appropriate coordinate systems) the solution of (2.4.12) – (2.4.14) on any time interval $[\sigma_n, \sigma_{n+1})$, $n \geq 1$, and therefore we obtain a strong solution of (2.4.12) – (2.4.14) defined for $t \geq 0$.

We summarize the above discussion in the following:

Theorem 2.4.9. *If $D_2 \subset \mathbb{R}^d$ is a polygonal domain, for any $X_0 = Y_0 \in \overline{D_2}$, there exists a strong solution of the system (2.4.12) - (2.4.14).*

Moreover, between successive hits of different bounding hyperplanes of D_2 (i.e. on each time interval $[\sigma_n, \sigma_{n+1})$ in the notation above), the solution is given by Lemma 2.4.8 in the appropriately chosen coordinate system.

We will refer to the solution $(X_t, Y_t)_{t\geq 0}$ constructed in the previous theorem as a *mirror coupling* of reflecting Brownian motions in $\left(\mathbb{R}^d, D_2\right)$ with starting point $X_0 = Y_0 \in \overline{D_2}$.

If $X_t \neq Y_t$, the hyperplane M_t of symmetry between X_t and Y_t (the hyperplane passing through $\frac{X_t+Y_t}{2}$ with normal $m_t = \frac{1}{\|Y_t - X_t\|}(Y_t - X_t)$) will be referred to as the *mirror of the coupling*. For definiteness, when $X_t = Y_t$ we let M_t denote any hyperplane passing through $X_t = Y_t$, for example we can choose M_t such that it is a left continuous function with respect to t.

In the particular case of a convex polygonal domain D_2, some of the properties of the mirror coupling are contained in the following:

Proposition 2.4.10. *If $D_2 \subset \mathbb{R}^d$ is a convex polygonal domain, for any $X_0 = Y_0 \in \overline{D_2}$, the mirror coupling given by the previous theorem has the following properties:*

i) If the reflection takes place in the bounding hyperplane $\mathcal{H}$ of D_2 with inward unitary normal $\nu_{\mathcal{H}}$, then the angle $\angle(m_t; \nu_{\mathcal{H}})$ decreases monotonically to zero.

ii) When processes are not coupled, the mirror M_t lies outside D_2.

iii) Coupling can take place precisely when $X_t \in \partial D_2$. Moreover, if $X_t \in D_2$, then $X_t = Y_t$.

iv) If $D_\alpha \subset D_\beta$ are two polygonal domains and $(Y_t^\alpha; X_t)$, $(Y_t^\beta; X_t)$ are the corresponding mirror coupling starting from $x \in \overline{D_\alpha}$, for any $t > 0$ we have

$$\sup_{s\leq t} \|Y_s^\alpha - Y_s^\beta\| \leq \operatorname{Dist}\left(D^\alpha, D^\beta\right) := \max_{\substack{x_\alpha\in\partial D_\alpha, x_\beta\in\partial D_\beta \\ (x_\beta - x_\alpha)\cdot\nu_{D_\alpha}(x_\alpha)\leq 0}} \|x_\alpha - x_\beta\|. \tag{2.4.25}$$

Proof. i) In the notation of Theorem 2.4.9, on the time interval $[\sigma_0, \sigma_1)$ we have $Y_t = X_t$, so $\angle(m_t, \nu_{\mathcal{H}}) = 0$ and therefore the claim is verified in this case.

On an arbitrary time interval $[\sigma_n, \sigma_{n+1})$, in the appropriately chosen coordinate system, Y_t is given by Lemma 2.4.8. For $t < \xi$, Y_t is given by the rotation $R(\varphi_t)$ of $G(Y_0 - X_0) X_t$ which leaves invariant the first $(d-2)$ coordinates, and therefore

$$\angle(m_t, \nu_{\mathcal{H}}) = \angle(m_0, \nu_{\mathcal{H}}) + \frac{L_t^0 - L_t^\pi}{2},$$

which proves the claim in this case (note that before the coupling time ξ only one of the non-decreasing processes L_t^0 and L_t^π is not identically zero).

Since for $t \geq \xi$ we have $Y_t = \left(X_t^1, \ldots, \left|X_t^d\right|\right)$, we have $\angle(m_t, \nu_{\mathcal{H}}) = 0$ which concludes the proof of the claim.

ii) On the time interval $[\sigma_0, \sigma_1)$ the processes are coupled, so there is nothing to prove in this case.

On the time interval $[\sigma_1, \sigma_2)$, in the appropriately chosen coordinate system we have $Y_t = \left(X_t^1, \ldots, \left|X_t^d\right|\right)$, thus the mirror M_t coincides with the boundary hyperplane

$$\mathcal{H}_d = \left\{\left(z^1, \ldots, z^d\right) \in \mathbb{R}^d : z^d = 0\right\}$$

of D_2 where the reflection takes place, and therefore $M_t \cap D_2 = \varnothing$ in this case.

Inductively, assume the claim is true for $t < \sigma_n$. By continuity, $M_{\sigma_n} \cap D_2 = \varnothing$, thus D_2 lies on one side of M_{σ_n}. By the previous proof, the angle $\angle(m_t, \nu_{\mathcal{H}})$ between m_t and the inward unit normal $\nu_{\mathcal{H}}$ to bounding hyperplane $\mathcal{H}$ of D_2 where the reflection takes place decreases to zero. Since D_2 is a convex domain, it follows that on the time interval $[\sigma_n, \sigma_{n+1})$ we have $M_t \cap D_2 = \varnothing$, concluding the proof.

iii) The first part of the claim follows from the previous proof (when the processes are not coupled, the mirror (hence X_t) lies outside D_2; by continuity, it follows that at the coupling time ξ we must have $X_\xi = Y_\xi \in \partial D_2$).

To prove the second part of the claim, consider an arbitrary time interval $[\sigma_n, \sigma_{n+1})$ between two successive hits of Y_t to different bounding hyperplanes of D_2. In the appropriately chosen coordinate system, Y_t is given by Lemma 2.4.8. After the coupling time ξ, Y_t is given by $Y_t = \left(X_t^1, \ldots, \left|X_t^d\right|\right)$, and therefore if $X_t \in D_2$ (thus $X_t^d \geq 0$) we have $Y_t = \left(X_t^1, \ldots, X_t^d\right) = X_t$, concluding the proof.

iv) Let M_t^α and M_t^β denote the mirrors of the coupling in D^α, respectively D^β, with the same driving Brownian motion X_t.

Since Y_t^α and X_t are symmetric with respect to M_t^α, and Y_t^β and X_t are symmetric with respect to M_t^β, it follows that Y_t^β is obtained from Y_t^β by a rotation which leaves invariant the hyperplane $M_t^\alpha \cap M_t^\beta$, or by a translation by a vector orthogonal to M_t^α (in the case when M_t^α and M_t^β are parallel).

The angle of rotation (respectively the vector of translation) is altered only when either Y_t^α or Y_t^β are on the boundary of D_α, respectively D_β. Since $D_\alpha \subset D_\beta$ are convex domains, the angle of rotation (respectively the vector of translation) decreases when $Y_t^\beta \in D_\beta$ or when $Y_t^\alpha \in \partial D_\alpha$ and $\left(Y_t^\beta - Y_t^\alpha\right) \cdot \nu_{D_\alpha}\left(Y_t^\alpha\right) > 0$ (in these cases Y_t^β and Y_t^α receive a push such that the distance $\|Y_t^\alpha - Y_t^\beta\|$ is decreased), thus the maximum distance $\|Y_t^\alpha - Y_t^\beta\|$ is attained when $Y_t^\alpha \in \partial D_\alpha$ and $\left(Y_t^\beta - Y_t^\alpha\right) \cdot \nu_{D_\alpha}\left(Y_t^\alpha\right) \leq 0$, and the formula follows. □

2.5 The proof of Theorem 2.4.1

By Remark 2.4.4, it suffices to consider the case when $D_1 = \mathbb{R}^d$ and $D_2 \subset \mathbb{R}^d$ is a convex bounded domain with smooth boundary. To simplify the notation, we will drop the index and write D for D_2 in the sequel.

Let $(D_n)_{n\geq 1}$ be an increasing sequence of convex polygonal domains in $\mathbb{R}^d$ with $\overline{D_n} \subset D_{n+1}$ and $\cup_{n\geq 1} D_n = D$.

Consider $(Y_t^n, X_t)_{t\geq 0}$ a sequence of mirror couplings in $\left(D_n, \mathbb{R}^d\right)$ with starting point $x \in D_1$ and driving Brownian motion $(W_t)_{t\geq 0}$ with $W_0 = 0$, given by Theorem 2.4.9.

By Proposition 2.4.10, for any $t > 0$ we have

$$\sup_{s\leq t} |Y_s^m - Y_s^n| \leq \operatorname{Dist}(D_n, D_m) = \max_{\substack{x_n \in \partial D_n, x_m \in \partial D_m \\ (x_m - x_n)\cdot \nu_{D_n}(x_n) \leq 0}} |x_n - x_m| \underset{n,m\to\infty}{\rightarrow} 0,$$

hence Y_t^n converges a.s. in the uniform topology to a continuous process Y_t.

Since $(Y^n)_{n\geq 1}$ are reflecting Brownian motions in $(D_n)_{n\geq 1}$ and $D_n \nearrow D$, the law of Y_t is that of a reflecting Brownian motion in D, that is Y_t is a reflecting Brownian motion in D starting at $x \in D$ (see [22]). Also note that since Y_t^n are adapted to the filtration $\mathcal{F}^W = (\mathcal{F}_t)_{t\geq 0}$ generated by the Brownian motion W_t, so is Y_t.

By construction, the driving Brownian motion Z_t^n of Y_t^n satisfies

$$Z_t^n = \int_0^t G\left(Y_t^n - X_t\right) dW_t, \qquad t \geq 0.$$

Consider the process

$$Z_t = \int_0^t G\left(Y_s - X_s\right) dW_s,$$

and note that since Y is $\mathcal{F}^W$-adapted and $||G|| = 1$, by Lévy's characterization of Brownian motion, Z_t is a free d-dimensional Brownian motion starting at $Z_0 = 0$, also adapted to the filtration $\mathcal{F}^W$.

We will show that Z is the driving process of the reflecting Brownian motion Y_t, that is, we will show that

$$Y_t = x + Z_t + L_t^Y = x + \int_0^t G\left(Y_s - X_s\right) dW_s + L_t^Y, \qquad t \geq 0.$$

Note that the mapping $z \longmapsto G(z)$ is continuous with respect to the norm $||A|| = ||(a_{ij})|| = \sum_{i,j=1}^d a_{ij}^2$ of $d \times d$ matrices at all points $z \in \mathbb{R}^d - \{0\}$, hence $G\left(Y_s^n - X_s\right) \underset{n\to\infty}{\to} G\left(Y_s - X_s\right)$ if $Y_s - X_s \neq 0$. If $Y_s - X_s = 0$, then either $Y_s = X_s \in D$ or $Y_s = X_s \in \partial D$.

If $Y_s = X_s \in D$, since $D_n \nearrow D$, there exists $N \geq 1$ such that $X_s \in D_N$, hence $X_s \in D_n$ for all $n \geq N$. By Proposition 2.4.10, it follows that $Y_s^n = X_s$ for all $n \geq N$, hence in this case we also have $G\left(Y_s^n - X_s\right) = G(0) \underset{n\to\infty}{\to} G(0) = G\left(Y_s - X_s\right)$.

If $Y_s = X_s \in \partial D$, since $\overline{D_n} \subset D$ we have $Y_s^n - X_s \neq 0$, and therefore by the definition (2.4.5) of G we have:

$$\begin{aligned}
& \int_0^t \left|\left|G\left(Y_s^n - X_s\right) - G\left(Y_s - X_s\right)\right|\right|^2 1_{Y_s = X_s \in \partial D} ds \\
= & \int_0^t \left|\left|H\left(\frac{Y_s^n - X_s}{||Y_s^n - X_s||}\right) - I\right|\right|^2 1_{Y_s = X_s \in \partial D} ds \\
= & \int_0^t \left|\left|I - 2\frac{Y_s^n - X_s}{||Y_s^n - X_s||}\left(\frac{Y_s^n - X_s}{||Y_s^n - X_s||}\right)' - I\right|\right|^2 1_{Y_s = X_s \in \partial D} ds \\
= & \int_0^t \left|\left|2\frac{Y_s^n - X_s}{||Y_s^n - X_s||}\left(\frac{Y_s^n - X_s}{||Y_s^n - X_s||}\right)'\right|\right|^2 1_{Y_s = X_s \in \partial D} ds \\
= & \ 4\int_0^t 1_{Y_s = X_s \in \partial D} ds \\
\leq & \ 4\int_0^t 1_{\partial D}\left(Y_s\right) ds \\
= & \ 0,
\end{aligned}$$

since Y_t is a reflecting Brownian motion in D, and therefore it spends zero Lebesgue time on the boundary of D.

Since $||G|| = 1$, using the above and the bounded convergence theorem we obtain

$$\lim_{n\to\infty} \int_0^t \left|\left|G\left(Y_s^n - X_s\right) - G\left(Y_s - X_s\right)\right|\right|^2 ds = 0,$$

and therefore by Doob's inequality it follows that

$$E\sup_{s\le t}\|Z_s^n-Z_s\|^2\le cE\|Z_t^n-Z_t\|^2\le cE\int_0^t\|G\left(Y_s^n-X_s\right)-G\left(Y_s-X_s\right)\|^2ds\underset{n\to\infty}{\to}0,$$

for any $t\ge 0$, which shows that Z_t^n converges uniformly on compact sets to $Z_t=\int_0^t G\left(Y_s-X_s\right)dW_s$.

By construction, Z_t^n is the driving Brownian motion for Y_t^n, that is

$$Y_t^n=x+Z_t^n+\int_0^t\nu_{D_n}\left(Y_s^n\right)dL_s^{Y_n},$$

and passing to the limit with $n\to\infty$ we obtain

$$Y_t=x+Z_t+A_t=x+\int_0^t G\left(Y_s-X_s\right)dW_s+A_t,\qquad t\ge 0,$$

where $A_t=\lim_{n\to\infty}\int_0^t\nu_{D_n}\left(Y_s^n\right)dL_s^{Y_n}$.

It remains to show that A_t is a process of bounded variation. For an arbitrary partition $0=t_0<t_1<\ldots t_l=t$ of $[0,t]$ we have

$$\begin{aligned}
E\sum_{i=1}^{l}\|A_{t_i}-A_{t_{i-1}}\| &= \lim_{n\to\infty}E\sum_{i=1}^{l}\left\|\int_{t_{i-1}}^{t_i}\nu_{D_n}\left(Y_s^n\right)dL_s^{Y_n}\right\|\\
&\le \limsup E\ L_t^{Y_n}\\
&= \limsup\int_0^t\int_{\partial D_n}p_{D_n}\left(s,x,y\right)\sigma_n\left(dy\right)ds\\
&\le c\sqrt{t},
\end{aligned}$$

where σ_n is the surface measure on ∂D_n, and the last inequality above follows from the estimates in [15] on the Neumann heat kernels $p_{D_n}\left(t,x,y\right)$ (see the remarks preceding Theorem 2.1 and the proof of Theorem 2.4 in [23]).

From the above it follows that $A_t=Y_t-x-Z_t$ is a continuous, $\mathcal{F}^W$-adapted process (since Y_t, Z_t are continuous, $\mathcal{F}^W$-adapted processes) of bounded variation.

By the uniqueness in the Doob-Meyer semimartingale decomposition of the reflecting Brownian motion Y_t in D, it follows that

$$A_t=\int_0^t\nu_D\left(Y_s\right)dL_s^Y,\qquad t\ge 0,$$

where L^Y is the local time of Y on the boundary ∂D. It follows that the reflecting Brownian motion Y_t in D constructed above is a strong solution to

$$Y_t=x+\int_0^t G\left(Y_s-X_s\right)dW_s+\int_0^t\nu_D\left(Y_s\right)dL_s^Y,\qquad t\ge 0,$$

or equivalent, the driving Brownian motion $Z_t=\int_0^t G\left(Y_s-X_s\right)dW_s$ of Y_t is a strong solution to

$$Z_t=\int_0^t G\left(\widetilde{\Gamma}\left(y+Z\right)_s-X_s\right)dW_s,\qquad t\ge 0,$$

concluding the proof of Theorem 2.4.1.

2.6 Extensions and applications

As an application of the construction of mirror coupling, we will present a unifying proof of the two most important results on Chavel's conjecture.

It is not difficult to prove that the Dirichlet heat kernel is an increasing function with respect to the domain. Since for the Neumann heat kernel $p_D(t,x,y)$ of a smooth bounded domain $D \subset \mathbb{R}^d$ we have

$$\lim_{t\to\infty} p_D(t,x,y) = \frac{1}{\mathrm{vol}(D)},$$

the monotonicity in the case of the Neumann heat kernel should be reversed.

The above observation was conjectured by Isaac Chavel ([29]), as follows:

Conjecture 2.6.1 (Chavel's conjecture, [29])**.** *Let $D_{1,2} \subset \mathbb{R}^d$ be smooth bounded convex domains in $\mathbb{R}^d$, $d \geq 1$, and let $p_{D_1}(t,x,y)$, $p_{D_2}(t,x,y)$ denote the Neumann heat kernels in D_1, respectively D_2. If $D_2 \subset D_1$, then*

$$p_{D_1}(t,x,y) \leq p_{D_2}(t,x,y), \tag{2.6.1}$$

for any $t \geq 0$ and $x, y \in D_1$.

Remark 2.6.2. *The smoothness assumption in the above conjecture is meant to insure the a.e. existence of the inward unit normal to the boundaries of D_1 and D_2, that is the boundaries should have a locally differentiable parametrization. Requiring that the boundary of the domain is of class $C^{1,\alpha}$ ($0<\alpha<1$) is a convenient hypothesis on the smoothness of the domains $D_{1,2}$.*

In order to simplify the proof, we will assume that $D_{1,2}$ are smooth C^2 domains (the proof can be extended to a more general setup, by approximating $D_{1,2}$ by less smooth domains).

Among the positive results on Chavel conjecture, the most general known results (and perhaps the easiest to use in practice) are due to I. Chavel ([29]) and W. Kendall ([51]), and they show that if there exists a ball B centered at either x or y such that $D_2 \subset B \subset D_1$, then the inequality (2.6.1) in Chavel's conjecture holds for any $t > 0$.

While there are also other positive results which suggest that Chavel's conjecture is true for certain classes of domains (see for example [27], [39]), in [12] R. Bass and K. Burdzy showed that Chavel's conjecture does not hold in its full generality (i.e. without additional hypotheses).

We note that both the proof of Chavel (the case when D_1 is a ball centered at either x or y) and Kendall (the case when D_2 is a ball centered at either x or y) relies in an essential way that one of the domains is a ball: the first uses an integration by parts technique, while the later uses a coupling argument of the radial parts of Brownian motion, and none of these proofs seem to be easily applicable to the other case.

Using the mirror coupling, we can derive a simple, unifying proof of these two important results, as follows:

Theorem 2.6.3. *Let $D_2 \subset D_1 \subset \mathbb{R}^d$ be smooth bounded domains and assume that D_2 is convex. If for $x, y \in D_2$ there exists a ball B centered at either x or y such that $D_2 \subset B \subset D_1$, then for all $t \geq 0$ we have*

$$p_{D_1}(t,x,y) \leq p_{D_2}(t,x,y). \tag{2.6.2}$$

Proof. Consider $x, y \in D_2$ arbitrarily fixed and assume that $D_2 \subset B = B(y,R) \subset D_1$ for some $R > 0$.

By eventually approximating the convex domain D_2 by convex polygonal domains, it suffices to prove the claim in the case when D_2 is a convex polygonal domain.

Let (X_t, Y_t) be a mirror coupling of reflecting Brownian motions in (D_1, D_2) starting at $x \in D_2$. The idea of the proof is to show that for all times $t \geq 0$, Y_t is at a distance from y is no greater than the distance from X_t to y.

Let $t_0 \geq 0$ be a time when the processes are at the same distance from y, and let $t_1 \geq t_0$ be the first time after t_0 when the process X_t hits the boundary of D_1.

Note that by the ball condition we have $\|X_t - y\| = R > \|Y_t - y\|$ for any $t \geq 0$, and in particular this holds for $t = t_1$. Since the processes X_t and Y_t are continuous, the distances from

X_t and Y_t to y are continuous functions of t, and therefore in order to prove the claim it suffices to show that $\|Y_t - y\| \le \|X_t - y\|$ for all $t \in [t_0, t_1]$. Also note that on the time interval $[t_0, t_1]$ the process X_t behaves like a free Brownian motion.

We distinguish the following cases:

i) The processes are coupled at time t_0 (i.e. $X_{t_0} = Y_{t_0}$);

In this case, the distances from X_t and Y_t to y will remain equal until the first time when the processes hit the boundary of D_2. Since on the time interval $[t_0, t_1]$ the process X_t behaves like a free Brownian motion, by Proposition 2.4.10 ii) it follows that when processes are not coupled, the mirror M_t of the coupling lies outside the domain D_2. Since the domain D_2 is assumed convex, this shows in particular that the mirror M_t of the coupling cannot separate the points Y_t and y, and therefore the distance from Y_t to y is smaller than or equal to the distance from X_t to y, for all $t \in [t_0, t_1]$.

ii) The processes are decoupled at time t_0;

In this case, since $|Y_{t_0} - y| = |X_{t_0} - y|$ and $X_{t_0} \neq Y_{t_0}$, the hyperplane M_{t_0} of symmetry between X_{t_0} and Y_{t_0} passes through the point y, so M_{t_0} does not separate the points Y_{t_0} and y.

The processes X_t and Y_t will remain at the same distance from y until the first time when $Y_t \in \partial D_2$. Since on the time interval $[t_0, t_1]$ the process X_t behaves like a free Brownian motion, by Theorem 2.4.9, it follows that between successive hits of different boundary hyperplanes of D_2, the mirror M_t of the coupling describes a rotation which leaves invariant $d-2$ coordinate axes. Moreover, by Proposition 2.4.10 the rotation is directed in such a way that the angle $\angle(m_t, \nu_{\mathcal{H}})$ between the normal $m_t = \frac{1}{\|Y_t - X_t\|}(Y_t - X_t)$ of M_t and the inner normal $\nu_{\mathcal{H}}$ of the bounding hyperplane $\mathcal{H}$ of D_2 where the reflection takes place decreases monotonically to zero (see Figure 2.7).

Since the hyperplane M_{t_0} does not separate the points Y_{t_0} and y, simple geometric considerations show that M_t will not separate the points Y_t and y for all $t \in [t_0, t_1]$, and therefore $\|Y_t - y\| \le \|X_t - y\|$ for all $t \in [t_0, t_1]$, concluding the proof of the claim.

We showed that for any $t \ge 0$ we have $\|Y_t - y\| \le \|X_t - y\|$, and therefore

$$P^x(\|X_t - y\| < \varepsilon) \le P^x(\|Y_t - y\| < \varepsilon),$$

for any $\varepsilon > 0$ and $t \ge 0$.

Dividing the above inequality by the volume of the ball $B(y, \varepsilon)$ and passing to the limit with $\varepsilon \searrow 0$, from the continuity of the transition density of the reflecting Brownian motion in the space variable we obtain

$$p_{D_1}(t, x, y) \le p_{D_2}(t, x, y), \qquad t \ge 0,$$

concluding the proof of the theorem. □

As also pointed out by Kendall in [51], we note that in the above theorem the convexity of the larger domain D_1 is not needed in order to derive the validity of condition (2.6.1) in Chavel's conjecture. We can also replace the hypothesis on the convexity of the smaller domain D_2 by the weaker hypothesis that D_2 is a star-shaped domain with respect to either x or y, as follows:

Theorem 2.6.4. *Let $D_2 \subset D_1 \subset \mathbb{R}^d$ be smooth bounded domains. If for $x, y \in D_2$ there exists a ball B centered at either x or y such that $D_2 \subset B \subset D_1$ and D_2 is star-shaped with respect to the center of the ball, then for all $t \ge 0$ we have*

$$p_{D_1}(t, x, y) \le p_{D_2}(t, x, y). \tag{2.6.3}$$

Proof. We will present an analytic proof which parallels the geometric proof of the previous theorem.

Consider $x, y \in D_2$ arbitrarily fixed and assume that $D_2 \subset B = B(y, R) \subset D_1$ for some $R > 0$ and D_2 is a star-shaped domain with respect to y.

By eventually approximating D_2 with star-shaped polygonal domains, it suffices to prove the claim in the case when D_2 is a polygonal star-shaped domain.

Let (X_t, Y_t) be a mirror coupling of reflecting Brownian motions in (D_1, D_2) starting at $x \in D_2$. The idea of the proof is to show that for all times $t \geq 0$, Y_t is at a distance from y is no greater than the distance from X_t to y.

We can reduce the proof to the case when $D_1 = \mathbb{R}^d$ as follows. Consider the sequences of stopping times $(\xi_n)_{n\geq 1}$ and $(\tau_n)_{n\geq 1}$ defined inductively by

$$\begin{aligned}
\tau_0 &= 0, \\
\xi_n &= \inf\{t > \tau_{n-1} : X_t \in \partial D_1\}, \qquad n \geq 1, \\
\tau_n &= \inf\{t > \xi_n : \|X_t - y\| = \|Y_t - y\|\}, \qquad n \geq 1.
\end{aligned}$$

Note that by the ball condition we have $\|X_{\xi_n} - y\| > \|Y_{\xi_n} - y\|$ for any $n \geq 1$, and therefore $\|X_t - y\| \geq \|Y_t - y\|$ for any $n \geq 1$ and any $t \in [\xi_n, \tau_n]$. In order to prove that the same inequality holds on the intervals $[\tau_n, \xi_{n+1}]$ for $n \geq 0$, we proceed as follows.

On the set $\{\tau_n < \infty\}$, the pair $(\widetilde{X}_t, \widetilde{Y}_t) = (X_{\tau_n+t}, Y_{\tau_n+t})$ defined for $t \leq \xi_{n+1} - \tau_n$ is a mirror coupling in $(\mathbb{R}^d, D_2)$ with driving Brownian motion $\widetilde{W}_t = W_{\tau_n+t} - W_{\tau_n}$ (and $\widetilde{Z}_t = Z_{\tau_n+t} - Z_{\tau_n}$), and starting points $(\widetilde{X}_0, \widetilde{Y}_0) = (X_{\tau_n}, Y_{\tau_n})$ independent of the filtration of $\widetilde{B_t}$ (see Remark 2.4.4). In order to prove the claim it suffices therefore to show that for any points $u \in \mathbb{R}^d$ and $v \in \overline{D_2}$ with $\|u - y\| = \|v - y\|$, the mirror coupling (X_t, Y_t) in $(\mathbb{R}^d, D_2)$ with starting points $(X_0, Y_0) = (u, v)$ verifies

$$\|X_t - y\| \geq \|Y_t - y\|, \qquad t \geq 0. \tag{2.6.4}$$

Consider therefore a mirror coupling (X_t, Y_t) in $(\mathbb{R}^d, D_2)$ with starting points $(X_0, Y_0) = (u, v) \in \mathbb{R}^d \times \overline{D_2}$ satisfying $\|u - y\| = \|v - y\|$.

If $u = v$, from the construction of the mirror coupling it follows that $X_t = Y_t$ until the process Y_t hits the boundary of D_2, and therefore the inequality in (2.6.4) holds for these values of t. After the process Y_t hits a bounding hyperplane of D_2, by Lemma 2.4.8 it follows that in an appropriate coordinate system Y_t is given by $Y_t = (X_t^1, \ldots, X_t^{d-1}, |X_t^d|)$, until the time σ when the process Y hits a different bounding hyperplane of D_2, and therefore the inequality in (2.6.4) is again verified for the corresponding values of t (in the chosen coordinate system we must have $y = (y^1, \ldots, y^d)$ with $y^d > 0$, and therefore $\|X_t - y\|^2 - \|Y_t - y\|^2 = 2y^d\left(|X_t^d| - X_t^d\right) \geq 0$). If at time σ the processes are coupled (i.e. $X_\sigma = Y_\sigma \in \partial D_2$), we can apply the above argument inductively, and find a time σ_1 when the processes are decoupled and $\|X_t - y\| \geq \|Y_t - y\|$ for all $t \leq \sigma_1$.

The above discussion shows that without loss of generality we may further reduce the proof of the claim to the case when $(u, v) \in \mathbb{R}^d \times \overline{D_2}$ with $u \neq v$ and $\|u - y\| \geq \|v - y\|$. Also, the above discussion shows that it is enough to prove (2.6.4) for all values of $t \leq \zeta$, where $\zeta = \inf\{s > 0 : X_s = Y_s\}$ is the first coupling time.

The mirror coupling defined by (2.4.1) – (2.4.3) becomes in the case

$$X_t = u + W_t \tag{2.6.5}$$

$$Y_t = v + Z_t + \int_0^t \nu_{D_2}(Y_s)\, dL_s^Y \tag{2.6.6}$$

$$Z_t = \int_0^t G(Y_s - X_s)\, dW_s \tag{2.6.7}$$

where G is given by (2.4.5). In order to prove the claim we will show that

$$R_t = \|X_t - y\|^2 - \|Y_t - y\|^2 \geq 0, \qquad t \leq \zeta, \tag{2.6.8}$$

where ζ is the first coupling time.

Using the Itô formula it can be shown that the process R_t verifies the stochastic differential equation

$$R_t = R_0 - 2\int_0^t R_s \frac{Y_s - X_s}{\|Y_s - X_s\|^2} \cdot dW_s - 2\int_0^t (Y_s - y) \cdot \nu_{D_2}(Y_s)\, dL_s^Y, \quad t \leq \zeta. \tag{2.6.9}$$

The process $B_t = -2\int_0^t \frac{Y_s - X_s}{\|Y_s - X_s\|^2} \cdot dW_s$ is a continuous local martingale on $[0, \zeta)$, with quadratic variation

$$A_t = 4\sum_{i=1}^{d} \int_0^{t\wedge\zeta} \frac{(Y_s^i - X_s^i)^2}{\|Y_s - X_s\|^4} ds = \int_0^{t\wedge\zeta} \frac{4}{\|Y_s - X_s\|^2} ds, \qquad t \geq 0, \tag{2.6.10}$$

and therefore by Lévy's characterization of Brownian motion it follows that $\widetilde{B}_t = B_{\alpha_t \wedge \zeta}$ is a 1-dimensional Brownian motion (possibly stopped at time ζ, if $A_\zeta < \infty$), where the time change $\alpha_t = \inf\{s \geq 0 : A_s > t\}$ is the inverse of the nondecreasing process A_t.

Setting $\widetilde{X}_t = X_{\alpha_t \wedge \zeta}$, $\widetilde{Y}_t = Y_{\alpha_t \wedge \zeta}$, $\widetilde{R}_t = R_{\alpha_t \wedge \zeta}$ and $\widetilde{L}_t^Y = L_{\alpha_t \wedge \zeta}^Y$, from (2.6.9) we obtain

$$\widetilde{R}_t = \widetilde{R}_0 + \int_0^t \widetilde{R}_s d\widetilde{B}_s - \int_0^t (\widetilde{Y}_s - y) \cdot \nu_{D_2}\left(\widetilde{Y}_s\right) d\widetilde{L}_s^Y, \qquad t \geq 0. \tag{2.6.11}$$

Since the polygonal domain D_2 is assumed star-shaped with respect to the point y, geometric considerations show that

$$(z - y) \cdot \nu_{D_2}(z) \leq 0, \tag{2.6.12}$$

for all the points $z \in \partial D_2$ for which the inside pointing normal $\nu_{D_2}(z)$ at the boundary point z of D_2 is defined, that is for all points $z \in \partial D_2$ not lying on the intersection of two bounding hyperplanes of D_2. Since the reflecting Brownian motion Y_t does not hit the set of these exceptional points with positive probability, we may assume that the above condition is satisfied for all points, and therefore

$$(\widetilde{Y}_s - y) \cdot \nu_{D_2}(\widetilde{Y}_s) \leq 0 \qquad \text{a.s,} \tag{2.6.13}$$

for all times $s \geq 0$ when $\widetilde{Y}_s \in \partial D_2$.

Since $\widetilde{L}_t^Y$ is a nondecreasing process of $t \geq 0$, a standard comparison argument for solutions of stochastic differential equations shows that the solution $\widetilde{R}_t$ of (2.6.11) satisfies $\widetilde{R}_t \geq \rho_t$ for all $t \geq 0$, where ρ_t is the solution of the stochastic differential equation

$$\rho_t = \widetilde{R}_0 + \int_0^t \rho_s d\widetilde{B}_s, \qquad t \geq 0. \tag{2.6.14}$$

The last equation has the explicit solution $\rho_t = R_0 e^{\widetilde{B}_t - \frac{1}{2}t}$, and since by hypothesis $R_0 = \|u - y\|^2 - \|v - y\|^2 \geq 0$, we obtain

$$R_{\alpha_t \wedge \zeta} = \widetilde{R}_t \geq \rho_t = \widetilde{R}_0 e^{\widetilde{B}_t - \frac{1}{2}t} \geq 0, \qquad t \geq 0, \tag{2.6.15}$$

and therefore $R_t = \|X_t - y\|^2 - \|Y_t - y\|^2 \geq 0$ for all $t \leq \zeta$, concluding the proof of the claim.

By the initial remarks, it follows that if (X_t, Y_t) is a mirror coupling in (D_1, D_2) with starting point $X_0 = Y_0 = x$, then

$$\|X_t - y\| \geq \|Y_t - y\|, \qquad t \geq 0. \tag{2.6.16}$$

As in the proof of the last theorem, this shows that $p_{D_1}(t, x, y) \leq p_{D_2}(t, x, y)$ for all $t \geq 0$, concluding the proof. □

We have chosen to carry out the construction of the mirror coupling in the case of smooth domains with $\overline{D_2} \subset D_1$ and D_2 convex, having in mind the application to Chavel's conjecture. However, although the technical details can be considerably longer, it is possible to construct the mirror coupling in a more general setup.

For example, in the case when D_1 and D_2 are disjoint domains, none of the difficulties encountered in the construction of the mirror coupling occur (the possibility of coupling/decoupling), so the constructions extends immediately to this case.

The two key ingredients in our construction of the mirror coupling were the hypothesis $\overline{D_2} \subset D_1$ (needed in order to reduce by a localization argument the construction to the case $D_1 = \mathbb{R}^d$) and the hypothesis on the convexity of the inner domain D_2 (which allowed us to construct a solution of the equation of the mirror coupling in the case $D_1 = \mathbb{R}^d$).

Replacing the first hypothesis by the condition that the boundaries ∂D_1 and ∂D_2 are not tangential (needed for the localization of the construction of the mirror coupling) and the second one by condition that $D_1 \cap D_2$ is a convex domain, the arguments in the present construction can be modified in order to give rise to a mirror coupling of reflecting Brownian motion in (D_1, D_2) (see Figure 2.9).

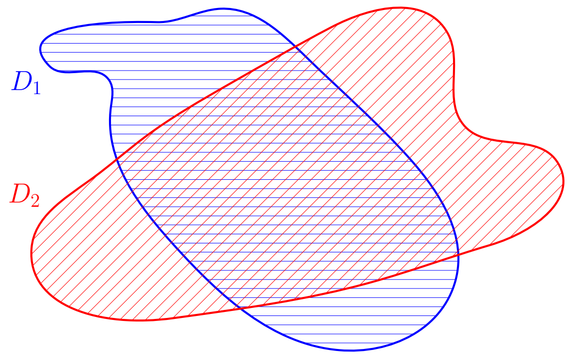

Figure 2.9: Generic smooth domains $D_{1,2} \subset \mathbb{R}^d$ for the mirror coupling: D_1, D_2 have non-tangential boundaries and $D_1 \cap D_2$ is a convex domain.

Remark 2.6.5. *Even though the construction of the mirror coupling was carried out without the additional assumption on the convexity of the inner domain D_2 in the case when D_2 is a polygonal domain (see Theorem 2.4.9), we cannot extend the construction of the mirror coupling to the general case of smooth domains $D_2 \subset D_1$.*

This is due to the fact that the stochastic differential equation which defines the mirror coupling has a singularity (discontinuity) when the processes couple, and we cannot prove the convergence of solutions in the approximating domains (as in the proof of Theorem 2.4.1). The convexity of the inner domain is an essential argument for this proof, which allowed us to handle the discontinuity of the stochastic differential equation which defines the mirror coupling: considering an increasing sequence of approximating domains $D_n \nearrow D_2$, the convexity of D_2 was used to show that if the coupling occurred in the case of the mirror coupling in $(\mathbb{R}^d, D_N)$, then coupling also occurred in the case of the mirror coupling in $(\mathbb{R}^d, D_n)$, for all $n \geq N$.

It is easy to construct an example of a non-convex domain D_2 and a sequence of approximating domains $D_n \nearrow D_2$ such that the mirror coupling (X_t, Y_t^n) in $(\mathbb{R}^d, D_n)$ does not have the above-mentioned property, and therefore we cannot prove the existence of the mirror coupling using the same ideas as in Theorem 2.4.1. However, this does not imply that the mirror coupling cannot be constructed by other methods in a more general setup.

We conclude with some remarks on the non-uniqueness of the mirror coupling in general domains. To simplify the ideas, we will restrict to the 1-dimensional case when $D_2 = (0, \infty) \subset D_1 = \mathbb{R}$.

Fixing $x \in (0, \infty)$ as starting point of the mirror coupling (X_t, Y_t) in (D_1, D_2), the equations of the mirror coupling are

$$X_t = x + W_t \tag{2.6.17}$$

$$Y_t = x + Z_t + L_t^Y \tag{2.6.18}$$

$$Z_t = \int_0^t G(Y_s - X_s)\, dW_s \tag{2.6.19}$$

where in this case

$$G(z) = \begin{cases} -1, & \text{if } z \neq 0 \\ +1, & \text{if } z = 0 \end{cases}.$$

Until the hitting time $\tau = \{s > 0 : Y_s \in \partial D_2\}$ of the boundary of ∂D_2 we have $L_t^Y \equiv 0$, and with the substitution $U_t = -\frac{1}{2}(Y_t - X_t)$, the stochastic differential for Y_t becomes

$$U_t = \int_0^t \frac{1 - G(Y_s - X_s)}{2} dW_s = \int_0^t \sigma(U_s) dW_s, \tag{2.6.20}$$

where

$$\sigma(z) = \frac{1 - G(z)}{2} = \begin{cases} 1, & \text{if } z \neq 0 \\ 0, & \text{if } z = 0 \end{cases}.$$

By a result of Engelbert and Schmidt ([35]) the solution of the above problem is not even weakly unique, for in this case the set of zeroes of the function σ is $N = \{0\}$ and σ^{-2} is locally integrable on $\mathbb{R}$.

In fact, more can be said about the solutions of (2.6.20) in this case. It is immediate that both $U_t \equiv 0$ and $U_t = W_t$ are solutions to 2.6.20, and it can be shown that an arbitrary solution can be obtained from W_t by delaying it when it reaches the origin (sticky Brownian motion with sticky point the origin).

Therefore, until the hitting time τ of the boundary, we obtain as solutions

$$Y_t = X_t = x + W_t \tag{2.6.21}$$

and

$$Y_t = X_t - 2W_t = x - W_t, \tag{2.6.22}$$

and an intermediate range of solutions, which agree with (2.6.21) for some time, then switch to (2.6.22) (see [69]).

Correspondingly, this gives rise to mirror couplings of reflecting Brownian motions for which the solutions stick to each other after they have coupled (as in (2.6.21)), or they immediately split apart after coupling (as in (2.6.22)), and there is a whole range of intermediate possibilities. The first case can be referred to as *sticky* mirror coupling, the second as *non-sticky* mirror coupling, and the intermediate possibilities as *weak/mild sticky* mirror coupling.

The same situation occurs in the general setup in $\mathbb{R}^d$, and it is the cause of lack uniqueness of the stochastic differential equations which defines the mirror coupling. In the present chapter we detailed the construction of the sticky mirror coupling, which we considered to be the most interesting, both from the point of view of the construction and of the applications, although the other types of mirror coupling might prove useful in other applications.

2.7 Open problems

Chapter 3

Fixed-distance coupling of Reflecting Brownian motions

In this paper we introduce three Markovian couplings of Brownian motions on smooth Riemannian manifolds without boundary which sit at the crossroad of two concepts. The first concept is the one of shy coupling put forward in [16] and the second concept is the lower bound on the Ricci curvature and the connection with couplings made in [84].

The first construction is the shy coupling, the second one is a fixed-distance coupling and the third is a coupling in which the distance between the processes is a deterministic exponentially function of time.

The simplest nontrivial manifold is the 2-dimensional sphere in $\mathbb{R}^3$, and in this case we give the explicit construction of all three types of couplings mentioned above and at first we use an extrinsic approach. Next, we construct part of these couplings on manifolds of constant curvature, this time using the intrinsic geometry.

Then we prove a full result which shows that an arbitrary Riemannian manifold satisfying some technical conditions supports shy couplings. Moreover, if in addition the Ricci curvature is non-negative, there exist fixed-distance couplings. Furthermore, if the Ricci curvature is bounded below by a positive constant, then there exists a coupling of Brownian motions for which the distance between the processes is deterministic and exponentially decaying. The constructions use the intrinsic geometry, and relies on an extension of the notion of frames which plays an important role for even dimensional manifolds.

As an application of the fixed-distance coupling we derive a maximum principle for the gradient of harmonic functions on manifolds with non-negative Ricci curvature.

3.1 Introduction

A first motivation of the study in the present chapter was the interest in the following (stochastic) modification of the classical *Lion and Man problem* of Rado ([56]) on manifolds. Consider a Brownian Lion X_t and a Brownian Man Y_t running on a d-dimensional Riemannian manifold M, for example the unit sphere in $\mathbb{R}^3$.

Problem 3.1.1 (Finite coupling time)**.** *Can the Lion capture the Man?*

More precisely, given two distinct starting points $x, y \in M$ and a Brownian motion Y_t on M starting at y, can one find a Brownian motion X_t on M starting at x such thatthe coupling time $\tau = \inf\{t \geq 0 : X_t = Y_t\}$ is almost surely finite (or almost surely bounded)? A weaker version of this problem is whether for a given $\epsilon > 0$ and a given Brownian motion Y_t on M starting at y one can find a Brownian motion X_t on M starting at x such that $\tau = \inf\{t \geq 0 : d(X_t, Y_t) = \epsilon\}$ is almost surely finite (or almost surely bounded). Here $d(x, y)$ stands for the Riemannian distance on M.

Problem 3.1.2 (ε-shy coupling)**.** *Can the Man avoid being eaten by the Lion indefinitely?*

More precisely, given two distinct starting points $x, y \in M$ and a Brownian motion X_t on M starting at x, can one find a Brownian motion Y_t on M starting at y such that almost surely $X_t \neq Y_t$ for all $t \geq 0$? A stronger version of the question is whether the Brownian motion Y_t can be chosen in such a way that there exists an $\epsilon > 0$ such that almost surely $d(X_t, Y_t) \geq \varepsilon$ for all $t \geq 0$.

A second motivation of the present work is related to the notion of *shy coupling* of Brownian motions introduced in [16] and subsequently studied in [17] and [52]. A shy coupling is a coupling for which, with positive probability, the distance between the two processes stays positive for all times. A stronger version of shyness (ϵ-shyness, $\epsilon > 0$), which we will use in this chapter, asserts that with positive probability the distance between the processes is greater than ϵ for all times.

We note that on the unit sphere S^2, there is an immediate affirmative answer to Problem 3.1.1: one can define X_t as the symmetric of Y_t with respect to the plane of symmetry of x and y. Since the Brownian motion Y_t hits this plane in finite time, τ is finite almost surely, so the Lion is sure to capture the Man in finite time.

The above mentioned coupling is known in the literature as the *mirror coupling* (see Chapter 2 for its construction in the case of smooth Euclidean domains), and it was introduced by Lindvall and Rogers [55] for processes defined on Euclidean spaces, and by Cranston in [32] and Kendall [50] in the case of processes defined on manifolds, the so-called *Cranston-Kendall mirror coupling.* As an application of this coupling, it can be shown ([50]) that in the case of manifolds with Ricci curvature bounded uniformly from below by a positive constant, the Man and the Lion must meet in finite time. Other related results regarding couplings of Brownian moptions can be found in [1], [2], and [81].

A synthetic notion of a lower bound on the Ricci curvature was settled in [57, 82, 83] and is a very useful tool in analysis on measure metric spaces. On the other hand, the notion of couplings and lower bound on Ricci curvature was pioneered in [61] and is particularly good for defining lower bounds on Ricci curvature in discrete spaces as it is for instance pointed out in [31, 54].

In this spirit, a third motivation of our work comes from [84, Corollary 1.4] which states the following.

Corollary 3.1.3. *On a complete Riemannian manifold M the Ricci tensor satisfies $Ric \geq k$ if and only if there exits a conservative Markov process $(\Omega, \mathcal{A}, \mathbb{P}^z, Z_t)_{z \in M \times M, t \geq 0}$ with values in $M \times M$ such that the coordinate processes $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ are Brownian motions on M and such that for all $z = (x, y)$ and all $t \geq 0$,*

$$d(X_t, Y_t) \leq e^{-kt/2} d(x, y), \quad \mathbb{P}^z - a.s. \tag{3.1.1}$$

A natural question, and one of our interests in the present chapter, is to see if one can find couplings of Brownian motions X_t, Y_t for which the equality in the inequality (3.1.1) is attained. For instance, in the case $k = 0$ this amounts to finding a fixed-distance coupling (a particular case of the strong version of shy coupling).

The structure of this chapter is the following. In Section 3.2 we introduce the notations and the basic results needed in the sequel. Next, in Section 3.3, we present a result about the existence of fixed-distance couplings on $\mathbb{R}^n$. Here we show that the only fixed-distance coupling in $\mathbb{R}^n$ case is the trivial one, namely the translation coupling, and that there is no distance-decreasing coupling in this case. This is to be contrasted to the fact that in the case of 2-dimensional sphere S^2 (presented in the following section) it is possible to construct distance-decreasing couplings.

In Section 3.4 we focus on the case of the 2-dimensional sphere S^2, and we prove that in this case it is possible to construct a fixed-distance coupling, a distance-decreasing coupling, and a distance-increasing coupling. The construction is carried out by using two main ingredients: Stroock's representation of spherical Brownian motion and Kendall's characterization of co-adapted couplings of Brownian motions in Euclidean spaces (see [52]). From a differential geometric perspective this construction is extrinsic, in the sense that the sphere S^2 is seen as a submanifold of $\mathbb{R}^3$, and we take advantage of this in order to reduce the problem at hand to that of finding unitary matrices in $\mathbb{R}^3$ satisfying certain conditions. The intriguing part about this construction is that the same argument does not extend to higher dimensional spheres.

In the last section of this chapter (Section 3.5) we present two general results. The first is in the case of manifolds of constant curvature in any dimension (Theorem 3.5.1), in which we prove the existence of fixed distance / fast approaching / fast repelling Brownian couplings depending on the sign of the curvature, and the second is in the case of complete d-dimensional Riemannian manifolds M with positive injectivity radius, for which the Ricci curvature is uniformly bounded from below and the sectional curvature uniformly bounded from above (Theorem 3.5.3), in which we prove the existence of shy couplings. Moreover, in this last case we also show that if the Ricci curvature is in addition non-negative, we can also construct fixed-distance couplings, and if the Ricci curvature is bounded from below by a positive constant, then we can also construct fast approaching couplings, for which the distance between processes decays exponentially fast to 0. We conclude with some applications of the couplings constructed in this chapter, by giving a resolution of the two problems presented in the beginning of this section, and a maximum principle for the gradient of harmonic functions on manifolds.

3.2 Preliminaries

We identify the vectors in $\mathbb{R}^3$ with the corresponding 3×1 column matrices, and for a vector $x \in \mathbb{R}^3$ we denote by x' the transpose of x. The dot product of two vectors $x, y \in \mathbb{R}^3$ can be written in terms of matrix multiplication as $x \cdot y = x'y$. The Euclidian length of a vector $x \in \mathbb{R}^3$ is $\|x\| = \sqrt{x'x}$.

We denote by $S^2 = \left\{x \in \mathbb{R}^3 : \|x\| = 1\right\}$ the unit sphere in $\mathbb{R}^3$, and for $x, y \in S^2$ we let $d(x, y)$ be the length of the geodesic joining x and y on S^2 (the length of the smaller of the two arcs of a great circle containing x and y, that is $d(x, y) = \arcsin\sqrt{1 - (x'y)^2} = 2\arcsin\left(\frac{1}{2}\|x - y\|\right)$).

There are various ways of describing the spherical Brownian motion on S^2, that is the Brownian motion on S^2 (see for example [19]). In what follows we exploit the Stroock's representation of spherical Brownian motion ([79]), as the solution X_t of the Itô stochastic differential equation

$$X_t = X_0 + \int_0^t (I - X_s X_s')\, dB_s - \int_0^t X_s ds, \tag{3.2.1}$$

where B_t is a 3-dimensional Brownian motion. The last term above may be thought as the pull needed in order to keep X_t on the surface of S^2.

Given two non-parallel vectors $x, y \in S^2$ (i.e. $y \neq \pm x$), we denote by $R_{x,y}$ the 3×3 rotation matrix with axis $u = x \times y$ (the cross product of the vectors x and y) and angle $\theta \in (0, \pi)$ equal to the angle between the vectors x and y, so in particular $R_{x,y}u = u$ and $R_{x,y}x = y$. It is known that $R_{x,y}$ is an orthogonal matrix ($R^{-1} = R'$) and the following (Rodrigues' rotation) formula holds

$$R_{x,y} = \cos\theta I + [u]_\times + \frac{1}{1 + \cos\theta} u \otimes u, \tag{3.2.2}$$

where $[u]_\times = yx' - xy'$ is the cross-product matrix of $u = x \times y$, $\otimes$ denotes the tensor product ($u \otimes u = uu'$) and I denotes the 3×3 identity matrix. Note that the above formula differs slightly from the usual one, due to the fact that we do not require the axis u to be a unit vector.

When $y = \pm x$, the cross product $u = x \times y$ is the zero vector, so the rotation matrix $R_{x,\pm x}$ is not well defined in this case. However, if we define $R_{x,\pm x} = \pm I$ we see that the $R_{x,\pm x}$ is still a unitary matrix and satisfies $R_{x,\pm x}x = \pm x$. Moreover, taking the limit as $\theta \to 0$ (or π, depending on whether $y = x$ or $y = -x$), we see that (3.2.2) still holds.

By M we denote Riemannian manifold. In this paper all Riemannian manifolds are assumed to be complete. For a given d-dimensional Riemannian manifold M, we use the standard notations from [40] or [81] to denote by $\mathcal{O}(M)$ the orthonormal frame bundle. For a given orthonormal frame U at a point $x \in M$ and $\xi \in \mathbb{R}^d$, $H_\xi(U)$ is the horizontal lift of $U\xi \in T_xM$ at the point $U \in \mathcal{O}(M)$.

We will use the simpler notation of H_i for H_{e_i}, with $\{e_i\}_{i=1,\dots,n}$ denoting the standard basis of $\mathbb{R}^d$.

We collect here some notions from differential geometry which will be used in the sequel. The reader is referred to [33] or [30] for basic notions and results. The curvature tensor R_x at x is $R_x(X,Y) = \nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]}$ and the Ricci tensor is the contraction $Ric_x(X,Y) = \sum_{i=1}^d \langle R_x(X,E_i)E_i, Y\rangle$, where $\{E_i\}_{i=1,\dots,d}$ is any orthonormal basis at x and $X, Y \in T_xM$. This definition of the Ricci tensor does not depend on the choice of orthonormal basis, and in the particular case of surfaces it simplifies to $Ric_x(X,Y) = K_x\langle X,Y\rangle$, where K is the Gauss curvature.

We denote by $d(x,y)$ the Riemannian distance between x and y.

A geodesic on M is a smooth curve $\gamma : [a,b] \to M$ such that $\ddot{\gamma}(s) = 0$ for each $s \in [a,b]$, where the dot represents the covariant derivative along γ. Throughout the paper we assume that the geodesics are running at unit speed. For a point $x \in M$, we define C_x to be the cutlocus of x, that is the set of points $y \in M$ for which there is more then one minimizing geodesic between x and y. We will also use the notation $Cut \subset M \times M$, defined as the set of all points (x,y) which are at each other's cut-locus. For points $x, y \in M$ which are not at each other's cut-locus, we define $\gamma_{x,y}$ to be the unique unit speed minimizing curve joining x and y.

The injectivity radius is the smallest number $i(M)$ such that any point $x \in M$, the exponential map at x is a diffeomorphism on the ball of radius $i(M)$ in the tangent space T_xM.

Given a geodesic γ, a Jacobi field along γ is a vector field $J(s)$ such that

$$\ddot{J}(s) + R_{\gamma(s)}(J(s), \dot{\gamma}(s))\dot{\gamma}(s) = 0, \tag{3.2.3}$$

where the dot represents the derivative along γ.

Given a vector field V along a geodesic γ defined on $[a,b]$, the index form $\mathcal{I}$ associated to it is defined as

$$\mathcal{I}(V,V) = \int_a^b (|\dot{V}(s)|^2 - \langle R_{\gamma(s)}(V(s), \dot{\gamma}(s))\dot{\gamma}(s), V(s)\rangle)ds, \tag{3.2.4}$$

and using polarization $\mathcal{I}$ can be extended to a bilinear form on the space of vector fields along the geodesic γ. In the particular case when J is a Jacobi field, an integration by parts formula shows that

$$\mathcal{I}(J,J) = \langle \dot{J}(b), J(b)\rangle - \langle \dot{J}(a), J(a)\rangle \tag{3.2.5}$$

where $[a,b]$ is the definition interval of γ.

A manifold has constant curvature r if the sectional curvature is r for all choices of the two dimensional plane, that is $\langle R_x(X,Y)Y, X\rangle = r$ for any $x \in M$ and any ortogonal unit vectors $X, Y \in T_xM$. In this case the Ricci curvature simplifies as well as the Jacobi field equation (3.2.3). We record here the calculation, as it will be used later on. Assume that $\gamma_{x,y}$ is the minimal geodesic between the points $x, y \in M$ which are not at each other's cut-locus, $\rho = d(x,y)$ and let $\xi \in T_xM$ and $\eta \in T_yM$ be two unit vectors. Consider $\xi(s)$ the extension of ξ by parallel transport along γ from x to y, and similarly let $\eta(s)$ be the extension of η by parallel transport from y to x. The Jacobi field $J_{\xi,\eta}$ whose value at x is ξ and η at y can be computed as follows

$$J_{\xi,\eta}(s) = w_1(s)\xi(s) + w_2(s)\eta(s) \tag{3.2.6}$$

where w_1, w_2 solve the boundary value problems

$$\begin{cases} \ddot{w}_1 + rw_1 = 0 \\ w_1(0) = 1 \\ w_1(\rho) = 0 \end{cases} \quad \text{and} \quad \begin{cases} \ddot{w}_2 + rw_2 = 0 \\ w_2(0) = 0 \\ w_2(\rho) = 1 \end{cases},$$

whose solutions are

$$w_1(s) = \begin{cases} \frac{\sin(\sqrt{r}(\rho - s))}{\sin(\sqrt{r}\rho)}, & r \neq 0 \\ \frac{\rho - s}{\rho}, & r = 0 \end{cases} \quad \text{and} \quad w_2(s) = \begin{cases} \frac{\sin(\sqrt{r}s)}{\sin(\sqrt{r}\rho)}, & r \neq 0 \\ \frac{s}{\rho}, & r = 0 \end{cases}. \tag{3.2.7}$$

Next, we introduce the main notions regarding couplings. Recall that in general by a coupling we understand a pair of processes (X_t, Y_t) defined on the same probability space, which are separately Markov, that is

$$P\left(X_{s+t} \in A \middle| X_s = z, X_u : 0 \le u \le s\right) = P^z\left(X_t \in A\right)$$
$$P\left(Y_{s+t} \in A \middle| Y_s = z, Y_u : 0 \le u \le s\right) = P^z\left(Y_t \in A\right)$$

for any measurable set A in the state space of the processes.

The notion of *Markovian coupling* as used in [16] requires that in addition to the above, the joint process (X_t, Y_t) is Markov and

$$\begin{aligned} P\left(X_{s+t} \in A \middle| X_s = z, X_u, Y_u : 0 \le u \le s\right) = P^z\left(X_t \in A\right) \\ P\left(Y_{s+t} \in A \middle| Y_s = z, X_u, Y_u : 0 \le u \le s\right) = P^z\left(Y_t \in A\right) \end{aligned} \tag{3.2.8}$$

for any measurable set A in the state space of the processes.

The notion of *co-adapted coupling* (introduced by Kendall, [52]) is the same as the above but without the Markov property of (X_t, Y_t).

The Markovian couplings are easily obtained for instance in the case when the process (X_t, Y_t) is actually a diffusion on the manifold. This would be the ideal case, but we still get a Markovian coupling if we patch together diffusion process in a nice way. For example this will be the case of the main construction on manifolds, where we start the coupling following a diffusion up to a certain stopping time, then, from the point it stopped we run it independently according to another diffusion and then stop this at another stopping time and so on. We do this quietly without further details.

3.3 Distance-decreasing couplings in $\mathbb{R}^d$

In this section we first examine the distance-decreasing couplings in the Euclidean space $\mathbb{R}^d$. To be precise, we want to find all possible co-adapted couplings (X_t, Y_t) of d -dimensional Brownian motions, for which the distance $\|X_t - Y_t\|$ is a (deterministic) non-increasing function of $t \ge 0$.

By a result on co-adapted couplings (Lemma 6 in [52]), a co-adapted coupling (X_t, Y_t) of Brownian motions in $\mathbb{R}^d$ can be represented as

$$Y_t = Y_0 + \int_0^t J_t dX_t + \int_0^t K_t dZ_t,$$

where Z is a d-dimensional Brownian motion independent of X (on a possibly larger filtration), and $J, K \in \mathcal{M}_{d\times d}$ are matrix-valued predictable random processes, satisfying the identity

$$J_t J_t' + K_t K_t' = I, \tag{3.3.1}$$

with I denoting the $d \times d$ identity matrix.

Setting $W_t = X_t - Y_t$ and using Itô's formula we obtain

$$d\left\|W_t\right\|^2 = 2W_t' dW_t + \sum_{i=1}^d d\langle W^i\rangle_t = 2\left(X_t - Y_t\right)'\left(I - J_t\right)dX_t - 2\left(X_t - Y_t\right)' K_t dZ_t + \sum_{i=1}^d d\langle W^i\rangle_t.$$

Using the independence of X and Z, and the relation 3.3.1 we obtain

$$\begin{aligned} \sum_{i=1}^d d\langle W^i\rangle_t &= \operatorname{tr}\left(\left(I - J_t\right)'\left(I - J_t\right) + K_t' K_t\right) dt \\ &= \operatorname{tr}\left(I - J_t - J_t' + J_t' J_t + K_t' K_t\right) dt \\ &= 2\left(\operatorname{tr}(I) - \operatorname{tr}(J_t)\right) dt \\ &= 2\left(d - \operatorname{tr}(J_t)\right) dt. \end{aligned}$$

From the last two equations we arrive at

$$d\|W_t\|^2 = 2(X_t - Y_t)'(I - J_t)\,dX_t - 2(X_t - Y_t)' K_t dZ_t + 2(d - \operatorname{tr}(J_t))\,dt,$$

so the differential of the quadratic variation of the martingale part of $\|W_t\|^2$ is given by

$$\begin{aligned}
&\Big(\big(2(X_t - Y_t)'(I - J_t)\big)\big(2(X_t - Y_t)'(I - J_t)\big)' + \big(2(X_t - Y_t)' K_t\big)\big(2(X_t - Y_t)' K_t\big)'\Big)\,dt \\
=\ & 4(X_t - Y_t)'(I - J_t - J_t' + J_t J_t' + K_t K_t')(X_t - Y_t)\,dt \\
=\ & 4(X_t - Y_t)'(2I - J_t - J_t')(X_t - Y_t)\,dt \\
=\ & 8(X_t - Y_t)'(I - J_t)(X_t - Y_t)\,dt,
\end{aligned}$$

and the differential of the bounded variation part of $\|W_t\|^2$ is given by

$$2(d - \operatorname{tr}(J_t))\,dt.$$

If $\|W_t\|$ is a (deterministic) non-increasing function of t, we must have

$$\operatorname{tr}(J_t) \geq d \qquad \text{and} \qquad (X_t - Y_t)'(I - J_t)(X_t - Y_t) = 0$$

for all $t \geq 0$.

Denoting by $a_{ij} = a_{ij}(t)$ the entries of J_t, observe that

$$\operatorname{tr}(J_t J_t') = \sum_{i,j=1}^{d} a_{ij}^2 \geq \sum_{i=1}^{d} a_{ii}^2 \geq \frac{\left(\sum_{i=1}^{d} a_{ii}\right)^2}{d} = \frac{\operatorname{tr}^2(J_t)}{d} \geq d,$$

with equality if and only if $J_t = I$.

On the other hand, from (3.3.1) it follows that $0 \leq x' J_t' J_t x \leq x'x$ for all $x \in \mathbb{R}^d$, so the eigenvalues $\lambda_i = \lambda_i(t)$ of $J_t J_t'$ satisfy $0 \leq \lambda_i \leq 1$, and therefore $\operatorname{tr}(J_t' J_t) = \sum_{i=1}^{d} \lambda_i \leq d$. Combining with the above we conclude that $\operatorname{tr}(J_t' J_t) = d$, and therefore $J_t = I$ for all $t \geq 0$. Equivalently, this shows that $dY_t = dX_t$ for all $t \geq 0$, or $Y_t = Y_0 - X_0 + X_t$, and we arrive at the following.

Theorem 3.3.1. *In the Euclidean space $\mathbb{R}^d$, $d \geq 1$, the only co-adapted coupling of Brownian motions with deterministic non-increasing distance is the translation coupling.*

As we will see later on in Theorem 3.5.1 there are distance increasing couplings on $\mathbb{R}^d$.

3.4 The 2-dimensional sphere case, the extrinsic approach

In this section we study the couplings of Brownian motions on the unit sphere S^2. The primary interest is the construction of couplings for which the distance between the processes is deterministic. Using Stroock's representation of the spherical Brownian motion, we construct three different couplings, as mentioned in the introduction. In the first one the distance is decaying at an exponential rate, in the second one the distance is increasing to the diameter of the sphere S^2 at an exponential rate, and in the third one, which is the most interesting and intriguing, the distance is constant in time.

We collect the results on the first two couplings mentioned above in the following result, and then treat separately the latter one.

Theorem 3.4.1. *Fix two points $x, y \in S^2$ with $y \neq \pm x$, and consider the spherical Brownian motion X_t on S^2 given by (3.2.1) with $X_0 = x$.*

a) Let Y_t be the solution to

$$Y_t = y + \int_0^t R_s dX_s \tag{3.4.1}$$

where $R_s = R_{X_s,Y_s}$ is the rotation matrix with axis $u_s = X_s \times Y_s$ and angle θ_s equal to the angle between X_s and Y_s (and $R_s = \pm I$ if $Y_s = \pm X_s$). Then Y_t is a spherical Brownian motion on S^2, and

$$\|X_t - Y_t\| = \|y - x\| e^{-t/2}, \qquad t \geq 0. \tag{3.4.2}$$

In particular, $d(X_t, Y_t) = 2 \arcsin\left(\frac{1}{2}\|y - x\| e^{-t/2}\right)$ decreases exponentially fast to 0 as $t \to \infty$.

b) Let $\widetilde{Y}_t$ be the solution to

$$\widetilde{Y}_t = y - \int_0^t R_s dX_s \tag{3.4.3}$$

where $R_s = R_{X_s,-\widetilde{Y}_s}$ is the rotation matrix with axis $u_s = -X_s \times \widetilde{Y}_s$ and angle θ_s equal to the angle between X_s and $-\widetilde{Y}_s$ (and $R_s = \mp I$ if $\widetilde{Y}_s = \pm X_s$). Then $\widetilde{Y}_t$ is a spherical Brownian motion on S^2, and

$$\left\|X_t - \widetilde{Y}_t\right\| = \sqrt{4 - \|y + x\|^2 e^{-t}}, \qquad t \geq 0. \tag{3.4.4}$$

In particular, $d\left(X_t, \widetilde{Y}_t\right) = \pi - 2 \arcsin\left(\frac{1}{2}\|y + x\| e^{-t/2}\right)$ increases exponentially fast to π as $t \to \infty$.

Notice that both (X_t, Y_t) and $\left(X_t, \widetilde{Y}_t\right)$ are both Markovian couplings. In fact they are diffusions on $S^2 \times S^2$.

Proof. Using (3.2.1), we first write

$$dY_t = R_t dX_t = R_t (I - X_t X_t') dB_t - R_t X_t dt.$$

By definition, R_t is a unitary matrix and $R_t X_t = Y_t$ all $t \geq 0$, from which we obtain

$$\begin{aligned} dY_t &= (R_t - R_t X_t X_t') dB_t - R_t X_t dt \\ &= \left(I - (R_t X_t)(R_t X_t)'\right) R_t dB_t - R_t X_t dt \\ &= (I - Y_t Y_t') R_t dB_t - Y_t dt \\ &= (I - Y_t Y_t') d\widetilde{B}_t - Y_t dt, \end{aligned}$$

where $\widetilde{B}_t = \int_0^t R_s dB_s$ is readily seen to be a 3-dimensional Brownian motion. Using again Stroock's characterization of spherical Brownian motion, the first claim follows.

To prove the second claim, we apply the Itô formula to the function $f(z) = z'z$ and to the process $Z_t = Y_t - X_t$. We get

$$d\|Z_t\|^2 = 2Z_t' dZ_t + \sum_{i=1}^3 d\langle Z^i \rangle_t.$$

Next, we'll show that $\|Z_t\|^2$ is a process of bounded variation. To do this, we write

$$\begin{aligned} dZ_t &= d(Y_t - X_t) \\ &= (R_t - I) dX_t \\ &= (R_t - I)(I - X_t X_t') dB_t - (R_t - I) X_t dt \\ &= (R_t - R_t X_t X_t' - I + X_t X_t') dB_t - (Y_t - X_t) dt = M_t dB_t - Z_t dt, \end{aligned} \tag{3.4.5}$$

where $M_t = R_t - R_t X_t X_t' - I + X_t X_t'$. Combining with the above, we get

$$d\|Z_t\|^2 = 2Z_t' M_t dB_t - 2Z_t' Z_t dt + \sum_{i=1}^3 d\langle Z^i \rangle_t, \tag{3.4.6}$$

and in order to prove the claim it suffices to show that $Z_t' M_t \equiv 0$. Notice that

$$\begin{aligned} Z_t' M_t &= (R_t X_t - X_t)' (R_t - R_t X_t X_t' - I + X_t X_t') \\ &= X_t' R_t' R_t - X_t' R_t' R_t X_t X_t' - X_t' R_t' + X_t' R_t' X_t X_t' - X_t' R_t + X_t' R_t X_t X_t' + X_t' - X_t' X_t X_t' \\ &= X_t' - X_t' - X_t' R_t' + X_t' R_t' X_t X_t' - X_t' R_t + X_t' R_t X_t X_t' + X_t' - X_t' \\ &= X_t' (R_t + R_t') (X_t X_t' - I). \end{aligned}$$

Using the representation in (3.2.2) for R_t, since $([u_t]_\times)' = (Y_t X_t' - X_t Y_t')' = -[u_t]_\times$ and $(u_t \otimes u_t)' = (u_t u_t')' = u_t \otimes u_t$, we obtain:

$$\begin{aligned} Z_t' M_t &= 2X_t' \left(\cos\theta_t I + \frac{1}{1+\cos\theta_t} (X_t \times Y_t)(X_t \times Y_t)' \right) (X_t X_t' - I) \\ &= 2\cos\theta_t (X_t' X_t X_t' - X_t') + \frac{2}{1+\cos\theta_t} X_t' (X_t \times Y_t)(X_t \times Y_t)' (X_t X_t' - I) = 0, \end{aligned}$$

where in the last equality we used $X_t' X_t = \|X_t\|^2 = 1$ and $X_t' (X_t \times Y_t) \equiv 0$ (the vector $X_t \times Y_t$ being orthogonal to X_t). It thus follows that $Z_t' M_t \equiv 0$ as we claimed, and therefore $\|Z_t\|^2$ is a process of bounded variation, given by

$$d\|Z_t\|^2 = -2\|Z_t\|^2 dt + \sum_{i=1}^{3} d\langle Z^i \rangle_t. \tag{3.4.7}$$

Finally, note that by using (3.4.5) we can write the last term in the above equation as

$$\sum_{i=1}^{3} d\langle Z^i \rangle_t = \operatorname{tr}(M_t M_t')\, dt,$$

and since X_t is on the unit sphere (so $X_t' X_t = 1$ and $(I - X_t X_t')^2 = I - X_t X_t'$), we can continue with

$$\begin{aligned} \operatorname{tr}(M_t M_t') &= \operatorname{tr}((R_t - I)(I - X_t X_t')^2 (R_t' - I)) = \operatorname{tr}((R_t' - I)(R_t - I)(I - X_t X_t')) \\ &= \operatorname{tr}((2I - R_t' - R_t)(I - X_t X_t')) = 2\operatorname{tr}(I - X_t X_t') - \operatorname{tr}((R_t' + R_t)(I - X_t X_t')) \\ &= 6 - 2X_t' X_t - 2\operatorname{tr}(R_t (I - X_t X_t')) = 4 - 2\operatorname{tr}(R_t) + 2\operatorname{tr}(R_t X_t X_t') = 4 - 2\operatorname{tr}(R_t) + 2Y_t' X_t, \end{aligned} \tag{3.4.8}$$

where in passing to the last line we used that $R_t X_t = Y_t$. Using the fact that the trace of the rotation matrix R_t equals the sum $1 + 2\cos\theta_t$ of its eigenvalues (recall that by construction the angle θ_t of rotation of R_t is the angle between X_t and Y_t), we can conclude that

$$\operatorname{tr}(M_t M_t') = 4 - 2(1 + 2Y_t' X_t) + 2Y_t' X_t = 2 - 2Y_t' X_t = \|X_t - Y_t\|^2 = \|Z_t\|^2.$$

Wrapping things up, we obtained

$$d\|Z_t\|^2 = -\|Z_t\|^2 dt, \qquad t \geq 0.$$

Setting $\tau = \inf\{t \geq 0 : Z_t = 0\}$, the above can be solved as an ordinary differential equation for $t < \tau = \tau(\omega)$ for any path $\omega \in \Omega$, and we obtain the solution

$$\|Z_t\| = \|Y_t - X_t\| = \|y - x\| e^{-t/2}, \qquad t < \tau. \tag{3.4.9}$$

In particular we see that for any $x \neq y$ we have $Z_t \neq 0$ a.s. for all $t \geq 0$, and therefore $\tau = \infty$ a.s. This shows that

$$\|Y_t - X_t\| = \|y - x\| e^{-t/2}, \qquad t \geq 0,$$

which concludes the proof of the first part of the theorem.

To prove the second part of the theorem, note that if $\widetilde{Y}_t$ solves (3.4.3), then $Y_t := -\widetilde{Y}_t$ solves (3.4.1) with y replaced by $-y$ (the process Y_t starts at $-y$ instead of y). If C_t denotes the circle on S^2 of radius 1 and passing through X_t and Y_t (since $x \neq -y$, by the previous proof we have that $X_t \neq Y_t$ for all $t \geq 0$, and thus C_t is well defined), it follows that $\widetilde{Y}_t = -Y_t \in C_t$ for all $t \geq 0$. The second part of the theorem follows now easily from the first part using simple geometric considerations. □

We now proceed to showing the existence of a fixed-distance coupling of Brownian motions on S^2, that is a Markovian coupling (X_t, Y_t) of spherical Brownian motions for which the distance $d(X_t, Y_t)$ is constant for all times $t \geq 0$.

Assume such a coupling exists, and that X_t and Y_t are given by

$$dX_t = (I - X_t X_t')\, dB_t - X_t dt \qquad \text{and} \qquad dY_t = (I - Y_t Y_t')\, dW_t - Y_t dt, \tag{3.4.10}$$

where B_t and W_t are the driving 3-dimensional Brownian motions, and $X_0 = x, Y_0 = y \in S^2$.

By a result on co-adapted couplings of free Brownian motions (assuming that the coupling is co-adapted, see Lemma 6 in [52]), there exist 3×3 matrices J_t and K_t with

$$J_t J_t' + K_t K_t' = I \tag{3.4.11}$$

and a 3-dimensional Brownian motion C_t independent of B_t such that

$$dW_t = J_t dB_t + K_t dC_t. \tag{3.4.12}$$

The idea is now very simple. We want to find the matrix-valued processes J_t and K_t such that the distance between X_t and Y_t does not change with time. The theorem below shows that such a construction is possible, and that in fact the resulting coupling is not only co-adapted, but also a Markovian coupling.

Theorem 3.4.2. *For any points $x, y \in S^2$, there exists a fixed-distance Markovian coupling of Brownian motions on the 2-dimensional unit sphere S^2 starting at x and y. As it turns out, the process (X_t, Y_t) is actually a diffusion on $S^2 \times S^2$.*

Proof. The claim is trivial if $x = \pm y$, so we may assume $x \neq \pm y$.

Denoting $Z_t = X_t - Y_t$, $U_t = I - X_t X_t'$ and $V_t = I - Y_t Y_t'$ (note that U_t and V_t are symmetric matrices, with $U_t^2 = U_t$ and $V_t^2 = V_t$), and using the above equations we obtain

$$dZ_t = U_t dB_t - V_t dW_t - Z_t dt = (U_t - V_t J_t)\, dB_t - V_t K_t dC_t - Z_t dt. \tag{3.4.13}$$

Itô's formula gives after expansion and rearrangements that

$$d\left\|Z_t\right\|^2 = 2Z_t' dZ_t + \sum_{i=1}^{3} d\langle Z^i \rangle_t = 2M_t dB_t + 2N_t dC_t - 2\left\|X_t - Y_t\right\|^2 dt + \sum_{i=1}^{3} d\langle Z^i \rangle_t,$$

with $M_t = -X_t' V_t J_t - Y_t' U_t$ and $N_t = -X_t' V_t K_t$.

The fact that B_t and C_t are independent Brownian motions allows us to compute the quadratic variation of $\left\|Z_t\right\|^2$ as follows:

$$\begin{aligned}
\frac{1}{4} d\left\langle \left\|Z\right\|^2 \right\rangle_t &= (M_t M_t' + N_t N_t')\, dt \\
&= (X_t' V_t (J_t J_t' + K_t K_t') V_t' X_t + X_t' V_t J_t U_t' Y_t + Y_t' U_t J_t' V_t' X_t + Y_t' U_t Y_t)\, dt \\
&= (X_t' V_t X_t + X_t' V_t J_t U_t Y_t + Y_t' U_t J_t' V_t X_t + Y_t' U_t Y_t)\, dt.
\end{aligned}$$

Note that $X_t' V_t X_t = X_t'(I - Y_t Y_t') X_t = X_t' X_t - X_t' Y_t Y_t' X_t = 1 - c_t^2$, where $c_t = Y_t' X_t$, and similarly $Y_t' V_t Y_t = 1 - c_t^2$. Since $X_t' V_t J_t U_t Y_t$ is a real number, it equals its transpose which is $Y_t' U_t J_t' V_t X_t$. Keeping in mind that $X_t' Y_t = Y_t' X_t = c_t$, we also get

$$\begin{aligned}
X_t' V_t J_t U_t Y_t &= X_t'(I - Y_t Y_t') J_t (I - X_t X_t') Y_t = (X_t' - c_t Y_t') J_t (Y_t - c_t X_t) \\
&= X_t' J_t Y_t - c_t X_t' J_t X_t - c_t Y_t' J_t Y_t + c_t^2 Y_t' J_t X_t,
\end{aligned}$$

and therefore

$$\frac{1}{4}d\left\langle \|Z\|^2\right\rangle_t = 2(1-c_t^2+X_t'J_tY_t-c_tX_t'J_tX_t-c_tY_t'J_tY_t+c_t^2Y_t'J_tX_t)dt.$$

If $\|Z_t\|^2$ is to be a constant process, then its quadratic variation must be identically zero, or

$$1-c_t^2+X_t'J_tY_t-c_tX_t'J_tX_t-c_tY_t'J_tY_t+c_t^2Y_t'J_tX_t=0, \tag{3.4.14}$$

and its bounded variation part must also be identically zero. To see what the latter equation is, from (3.4.13) we gain

$$\begin{aligned}
&-2\|X_t-Y_t\|^2 dt+\sum_{i=1}^{3} d\langle Z^i\rangle_t = \\
=\ &\Big(-2\|X_t-Y_t\|^2+\operatorname{tr}\big((U_t-V_tJ_t)(U_t-V_tJ_t)'+(V_tK_t)(V_tK_t)'\big)\Big)dt\\
=\ &(-4(1-c_t)+\operatorname{tr}(U_t-U_tJ_t'V_t'-V_tJ_tU_t'+V_t(J_tJ_t'+K_tK_t')V_t'))dt\\
=\ &(4c_t-2\operatorname{tr}(V_tJ_tU_t))dt,
\end{aligned}$$

which continues with

$$\begin{aligned}
\operatorname{tr}(V_tJ_tU_t) &= \operatorname{tr}(J_tU_tV_t)\\
&= \operatorname{tr}(J_t(I-X_tX_t')(I-Y_tY_t'))\\
&= \operatorname{tr}(J_t-J_tX_tX_t'-J_tY_tY_t'+c_tJ_tX_tY_t')\\
&= \operatorname{tr}(J_t)-\operatorname{tr}(J_tX_tX_t')-\operatorname{tr}(J_tY_tY_t')+c_t\operatorname{tr}(J_tX_tY_t')\\
&= \operatorname{tr}(J_t)-X_t'J_tX_t-Y_t'J_tY_t+c_tY_t'J_tX_t,
\end{aligned}$$

finally arriving at

$$X_t'J_tX_t+Y_t'J_tY_t-c_tY_t'J_tX_t=\operatorname{tr}(J_t)-2c_t. \tag{3.4.15}$$

The above shows that we can construct a co-adapted fixed-distance coupling of Brownian motions on S^2 iff we can find matrices satisfying (3.4.11), (3.4.14) and (3.4.15).

We can go one step further and simplify (3.4.14). Indeed, because of (3.4.15), it is easy to see that equation (3.4.14) is equivalent to

$$1-c_t^2+X_t'J_tY_t-c_t(\operatorname{tr}(J_t)-2c_t)=0.$$

Consequently, the existence of a fixed-distance co-adapted coupling of Brownian motions on S^2 is equivalent to solving for J_t and K_t from the following system

$$\begin{cases}
X_t'J_tY_t=-c_t^2+c_t\operatorname{tr}(J_t)-1 & (\text{or } X_t'V_tJ_tU_tY_t=c_t^2-1)\\
X_t'J_tX_t+Y_t'J_tY_t-c_tY_t'J_tX_t=\operatorname{tr}(J_t)-2c_t & (\text{or } \operatorname{tr}(V_tJ_tU_t)=2c_t)\\
J_tJ_t'+K_tK_t'=I.
\end{cases}$$

To find a solution of this system, we are easing a little bit the notations by dropping for the moment the dependence on t. Hence, given two vectors X, Y on the unit sphere we want to find two 3×3 matrices J and K such that

$$\begin{cases}
X'JY=c\,\operatorname{tr}(J)-1-c^2\\
X'JX+Y'JY-cY'JX=\operatorname{tr}(J)-2c\\
JJ'+KK'=I
\end{cases} \tag{3.4.16}$$

where $c=X'Y$. The first two equations above involve only J. Assuming that we can determine J which satisfies these equations, from the third equation of the system we can find the matrix K such that $KK'=I-JJ'$ if and only if $JJ'\le I$ in the operator sense, i.e. $\xi'JJ'\xi\le 1$ for any unit vector $\xi \in \mathbb{R}^3$, or equivalently $\|J'\xi\| \leq 1$. The latter condition is the same as the operator norm of J' is less than 1, or $\|J\| \leq 1$, since the operator norm of J and J' are the same.

Assume now that $X, Y \in S^2$ with $X \neq \pm Y$ are fixed. We can find an orthogonal matrix $O_{X,Y}$ such that

$$O_{X,Y}e_1 = X \text{ and } O_{X,Y}(ce_1 + \sqrt{1-c^2}e_2) = Y$$

where here $(e_i)_{i=1,2,3}$ is the standard basis in $\mathbb{R}^3$,

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \; e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \text{ and } e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

One way of choosing such a matrix $O_{X,Y}$ is for example by taking

$$O_{X,Y}[e_1, ce_1 + \sqrt{1-c^2}e_2, \sqrt{1-c^2}e_3] = [X, Y, X \times Y],$$

where $[X, Y, Z]$ denotes the matrix whose columns are the vectors X, Y, Z. It is worth mentioning that if the matrix $O_{X,Y}$ is to be orthogonal, then it has to map e_3 into an unitary vector which is collinear to $X \times Y$, which in this case gives $O_{X,Y}e_3 = \pm\frac{1}{\sqrt{1-c^2}}X \times Y$, so there are essentially two choices for the matrix $O_{X,Y}$.

Computing the inverse

$$[e_1, ce_1 + \sqrt{1-c^2}e_2, \sqrt{1-c^2}e_3]^{-1} = \begin{bmatrix} 1 & -\frac{c}{\sqrt{1-c^2}} & 0 \\ 0 & \frac{1}{\sqrt{1-c^2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{1-c^2}} \end{bmatrix},$$

we obtain an explicit formula for $O_{X,Y}$

$$O_{X,Y} = \begin{bmatrix} x_1 & y_1 & x_2y_3 - y_2x_3 \\ x_2 & y_2 & x_3y_1 - y_3x_1 \\ x_3 & y_3 & x_1y_2 - y_1x_2 \end{bmatrix} \begin{bmatrix} 1 & -\frac{c}{\sqrt{1-c^2}} & 0 \\ 0 & \frac{1}{\sqrt{1-c^2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{1-c^2}} \end{bmatrix} = \begin{bmatrix} x_1 & \frac{-cx_1+y_1}{\sqrt{1-c^2}} & \frac{x_2y_3-y_2x_3}{\sqrt{1-c^2}} \\ x_2 & \frac{-cx_2+y_2}{\sqrt{1-c^2}} & \frac{x_3y_1-y_3x_1}{\sqrt{1-c^2}} \\ x_3 & \frac{-cx_3+y_3}{\sqrt{1-c^2}} & \frac{x_1y_2-y_1x_2}{\sqrt{1-c^2}} \end{bmatrix}. \tag{3.4.17}$$

Note that since $X \neq \pm Y$, $c \neq \pm 1$, so the matrix $O_{X,Y}$ is well defined.

Finding a solution J to the system (3.4.16) is equivalent to finding a solution

$$\widetilde{J} = O'_{X,Y}JO_{X,Y} \text{ and } \widetilde{K} = O'_{X,Y}KO_{X,Y}$$

to the system obtained from (3.4.16) by replacing X by e_1, and Y by $ce_1 + \sqrt{1-c^2}e_2$, which becomes

$$\begin{cases} ce_1'\widetilde{J}e_1 + \sqrt{1-c^2}e_1'\widetilde{J}e_2 = c\,\mathrm{tr}(\widetilde{J}) - 1 - c^2 \\ e_1'\widetilde{J}e_1 + c\sqrt{1-c^2}e_1'\widetilde{J}e_2 + (1-c^2)e_2'\widetilde{J}e_2 = \mathrm{tr}(\widetilde{J}) - 2c \\ \widetilde{J}\widetilde{J}' + \widetilde{K}\widetilde{K}' = I. \end{cases}$$

Now let

$$\widetilde{J} = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{bmatrix},$$

which turns the first two equations of the above system into

$$\begin{cases} \sqrt{1-c^2}\alpha_2 - c\beta_2 - c\gamma_3 = -1 - c^2 \\ c\sqrt{1-c^2}\alpha_2 - c^2\beta_2 - \gamma_3 = -2c. \end{cases} \tag{3.4.18}$$

This is a system of two equations with three unknown which can be reduced to

$$\begin{cases} \beta_2 = \frac{\sqrt{1-c^2}\alpha_2 + 1}{c} \\ \gamma_3 = c. \end{cases}$$

Of course the case $c = 0$ needs to be treated separately, in which case, it is obvious that $\alpha_2 = -1$ and $\gamma_3 = 0$.

In the case $c \neq 0$, the simplest matrix $\widetilde{J}$ which satisfies the above conditions is the one whose entries are all 0 except for α_2, β_2, and γ_3, so we may try

$$\widetilde{J} = \begin{bmatrix} 0 & \alpha_2 & 0 \\ 0 & \frac{\sqrt{1-c^2}\alpha_2+1}{c} & 0 \\ 0 & 0 & c \end{bmatrix}.$$

The main restriction now is that we want the operator norm of $\widetilde{J}$ to be at most 1. Because of the block diagonal structure, this is equivalent to

$$\alpha_2^2 + \frac{(\sqrt{1-c^2}\alpha_2 + 1)^2}{c^2} \leq 1$$

or

$$(\alpha_2 + \sqrt{1-c^2})^2 \leq 0,$$

whose solution is $\alpha_2 = -\sqrt{1-c^2}$, and consequently

$$\widetilde{J} = \begin{bmatrix} 0 & -\sqrt{1-c^2} & 0 \\ 0 & c & 0 \\ 0 & 0 & c \end{bmatrix}.$$

This matrix now is well defined also for $c = 0$ and is consistent with the solutions provided by the system (3.4.18).

For the above choice of $\widetilde{J}$ we need to find $\widetilde{K}$ such that

$$\widetilde{J}\widetilde{J}' + \widetilde{K}\widetilde{K}' = I,$$

which reduces to

$$\widetilde{K}\widetilde{K}' = \begin{bmatrix} c^2 & c\sqrt{1-c^2} & 0 \\ c\sqrt{1-c^2} & 1-c^2 & 0 \\ 0 & 0 & 1-c^2 \end{bmatrix}.$$

There are several possible choices here, one of them being

$$\widetilde{K} = \begin{bmatrix} 0 & c & 0 \\ 0 & \sqrt{1-c^2} & 0 \\ 0 & 0 & \sqrt{1-c^2} \end{bmatrix}.$$

Going back to initial problem, we obtain the solution

$$J = O_{X,Y}\widetilde{J}O'_{X,Y} \text{ and } K = O_{X,Y}\widetilde{K}O'_{X,Y}.$$

The only possible problem with this choice of the matrices J and K is that when the particles X and Y get close or antipodal ($X = \pm Y$), the above matrices are undefined because $O_{X,Y}$ does not. However, this does not happen, since by hypothesis $x \neq \pm y$, and with the above choices of J and K the Brownian motions X_t and Y_t are at a fixed-distance (the initial distance).

Finally, since (X_t, Y_t) solves a stochastic differential equation and the matrices J_t and K_t are actually functions of (X_t, Y_t), this means that the process (X_t, Y_t) is in fact a diffusion on $S^2 \times S^2$. This is in fact stronger than mere Markovianity. □

Remark 3.4.3. *It is tempting to extend this argument to higher dimensional spheres. If we follow the same argument we do not have to change anything up to* (3.4.16)*. The attempt on solving* (3.4.16) *was based on arranging the vectors X, Y in a certain position, in other words, make X for instance to be e_1 and Y a linear combination of e_1 and e_2. Since there is essentially (up to a sign choice) a unique perpendicular unit vector to both X and Y, the condition that $O_{X,Y}$*

sends this into e_3 determines the matrix $O_{X,Y}$ perfectly well. In higher dimensions this becomes an issue because there is no canonical choice of the matrix $O_{X,Y}$. Indeed, given two vectors X, Y it is not clear that one can produce a number of vectors which depend smoothly on X, Y and be a basis of the orthogonal complement of the span of X, Y. In more abstract terms, if $V_{k,n}$ is the Stiefel manifold of k orthogonal frames ($k \geq 2$) in $\mathbb{R}^n$, then our problem becomes equivalent to the problem of finding a cross section of the projection $V_{k,n} \to V_{2,n}$. The projection used here is sending the frame $f_1, f_2, \ldots, f_k$ into f_1, f_2. It is know that this is possible (cf. [46, Theorem 1.7]) if and only if $n = 3$ and $k = 3$, and this shows that the proof above works essentially only for the 2-dimensional sphere.

For the higher dimensional spheres, we are going to use a different approach. So far, we have only used the extrinsic approach which is very versatile in the present context, but could become a weakness when one wants to extend it to other manifolds.

Remark 3.4.4. *Without much extra work one can refine the result in Theorem 3.4.2 and show that for any $0 \leq k \leq 1$ and $x, y \in S^2$, there is a Markovian coupling (X_t, Y_t) starting at (x, y) such that $\|X_t - Y_t\| = e^{-kt/2}\|x - y\|$ for all $t \geq 0$.*

If $k < 0$, then there is a coupling (X_t, Y_t) initiated at (x, y) such that $\|X_t - Y_t\| = e^{-kt/2}\|x - y\|$ but only for $0 \leq t \leq \delta$ where δ is a constant determined by k and $\|x - y\|$. Notice that the distance increases exponentially fast in the case $k < 0$, and because of the compactness of S^2 this coupling exists only for short time.

The proof is just a straightforward refinement of the one of Theorem 3.4.2 and is left to the reader. An interesting feature of the proof is that the upper limit of k for which we can get the exponential distance is $k = 1$. This is perhaps a reflection of the fact that the curvature of S^2 is actually 1.

3.5 Extensions and applications

From the geometric point of view, the construction in the previous section is extrinsic, in the sense that we considered the manifold in discussion (the S^2 sphere) imbedded into another manifold ($\mathbb{R}^3$). Since Brownian motion on a manifold is essentially an intrinsic object, it is natural to try to find couplings which are defined in terms of the intrinsic structure of the manifold, that is in terms of its own Riemannian structure, and without embedding it into another manifold.

In [73] we obtained a general intrinsic proof of the existence of couplings in case of manifolds of constant curvature, which present below.

We start with a d-dimensional Riemannian manifold M and we will use the notations introduced in Section 3.4. On M, one can construct the Brownian motion as the solution to a martingale problem associated to the Laplacian (for more details, see [40] or [81]).

Following [40, Section 6.5], we want to define the coupling as a solution to a certain stochastic differential system at the level of orthonormal frame bundle. If U_0 is a given orthonormal frame bundle at x_0 and $V_0 = O_{x_0,y_0}U_0$ is an orthonormal frame bundle at y_0, the system we consider is

$$\begin{cases} dU_t = \sum_{i=1}^d H_i(U_t) \circ dW_t^i \\ dV_t = \sum_{i=1}^d H_i(V_t) \circ dB_t^i \\ dB_t = V_t^{-1} O_{X_tY_t} U_t dW_t \\ X_t = \pi U_t \\ Y_t = \pi V_t \end{cases} \tag{3.5.1}$$

where $O_{x,y}$ is an isometry from T_xM into T_yM.

Among the candidates to the role of the isometry $O_{x,y}$ one is the parallel transport along the minimizing geodesic from x to y, and the resulting coupling is called *synchronous coupling*. The other choice which fits into the picture is the one when $O_{x,y}$ preserves the tangential component of the geodesic from x to y, but changes the sign of the vertical component after performing the parallel transport. Geometrically this is a version of *perverse coupling* and we will refer it so. With this choice, perpendicular to the geodesic the particles move in opposite directions.

To be precise, let $\tau_{x,y}$ be the parallel transport of T_xM into T_yM along the minimizing geodesic γ and let $T_xM = \mathbb{R}\dot\gamma(0) + T_{xy}^{\perp}$ be the orthogonal decomposition of T_xM into the geodesic direction and the perpendicular direction. Similarly let $T_y = \mathbb{R}\dot\gamma(\rho) + T_{yx}^{\perp}$ with $\rho = d(x,y)$. The two choices described above are given by

$$O_{x,y}\dot\gamma(0) = \dot\gamma(\rho) \text{ and } O_{x,y}\xi = \tau_{x,y}\xi, \tag{3.5.2}$$

respectively by

$$O_{x,y}\dot\gamma(0) = \dot\gamma(\rho) \text{ and } O_{x,y}\xi = -\tau_{x,y}\xi, \tag{3.5.3}$$

for any $\xi \in T_{xy}^{\perp}$.

In [73] we obtained the following result which summarizes the main properties of the coupling in the case of constant curvature manifolds.

Theorem 3.5.1. *Let M be a complete d-dimensional Riemannian manifold of constant sectional curvature r. For simplicity consider only the cases $r = -1$, 0 or 1, the general case following by a scaling argument.*

If the starting points x_0, y_0 are chosen such that $\rho_0 < i(M)/2$, then the following hold.

a) For the choice of $O_{x,y}$ as in (3.5.2), the coupling of the Brownian motions satisfies the property that

$$\begin{cases} \text{if } r=-1, & \rho_t \geq \rho_0 \text{ for all } t\geq 0 \\ \text{if } r=0, & \rho_t = \rho_0 \text{ for all } t \geq 0 \\ \text{if } r=1, & 0<\rho_t \leq Ce^{-(d-1)t/2} \text{ for all } t\geq 0 \text{ and some constant } C>0. \end{cases} \tag{3.5.4}$$

b) For the choice of $O_{x,y}$ as in (3.5.3), in all cases,

$$\rho_t \geq \rho_0 \text{ for all } t \geq 0. \tag{3.5.5}$$

Moreover, in the case of the model spaces, namely the hyperbolic space ($r = -1$), the sphere ($r = 1$), and the plane ($r = 0$), for any starting points $x_0 \neq y_0$ which are not at each other's cut-locus, the following hold true.

c) For the choice of $O_{x,y}$ as in (3.5.2),

$$\begin{cases} \text{if } r=-1, & \rho_t = 2\operatorname{arcsinh}(e^{(d-1)/2}\sinh(\rho_0/2)) \text{ for all } t\geq 0 \\ \text{if } r=0, & \rho_t = \rho_0 \text{ for all } t \geq 0 \\ \text{if } r=1, & \rho_t = 2\arcsin(e^{-(d-1)t/2}\sin(\rho_0/2)) \text{ for all } t\geq 0. \end{cases} \tag{3.5.6}$$

d) For the choice of $O_{x,y}$ as in (3.5.3),

$$\begin{cases} \text{if } r=-1, & \rho_t = 2\operatorname{arccosh}(e^{(d-1)t/2}\cosh(\rho_0/2)) \text{ for all } t\geq 0 \\ \text{if } r=0, & \rho_t = \sqrt{\rho_0^2+4(d-1)t} \text{ for all } t \geq 0 \\ \text{if } r=1, & \rho_t = 2\arccos(e^{-(d-1)t/2}\cos(\rho_0/2)) \text{ for all } t\geq 0. \end{cases} \tag{3.5.7}$$

Proof. See [73]. ☐

Remark 3.5.2. *In the case of the sphere S^2, the construction in the above theorem matches the one in Theorem 3.4.1, but also covers the case of spheres in all dimensions, and has the virtue of being intrinsic.*

In [73] we also proved the following general result about the existence of shy coupling on Riemannian manifolds.

Theorem 3.5.3. *Let M be a complete d-dimensional Riemannian manifold with positive injectivity radius and such that for some real number k:*

$$k \leq Ric_x \text{ for all } x \in M \text{ and } \sup_{x \in M} K_x < \infty, \tag{3.5.8}$$

where Ric is the Ricci tensor and K_x stands for the maximum of the sectional curvatures at $x \in M$.

1. *For $k < 0$, there exists $\epsilon, \delta > 0$ such that for any points $x_0, y_0 \in M$ with $d(x_0, y_0) < \epsilon$ we can find a Markovian coupling of Brownian motions X_t, Y_t starting at x_0, y_0 such that $d(X_t, Y_t) \geq d(x_0, y_0)$ for all $t \geq 0$ and $d(X_t, Y_t) = e^{-kt/2} d(x_0, y_0)$ for $0 \leq t \leq \delta$.*

2. *If $k \geq 0$, then there exists $\epsilon > 0$ such that for any $x_0, y_0 \in M$ with $d(x_0, y_0) < \epsilon$, there exists a Markovian coupling of Brownian motions X_t, Y_t starting at x_0, y_0 with $d(X_t, Y_t) = d(x_0, y_0)$ for all $t \geq 0$.*

3. *Moreover, if $k > 0$, then there exists $\epsilon > 0$ such that for any $x_0, y_0 \in M$ with $d(x_0, y_0) < \epsilon$, there exists a Markovian coupling of Brownian motions X_t, Y_t starting at x_0, y_0 with $d(X_t, Y_t) = d(x_0, y_0) e^{-kt/2}$ for all $t \geq 0$.*

Proof. See [73]. □

As a first application of our results, we give a resolution of Problem 3.1.1 and Problem 3.1.2 presented in Section 3.1. Assume M is a Riemannian manifold satisfying the condition in Theorem 3.5.3. According to the theorem it follows that given a Brownian Lion running on M, there is a strategy for the Brownian Man which keeps him at the safe positive distance from the Lion for all times, thus giving an affirmative answer to Problem 3.1.2. Moreover, if the Ricci curvature of M is non-negative, then the Brownian Man can choose a strategy which keeps him at fixed distance from the Brownian Lion.

Theorem 3.5.3 also shows that if the Ricci curvature is bounded below by a positive constant, then given a Brownian Man running on M, the Brownian Lion has a strategy which will bring him arbitrarily close to its meal, thus giving an affirmative answer to Problem 3.1.1.

As another application, we obtain a proof of the following maximum principle for the gradient of harmonic functions.

Theorem 3.5.4. *Let M be a Riemannian manifold with non-negative Ricci curvature and let $u : M \to \mathbb{R}$ be a harmonic function on M. Then, for any relatively compact open set D with smooth boundary we have*

$$\max_{x \in \overline{D}} |\nabla u(x)| = \max_{x \in \partial D} |\nabla u(x)|. \tag{3.5.9}$$

Proof. Fix an arbitrary point $x \in D$. Then there is a geodesic γ such that $\gamma(0) = x$ and

$$v(x) := |\nabla u(x)| = \lim_{h \to 0} \frac{u(\gamma(h)) - u(x)}{h}.$$

In particular, for a small enough $h > 0$, we can consider a fixed distance coupling started at $\gamma(h)$ and x, and run it up until the stopping time ζ, defined as the first time when either of the processes X_t or Y_t hit the boundary of D. On the other hand, since the function u is harmonic, $u(X_t)$ and $u(Y_t)$ are local martingales, and in fact, since u is bounded on D, we can write

$$u(\gamma(h)) - u(x) = \mathbb{E}[u(X_\zeta) - u(Y_\zeta)].$$

In particular, the above equality shows that there must be an ω in the probability space where the processes X_t and Y_t are defined, such that

$$u(\gamma(h)) - u(x) \leq u(X_\zeta(\omega)) - u(Y_\zeta(\omega))$$

Since $d(X_\zeta, Y_\zeta) = d(\gamma(h), x) = h$, we can find a point ξ on the geodesic joining X_ζ with Y_ζ such that

$$u(\gamma(h)) - u(x) \leq u(X_\zeta(\omega)) - u(Y_\zeta(\omega)) \leq |\nabla u(\xi)| h.$$

Since either X_t or Y_t are on the boundary of D, we conclude that ξ is distance h or less from the boundary ∂D.

Thus, as h goes to 0, from the compactness of ∂D, we can find a point $\alpha \in \partial D$ such that

$$v(x) = |\nabla u(x)| \leq |\nabla u(\alpha)| \leq \max_{x \in \partial D} v(x),$$

which concludes the proof. □

For other applications of the couplings constructed in this chapter, see [73].

Chapter 4

A maximum modulus principle for non-analytic functions defined in the unit disk

In the present chapter we present some extensions of the classical maximum modulus principle for analytic functions to certain classes of non-analytic function defined on the unit disk, obtained by the author in [36] and [37]. As corollaries we obtain a new proof of the classical maximum modulus principle for analytic functions, simple conditions on the coefficients of the series development under which the maximum modulus principle holds, as well as as applications to the case of real-valued functions of two variables.

4.1 Introduction

Maximum principles are important tools in many areas of mathematics, such as Differential equations, Potential theory, Complex analysis, Harmonic analysis, and so on.

The classical maximum modulus principle in Complex analysis states that the maximum modulus of a non-constant analytic function defined on a simply connected domain cannot be attained at an interior point of the domain. Maximum modulus principle does not apply without the assumption that the function is analytic, as it can be easily seen by considering the function $f(z,\bar{z}) = \frac{1}{2} - z\bar{z}$ defined for $z \in U = \{z \in \mathbb{C} : |z| < 1\}$ (this function attains its maximum modulus at $z = 0$ without being constant in U). Therefore, by removing the hypothesis on the analyticity of the function, one needs to add supplementary hypotheses in order to insure that the maximum modulus principle holds.

In the present chapter we are concerned with extending the maximum modulus principle to the class of non-analytic functions f defined on the open unit disk $U \subset \mathbb{C}$ which have a series expansion of the form

$$f(z,\bar{z}) = \sum_{n=1}^{\infty} f_n(z,\bar{z}), \quad z \in U, \tag{4.1.1}$$

where $f_n = f_n(z,\bar{z})$ are complex functions defined for $z = x + iy \in \overline{U}$, (real positive) homogeneous of degree n and satisfying a certain inequality on the boundary ∂U.

In Theorem 4.3.1 we show that the maximum modulus principle holds for the functions of this class, and moreover that $|f(z,\bar{z})|$ is a radially increasing function in the open unit disk U.

As a consequence (Corollary 4.3.2), by considering the functions $f_n(z,\bar{z}) = a_n z^n$ in Theorem 4.3.1, we obtain a new proof of the maximum modulus principle for analytic functions (with an additional hypothesis on the coefficients of the corresponding Taylor series).

More generally, considering functions of the form

$$f_n(z,\bar{z}) = \sum_{k=0}^{n} a_n z^{k_n} \bar{z}^{n-k_n} \qquad (k_n \in \{0,1,2,\ldots,n\})$$

and using Theorem 4.3.1, we obtain a simple sufficient condition on the coefficients a_n under which the function $f(z,\bar{z})$ defined by (4.1.1) satisfies a maximum modulus principle (see Corollary 4.3.2). Finally, by means of a counterexample, we show that the condition found in Corollary 4.3.2 is sharp, and we conclude with a weaker version of Theorem 4.3.1 (Theorem 4.3.5), but more useful in applications.

In the present paper we give sufficient conditions for a non-analytic function defined in the unit disk to satisfy a maximum modulus principle. We consider the class of non-analytic functions having a power series expansion of the form

$$f(z,\bar{z}) = \sum_{n=1}^{\infty}\sum_{k=0}^{n} a_{kn} z^{n-k}\bar{z}^{k}, \quad z \in U. \tag{4.1.2}$$

In Theorem 4.4.1 we show that under certain conditions on the coefficients a_{kn} the maximum modulus principle holds for $f(z,\bar{z})$; moreover, we show that $|f(z,\bar{z})|$ is an increasing function of $|z|$. The proof uses a result in [36] on the maximum principle for non-analytic functions.

Our choice of the class of functions in 4.1.2 is motivated by the fact that it is a large enough class of functions which includes some important classes of functions. In particular, it includes the class of real-valued functions of two variables having a Taylor series expansion in the whole unit disk (see Theorem 4.4.5).

4.2 Preliminaries

We denote by $U = \{z \in \mathbb{C} : |z| < 1\}$ the open unit disk in $\mathbb{C}$, and for a complex number $z = x + iy \in \mathbb{C}$, we denote by $\bar{z} = x - iy$, $|z| = \sqrt{x^2+y^2}$ the complex conjugate, respectively the modulus of z.

Recall that a function $f : U \to \mathbb{C}$ analytic in U is called *convex* if f is univalent in U and it maps U onto a convex domain in $\mathbb{C}$. The following result gives a sufficient condition for convexity (see for example [74] or [38]).

Lemma 4.2.1. *If $f : U \to \mathbb{C}$ is analytic in U and it has a Taylor series expansion*

$$f(z) = \sum_{n=1}^{\infty} a_n z^n, \quad z \in U,$$

where $a_1 \neq 0$ and the coefficients $a_n \in \mathbb{C}$ satisfy the inequality

$$\sum_{n=2}^{\infty} n^2 |a_n| \le |a_1|, \tag{4.2.1}$$

then f is a convex function.

We also need the following property, which shows that for a convex function f, $|zf'(z)|$ is an increasing function of $|z|$.

Lemma 4.2.2. *If $f : U \to \mathbb{C}$ is a convex function then for any $\theta \in [0,2\pi)$ arbitrarily fixed, $\left|rf'(re^{i\theta})\right|$ is an increasing function of $r \in (0,1)$, that is*

$$r_1 \left|f'(r_1e^{i\theta})\right| < r_2 \left|f'(r_2e^{i\theta})\right|,$$

for any $0 < r_1 < r_2 < 1$.

Proof. For $\theta \in [0, 2\pi)$ arbitrarily fixed, consider the function $\varphi : (0,1) \to \mathbb{R}$ defined by $\varphi(r) = \ln \left| r f'(re^{i\theta}) \right|$.

The function φ is differentiable on $(0,1)$, and we have

$$\begin{aligned}\frac{d}{dr}\varphi(r) &= \frac{\partial}{\partial r}\left(\ln r + \ln \left|f'(re^{i\theta})\right|\right) \\ &= \frac{1}{r} + \frac{\partial}{\partial r}\operatorname{Re}\left(\log f'(re^{i\theta})\right) \\ &= \frac{1}{r} + \operatorname{Re}\left(\frac{f''(re^{i\theta})}{f'(re^{i\theta})}e^{i\theta}\right) \\ &= \frac{1}{r}\left(1 + re^{i\theta}\frac{f''(re^{i\theta})}{f'(re^{i\theta})}\right) \\ &> 0,\end{aligned}$$

for all $r \in (0,1)$, since f is a convex function in U, and therefore it satisfies

$$\operatorname{Re}\left(1 + z\frac{f''(z)}{f'(z)}\right) > 0, \qquad z \in U.$$

Since $\varphi' > 0$ on $(0,1)$, it follows that φ is an increasing function on $(0,1)$, and therefore $\left|rf'(re^{i\theta})\right|$ is also an increasing function of $r \in (0,1)$, concluding the proof. □

4.3 An extended maximum modulus principle

We are now ready to prove the main result, as follows.

Theorem 4.3.1. *Let $f(z,\bar{z})$ defined for $z \in U$ have a series expansion of the form*

$$f(z,\bar{z}) = \sum_{n=1}^{\infty} f_n(z,\bar{z}), \quad z \in U, \tag{4.3.1}$$

where $f_n(z,\bar{z})$ are functions of $z \in \overline{U}$ satisfying

$$f_n(rz, r\bar{z}) = r^n f_n(z,\bar{z}), \tag{4.3.2}$$

for all $z \in \overline{U}$ and real numbers $r > 0$ for which $rz \in \overline{U}$, $n = 1, 2, 3, \ldots$.

If for some $\theta \in [0, 2\pi)$ we have

$$\sum_{n=2}^{\infty} n\left|f_n(e^{i\theta}, e^{-i\theta})\right| \le \left|f_1(e^{i\theta}, e^{-i\theta})\right| \ne 0, \tag{4.3.3}$$

then $f(z,\bar{z})$ is an increasing function of $|z|$ on $\arg z = \theta$, that is

$$|f(z_1,\overline{z_1})| < |f(z_2,\overline{z_2})|, \tag{4.3.4}$$

for any $z_1 = r_1e^{i\theta}$, $z_2 = r_2e^{i\theta} \in U$ with $0 < r_1 < r_2 < 1$.

In particular, if the condition (4.3.3) holds for all $\theta \in [0, 2\pi)$, then $|f|$ is radially increasing in the whole open unit disk U, and it cannot therefore attain its maximum at an interior point of U.

Proof. Consider $z = re^{i\theta} \in U - \{0\}$ arbitrarily fixed, where $r \in (0,1)$ and $\theta \in [0, 2\pi)$.

The series

$$\sum_{n=1}^{\infty} \frac{f_n(z,\bar{z})}{nz^n}u^n$$

has a radius R of convergence given by

$$\begin{aligned} R &= \frac{1}{\limsup_{n\to\infty} \sqrt[n]{\left|\frac{f_n(z,\bar{z})}{nz^n}\right|}} \\ &= \frac{1}{\limsup_{n\to\infty} \sqrt[n]{\left|\frac{f_n(e^{i\theta},e^{-i\theta})\,|z|^n}{nz^n}\right|}} \\ &= \frac{1}{\limsup_{n\to\infty} \sqrt[n]{\left|\frac{f_n(e^{i\theta},e^{-i\theta})}{n}\right|}}. \end{aligned}$$

From the hypothesis (4.3.3) it follows that $\left|f_n(e^{i\theta},e^{-i\theta})\right| \leq \frac{1}{n}\left|f_1(e^{i\theta},e^{-i\theta})\right|$ for all $n = 1,2,3,\ldots$, and therefore we obtain:

$$R \geq \frac{1}{\limsup_{n\to\infty} \sqrt[n]{\frac{1}{n^2}\left|f_1(e^{i\theta},e^{-i\theta})\right|}} = \frac{1}{\limsup_{n\to\infty} \frac{\frac{1}{(n+1)^2}\left|f_1(e^{i\theta},e^{-i\theta})\right|}{\frac{1}{n^2}\left|f_1(e^{i\theta},e^{-i\theta})\right|}} = 1.$$

It follows that the function $F_z : U \to \mathbb{C}$ defined by

$$F_z(u) = \sum_{n=1}^{\infty} \frac{f_n(z,\bar{z})}{nz^n} u^n, \quad u \in U,$$

is analytic in U.

Moreover, using the hypothesis (4.3.3) and the homogeneity of functions $f_n(z,\bar{z})$, we obtain

$$\begin{aligned} \sum_{n=2}^{\infty} n^2 \left|\frac{f_n(z,\bar{z})}{nz^n}\right| &= \sum_{n=2}^{\infty} n \left|\frac{f_n(z,\bar{z})}{z^n}\right| \\ &= \sum_{n=2}^{\infty} n \left|f_n\left(\frac{z}{|z|},\frac{\bar{z}}{|z|}\right)\right| \\ &= \sum_{n=2}^{\infty} n \left|f_n(e^{i\theta},e^{-i\theta})\right| \\ &\leq \left|f_1(e^{i\theta},e^{-i\theta})\right| \\ &= \left|\frac{f_1(z,\bar{z})}{z}\right|, \end{aligned}$$

which shows that the coefficients of the Taylor series expansion of $F_z(u)$ satisfy the hypothesis (4.2.1) of Lemma 1, and therefore $F_z(u)$ is a convex function.

By Lemma 2, it follows that $|uF_z'(u)|$ is an increasing function of $|u| \in (0,1)$, that is

$$|u_1F_z'(u_1)| < |u_2F_z'(u_2)|,$$

for any $u_1 = \rho_1 e^{i\varphi}$ and $u_2 = \rho_2 e^{i\varphi}$ with $0 < \rho_1 < \rho_2 < 1$ and $\varphi \in [0,2\pi)$.

In particular, for $u_1 = z$ and $u_2 = rz$ with $1 < r < \frac{1}{|z|}$, we obtain

$$|zF_z'(z)| < |rzF_z'(rz)|,$$

or equivalent

$$\left| z\sum_{n=1}^{\infty} \frac{f_n(z,\bar{z})}{nz^n} nz^{n-1} \right| < \left| rz\sum_{n=1}^{\infty} \frac{f_n(z,\bar{z})}{nz^n} nr^{n-1}z^{n-1} \right|.$$

By hypothesis (4.3.2) we have $a_n(z,\bar{z})r^n = a_n(rz, r\bar{z})$ for all $n = 1, 2, 3, \ldots$, and the relation above becomes

$$\left| \sum_{n=1}^{\infty} f_n(z,\bar{z}) \right| < \left| \sum_{n=1}^{\infty} f_n(rz, r\bar{z}) \right|,$$

or equivalent

$$|f(z,\bar{z})| < |f(rz, r\bar{z})|,$$

for any $r \in \left(1, \frac{1}{|z|}\right)$. Since $z \in U - \{0\}$ was arbitrarily chosen, this completes the first part of the proof.

The last part of the proof follows from the strict monotonicity of $|f(z,\bar{z})|$ in the radial direction. □

As a first consequence of the above theorem, we obtain a new proof of the maximum modulus principle for (a class) of analytic functions in the unit disk, as follows:

Corollary 4.3.2 (Maximum modulus principle for analytic functions)**.** *If $f : U \to \mathbb{C}$ is analytic in the open unit disk U having a Taylor series expansion*

$$f(z) = \sum_{n=1}^{\infty} a_n z^n, \quad z \in U,$$

where $a_1 \neq 0$ and the coefficients a_n satisfy the inequality

$$\sum_{n=2}^{\infty} n\,|a_n| \leq |a_1| \tag{4.3.5}$$

then $|f(z)|$ is an increasing function of $|z|$.

In particular $|f(z)|$ cannot be attained at an interior point of U.

Proof. The claim follows from Theorem 4.3.1 by considering the functions $f_n(z,\bar{z}) = a_n z^n$. □

More generally, we can obtain a maximum modulus principle for a class of non-analytic functions as follows:

Corollary 4.3.3. *If $f(z,\bar{z})$ defined for $z \in U$ has a series expansion of he form*

$$f(z,\bar{z}) = \sum_{n=1}^{\infty} a_n z^{k_n} \bar{z}^{n-k_n}, \quad z \in U,$$

where $k_n \in \{0, 1, \ldots, n\}$, $a_1 \neq 0$ and the coefficients $a_n \in \mathbb{C}$ satisfy the inequality

$$\sum_{n=2}^{\infty} n\,|a_n| \leq |a_1|, \tag{4.3.6}$$

then $|f(z,\bar{z})|$ is an increasing function of $|z|$.

In particular, $|f|$ cannot attain its maximum at an interior point of U.

Proof. Follows from Theorem 4.3.1 by considering the functions $f_n(z,\bar{z}) = a_n z^{k_n} \bar{z}^{n-k_n}$ with $k_n \in \{0, 1, 2, \ldots, n\}$, $n \geq 1$. □

Example 4.3.4. *For $a \in \mathbb{R}$, consider the function $f_a(z, \bar{z}) = z - az\bar{z}^2 = z - az|z|^2$ defined for $z \in U$.*

Note that we have $a_1 = 1$, $a_3 = a$ and $a_n = 0$ for $n \in N^* - \{1, 3\}$ in the notation of the above corollary, and therefore f_a satisfies the hypothesis (4.3.6) of the previous corollary iff $|a| \leq \frac{1}{3}$. For these values of a it follows that f_a does not attain its maximum modulus inside U (the maximum value of $|f_a(z, \bar{z})|$ for $z \in \overline{U}$ is $1 - a$, and it is attained only on the boundary ∂U of U, as it can easily be checked).

However, for $\frac{1}{3} < a < \frac{4}{3}$, the maximum modulus principle fails for the function $f_a(z, \bar{z})$ (the maximum value of the modulus of f_a is attained inside the unit disk, on the circle $|z| = \frac{1}{\sqrt{3a}}$).

The above example shows that the inequality (4.3.6) cannot be relaxed (it is sharp), in the sense that if we replace the constant 1 in the hypothesis

$$\frac{\sum_{n=2}^{\infty} n |a_n|}{|a_1|} \leq 1$$

of Corollary 4.3.3 by a larger constant, then there exists functions for which the conclusion of the corollary fails.

We conclude with a weaker version of Theorem 4.3.1, more convenient for applications.

Using the fact that the series

$$\zeta(a) = \sum_{n=1}^{\infty} \frac{1}{n^a} < +\infty$$

converges for $a > 0$ and using Theorem 4.3.1, we obtain the following.

Theorem 4.3.5. *Let $f(z, \bar{z})$ defined for $z \in U$ have a series expansion of the form*

$$f(z, \bar{z}) = \sum_{n=1}^{\infty} f_n(z, \bar{z}), \quad z \in U, \tag{4.3.7}$$

where $f_n(z, \bar{z})$ are functions of $z \in \overline{U}$ satisfying

$$f_n(rz, r\bar{z}) = r^n f_n(z, \bar{z}), \tag{4.3.8}$$

for all $z \in \overline{U}$ and real numbers $r > 0$ for which $rz \in \overline{U}$, $n = 1, 2, 3, \ldots$ and

$$f_1(z, \bar{z}) \neq 0$$

for all $z \in \overline{U}$.

If for some $\theta \in [0, 2\pi)$ we have

$$\left|f_n(e^{i\theta}, e^{-i\theta})\right| \leq \frac{1}{n^{2+\alpha}} \frac{\min_{\theta \in [0,2\pi)} \left|f_1(e^{i\theta}, e^{-i\theta})\right|}{\zeta(1+a) - 1}, \quad n = 2, 3, 4, \ldots, \tag{4.3.9}$$

then $f(z, \bar{z})$ is an increasing function of $|z|$ on $\arg z = \theta$, that is

$$|f(z_1, \overline{z_1}| < |f(z_2, \overline{z_2}|, \tag{4.3.10}$$

for any $z_1 = r_1 e^{i\theta}$, $z_2 = r_2 e^{i\theta} \in U$ with $0 < r_1 < r_2 < 1$.

In particular, if (4.3.9) holds for all $\theta \in [0, 2\pi)$, then $|f|$ is radially increasing in the whole open unit disk U, and it cannot therefore attain its maximum at an interior point of U.

4.4 A maximum modulus principle for a class of non-analytic functions defined in the unit disk

In this section we present a maximum modulus principle for a large class of non-analytic functions defined in the unit disk. The main result is the following.

Theorem 4.4.1. *If the function $f(z,\bar{z})$ defined for $z \in U$ has a series expansion of the form:*

$$f(z,\bar{z}) = \sum_{n=1}^{\infty}\sum_{k=0}^{n} a_{kn} z^{n-k}\bar{z}^{k}, \quad z \in U, \tag{4.4.1}$$

where the coefficients $a_{kn} \in \mathbb{C}$ satisfy the inequality

$$\sum_{n=2}^{\infty} n \sum_{k=0}^{n} |a_{kn}| \le ||a_{01}| - |a_{11}|| \ne 0, \tag{4.4.2}$$

then $|f(z,\bar{z})|$ is a radially increasing function in the unit disk, that is

$$|f(z_1,\bar{z}_1)| < |f(z_2,\bar{z}_2)|,$$

for any $z_1 = r_1 e^{i\theta}$, $z_2 = r_2 e^{i\theta} \in U$ with $0 < r_1 < r_2 < 1$.

In particular, $f(z,\bar{z})$ cannot attain its maximum modulus at an interior point of U.

Proof. Consider the functions $f_n(z,\bar{z}) = \sum_{k=0}^{n} a_{kn} z^{n-k}\bar{z}^{k}$ defined for $z \in \overline{U}$, where $n = 1, 2, \ldots$.

We will show that with this choice the hypotheses of Theorem 4.3.1 are satisfied, and therefore the claim of the theorem will follow.

Let us note first that from the definition of the functions $f_n(z,\bar{z})$, they are (positive real) homogeneous of degree n, that is:

$$\begin{aligned} f_n(rz, r\bar{z}) &= \sum_{k=0}^{n} a_{kn} (rz)^{n-k}\, \overline{(rz)}^{k} \\ &= r^n \sum_{k=0}^{n} a_{kn} z^{n-k}\bar{z}^{k} \\ &= r^n f_n(z,\bar{z}), \end{aligned}$$

for all $z \in \overline{U}$ and $r > 0$ for which $rz \in \overline{U}$, and all $n = 1, 2, \ldots$, and therefore the hypothesis (4.3.2) of Theorem 4.3.1 is satisfied.

To verify condition (4.3.3), let us note that from the hypothesis (4.4.2), we have

$$\left|f_n(e^{i\theta}, e^{-i\theta})\right| = \left|\sum_{k=0}^{n} a_{kn} e^{i(n-k)\theta} e^{-ik\theta}\right| \le \sum_{k=0}^{n} \left|a_{kn} e^{i(n-k)\theta} e^{-ik\theta}\right| = \sum_{k=0}^{n} |a_{kn}|,$$

for all $n = 1, 2, \ldots$ and $\theta \in [0, 2\pi)$, and also

$$\left|f_1(e^{i\theta}, e^{-i\theta})\right| = \left|a_{01}e^{i\theta} + a_{11}e^{-i\theta}\right| \ge \left|\left|a_{01}e^{i\theta}\right| - \left|a_{11}e^{-i\theta}\right|\right| = ||a_{01}| - |a_{11}||,$$

for all $\theta \in [0, 2\pi)$.

We obtain therefore

$$\sum_{n=2}^{\infty} n \left|f_n(e^{i\theta}, e^{-i\theta})\right| \le \sum_{n=2}^{\infty}\sum_{k=0}^{n} |a_{kn}| \le ||a_{01}| - |a_{11}|| \le \left|f_1(e^{i\theta}, e^{-i\theta})\right|,$$

for all $\theta \in [0, 2\pi)$, which shows that the hypothesis (4.3.3) of Theorem 4.3.1 is also verified.

The claim of the theorem follows now by using Theorem 4.3.1, concluding the proof. □

Remark 4.4.2. *Let us note that the maximum modulus principle in the previous theorem also holds in the case $|a_{01}| = |a_{11}|$, provided $f(z,\bar{z})$ is not identically zero in U.*

To see this, note that if $|a_{01}| = |a_{11}|$, by using the hypotheses (4.4.1) and (4.4.2) it follows that $a_{kn} = 0$ for all $n = 2, 3, \ldots$ and $k \in \{0, 1, \ldots, n\}$, and therefore we have

$$f(z,\bar{z}) = a_{01}z + a_{11}\bar{z}, \quad z \in U.$$

Also note that since $f(z,\bar{z})$ is not identically zero in U, we have $|a_{01}| = |a_{11}| \neq 0$. Letting $a_{01} = \rho e^{i\alpha}$ and $a_{11} = \rho e^{i\beta}$ with $\rho \in (0,1)$ and $\alpha, \beta \in [0, 2\pi)$, we obtain:

$$\begin{aligned} |f(z,\bar{z})|^2 &= \left|\rho e^{i\alpha}z + \rho e^{i\beta}\bar{z}\right|^2 \\ &= \rho r^2 \left|e^{i(\alpha+\theta)} + e^{i(\beta-\theta)}\right|^2 \\ &= \rho^2 r^2 (2 + 2\cos(\alpha+\theta)\cos(\beta-\theta) + 2\sin(\alpha+\theta)\sin(\beta-\theta)) \\ &= \rho^2 r^2 (2 + 2\cos(\alpha - \beta + 2\theta)) \\ &= 4\rho^2 r^2 \cos^2\left(\theta + \frac{\alpha-\beta}{2}\right), \end{aligned}$$

for all $z = re^{i\theta} \in U$, which shows that $|f(z,\bar{z})|$ is an increasing function of $|z| = r \in (0,1)$, for all $\theta \in [0, 2\pi)$ for which $\cos\left(\theta + \frac{\alpha-\beta}{2}\right) \neq 0$.

However, for the values of θ for which $\cos\left(\theta + \frac{\alpha-\beta}{2}\right) = 0$, we have $\left|f(re^{i\theta}, e^{-i\theta})\right| = 0$, and since $f(z,\bar{z})$ is not identically constant, it follows that the maximum modulus principle still holds for $f(z,\bar{z})$ in this case.

Using the fact that the series $\zeta(a) = \sum_{n=1}^{\infty} \frac{1}{n^a}$ converges for $a > 1$, we can obtain a maximum principle for functions having a series expansion of the form (4.4.1) for which the coefficients satisfy a simple inequality, as follows:

Corollary 4.4.3. *If the function $f(z,\bar{z})$ defined for $z \in U$ has a series expansion of the form:*

$$f(z,\bar{z}) = \sum_{n=1}^{\infty}\sum_{k=0}^{n} a_{kn} z^{n-k}\bar{z}^k, \quad z \in U, \tag{4.4.3}$$

where for some real number $a > 0$ the coefficients $a_{kn} \in \mathbb{C}$ satisfy the inequality

$$\max_{0\le k\le n} |a_k| \le \frac{1}{(n+1)n^{2+a}} \frac{||a_{01}| - |a_{11}||}{\zeta(1+a) - 1} \neq 0, \quad n = 2, 3, \ldots \tag{4.4.4}$$

then $|f(z,\bar{z})|$ is a radially increasing function in U, that is

$$|f(z_1,\bar{z}_1)| < |f(z_2,\bar{z}_2)|, \tag{4.4.5}$$

for any $z_1 = r_1e^{i\theta}$, $z_2 = r_2e^{i\theta} \in U$ with $0 < r_1 < r_2 < 1$.

In particular $f(z,\bar{z})$, cannot attain its maximum modulus at an interior point of U.

Remark 4.4.4. *The inequalities (4.4.4) show essentially that the coefficients a_{kn} converge in absolute value to zero faster than $\frac{1}{n^3}$ as $n \to \infty$, and it provides therefore a large class of functions $f(z,\bar{z})$ for which the maximum principle holds (note that by Remark 4.4.2, the maximum principle still holds in the case $|a_{01}| = |a_{11}|$, provided $f(z,\bar{z})$ is not identically zero in U).*

For the convenience of the reader, in Table 4.1 we listed some approximate values of the Riemann zeta function $\zeta(a)$.

As an application of the previous corollary, we derive a maximum modulus principle for a class of real-valued functions of two real variables defined in the unit disk, as follows.

a	1.1	1.2	1.3	1.4	1.5	2.0
$\zeta(a)$	10.5844	5.9158	3.93195	3.10555	2.61238	$\frac{\pi^2}{6} \approx 1.64493$

Table 4.1: Table of values of Riemann zeta function $\zeta(a)$ for some values of a.

Theorem 4.4.5. *If the real-valued function of two real variables $f = f(x, y) : U \subset \mathbb{R} \to \mathbb{R}$ can be represented by a Taylor series*

$$f(x, y) = \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} C_n^k \frac{\partial^n f}{\partial x^{n-k} \partial y^k}(0, 0) x^{n-k} y^k, \quad (x, y) \in U, \tag{4.4.6}$$

where for some positive real number $a > 0$ we have:

$$\max_{0 \le k \le n} C_n^k \left| \frac{\partial^n f}{\partial x^{n-k} \partial y^k}(0, 0) \right| \le \frac{n!}{(n+1)n^{2+a}} \frac{\left| \left| \frac{\partial f}{\partial x}(0, 0) \right| - \left| \frac{\partial f}{\partial y}(0, 0) \right| \right|}{\zeta(1 + a) - 1} \neq 0,$$

for all $n = 2, 3, \ldots$, then $|f(x, y)|$ is a radially increasing function in U, that is

$$|f(r_1 \cos\theta, r_1 \sin\theta)| < |f(r_2 \cos\theta, r_2 \sin\theta)|, \tag{4.4.7}$$

for any $0 < r_1 < r_2 < 1$ and $\theta \in [0, 2\pi)$.

In particular, $f(x, y)$ cannot attain its maximum modulus at an interior point of U.

Chapter 5

Univalent approximations of analytic functions

When an analytic function is not univalent, it is often of interest to approximate it by univalent functions. In this chapter we introduce a measure of the non-univalency of a function and we derive a method for constructing the best starlike and convex univalent approximations of analytic functions with respect to it, suitable for both practical problems and numerical implementation.

5.1 Introduction

The univalency of an analytic function is an important problem of the Geometric function theory, and there are many sufficient conditions for univalency in the literature (see for example the monographs [34], [74] or [75]). If a function is not univalent, then, in practical problems, it is of interest to find a "best approximation" of it by a univalent function.

In the present chapter we introduce a measure of the non-univalency of an analytic function, and we use it in order to find the best approximation of a normed analytic function in certain subclasses of univalent functions (starlike, respectively convex functions), in the sense of $L^2(U)$ norm. We show that the corresponding problems can be reduced to certain semi-infinite quadratic programming problems, which we solve explicitly in Theorem 5.3.1 and Theorem 5.3.5, thus leading to a method for finding the best starlike, respectively convex approximation: our main results in Theorem 5.4.1 and Theorem 5.5.1 provide constructive algorithms for finding explicitly the measures $\operatorname{dist}(f, \mathcal{S}^*)$ and $\operatorname{dist}(f, \mathcal{K}^*)$ of the (non)starlikeness, respectively (non)convexity of an analytic function (see also Theorem 5.2.3), as well as for finding the corresponding best starlike approximation, respectively the best convex approximation.

The structure of the paper is the following. In Section 5.2 we introduce the measures $\operatorname{dist}(f, \mathcal{U})$, $\operatorname{dist}(f, \mathcal{S})$, and $\operatorname{dist}(f, \mathcal{K})$, which show how far is the function f from being univalent, starlike, respectively convex. In Lemma 5.2.2 we find a convenient representation of $\operatorname{dist}(f, \mathcal{U})$ in terms of the coefficients of the Taylor series of f. Although $\operatorname{dist}(\cdot, \mathcal{U})$ it is not a norm, in Theorem 5.2.3 we show that $\operatorname{dist}(f, \mathcal{U}) = 0$ iff f is a univalent function, so $\operatorname{dist}(f, \mathcal{U})$ is a measure showing how far is the function f from being univalent (the same holds for $\operatorname{dist}(f, \mathcal{S})$ and $\operatorname{dist}(f, \mathcal{K})$).

In Section 5.3, we first present some basic results about the quadratic programming (the Karush-Kuhn-Tucker conditions). In Subsection 5.3.2 we consider a particular quadratic programming problem with an infinite number of variables, for which we show (see Remark 5.3.2) that the Karush-Kuhn-Tucker conditions can be applied. Next, in Theorem 5.3.1 we show that the particular quadratic problem can be solved explicitly: we determine the minimum value of the problem, as well as the extremal function. In Subsection 5.3.3 we consider another particular quadratic problem, for which we also determine explicitly the minimum value of the objective function and the extremum point (Theorem 5.3.5).

In Section 5.4 we apply the results in the previous section in order to find the best starlike approximation of a normed analytic function. The main result is contained in Theorem 5.4.1, which gives an explicit method for constructing the starlike approximation of an analytic function, suitable for numerical implementation and applications. The section concludes with some examples, which show how to construct the starlike approximations for some particular functions. For the values of the parameters involved for which $\operatorname{dist}(f, \mathcal{S}^*)$ is not too large, the numerical results show that the images of the unit disk under the two maps (the original analytic function and its starlike approximation) are close to each other, so the method produces good numerical results.

As an application of Theorem 5.3.5, in the last section we obtain (Theorem 5.5.1) the best convex approximation of a normed analytic function defined in the unit disk. The result gives explicitly the value of $\operatorname{dist}(f, \mathcal{K}^*)$ and of the extremal convex function, so it is again suitable for both numerical implementation and applications. The section concludes with two examples, which show that when $\operatorname{dist}(f, \mathcal{K}^*)$ is not too large, the method of Theorem 5.5.1 produces a good convex approximation of a given analytic function $f \in \mathcal{A}$ as shown in Figure 5.2.

5.2 Univalent approximation of analytic functions

Let $\mathcal{A}$ denote the class of analytic functions $f : U \to \mathbb{C}$ satisfying the normalization condition $f(0) = f'(0) - 1 = 0$, and let $\mathcal{U}$ denote the subclass of $\mathcal{A}$ consisting of univalent functions. Further, let $\mathcal{S}$, $\mathcal{K}$ denote the subclasses of $\mathcal{U}$ consisting of starlike univalent, respectively convex univalent functions in the unit disk U (functions which map the unit disk U univalently onto a starlike, respectively a convex domain).

It is known (see for example [38]) that if $f \in \mathcal{A}$ has the series expansion

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n, \qquad z \in \mathbb{C}, \tag{5.2.1}$$

and the coefficients a_n satisfy the inequality

$$\sum_{n=2}^{\infty} n |a_n| \leq 1 \tag{5.2.2}$$

then $f \in \mathcal{S}$, and if the coefficients a_n satisfy the inequality

$$\sum_{n=2}^{\infty} n^2 |a_n| \leq 1, \tag{5.2.3}$$

then $f \in \mathcal{K}$. We denote by $\mathcal{S}^*$ and $\mathcal{K}^*$ the subclasses of $\mathcal{S}$ and $\mathcal{K}$ defined by (5.2.2), respectively by (5.2.3).

As a measure of the non-univalency of a function we introduce the following.

Definition 5.2.1. *For $f \in A$ we define*

$$\operatorname{dist}(f, \mathcal{U}) = \inf_{g \in \mathcal{U}} \left(\int_U |f(x+iy) - g(x+iy)|^2 \, dxdy \right)^{1/2}, \tag{5.2.4}$$

with similar definitions for $\operatorname{dist}(f, \mathcal{S})$*,* $\operatorname{dist}(f, \mathcal{S}^*)$*,* $\operatorname{dist}(f, \mathcal{K})$*, and* $\operatorname{dist}(f, \mathcal{K}^*)$.

Although $\operatorname{dist}(\cdot, \mathcal{U})$ is not a norm in $\mathcal{A}$ (see Theorem 5.2.3), $\operatorname{dist}(f, \mathcal{U})$ is a measure showing how "far" is the function f from being univalent. Similarly, $\operatorname{dist}(f, \mathcal{S})$, $\operatorname{dist}(f, \mathcal{K})$, $\operatorname{dist}(f, \mathcal{S}^*)$, and $\operatorname{dist}(f, \mathcal{S}^*)$ measure how far is the function f from being starlike, convex, in the class $\mathcal{S}^*$, respectively in the class $\mathcal{K}^*$.

A first preliminary result is the following.

Lemma 5.2.2. *If $f : U \to \mathbb{C}$ is analytic in U and has series expansion*

$$f(z) = \sum_{n=0}^{\infty} a_n z^n, \qquad z \in U, \tag{5.2.5}$$

then

$$\int_U |f(x+iy)|^2 \, dxdy = \pi \sum_{n=0}^{\infty} \frac{|a_n|^2}{n+1}.$$

Proof. For $R \in (0,1)$ arbitrarily fixed, the series in (5.2.5) converges uniformly on $U_R = \{z \in \mathbb{C} : |z| < R\}$. Passing to polar coordinates and using Fubini's theorem, we obtain:

$$\begin{aligned}
\int_{U_R} |f(x+iy)|^2 \, dxdy &= \int_0^R \left(\int_0^{2\pi} \left| \sum_{n=0}^{\infty} a_n r^n e^{in\theta} \right|^2 r dr d\theta \right) \\
&= \int_0^R \left(\int_0^{2\pi} \left(\sum_{n=0}^{\infty} a_n r^n e^{in\theta} \right) \left(\sum_{m=0}^{\infty} \overline{a_m} r^m e^{-im\theta} \right) r dr d\theta \right) \\
&= \int_0^R \left(\sum_{m,n=0}^{\infty} \int_0^{2\pi} a_n \overline{a_m} r^{m+n+1} e^{i(n-m)\theta} d\theta \right) dr \\
&= \int_0^R \left(\sum_{m,n=0}^{\infty} 2\pi a_n \overline{a_m} \delta_{mn} r^{m+n+1} \right) dr \\
&= \pi \sum_{n=0}^{\infty} \frac{|a_n|^2}{n+1} R^{2n+2}.
\end{aligned}$$

Letting $R \nearrow 1$ and using the monotone convergence theorem the result follows. □

If $f \in \mathcal{U}$ is then obviously $\operatorname{dist}(f, \mathcal{U}) = 0$. The following theorem shows that the converse is also true.

Theorem 5.2.3. *For $f \in \mathcal{A}$,* $\operatorname{dist}(f, \mathcal{U}) = 0$ *iff $f \in \mathcal{U}$.*

Proof. The converse implication is obvious. To prove the direct implication we will show that if $\operatorname{dist}(f, \mathcal{U}) = 0$, we can find a sequence of univalent functions $f_n \in \mathcal{U}$ such that $f_n \to f$ uniformly on compact subsets of U, and therefore either f is identically constant in U or $f \in \mathcal{U}$ (the first possibility is however ruled out by the normalization condition).

Since $\operatorname{dist}(f, \mathcal{U}) = 0$, we can find a sequence $(f_n)_{n \geq 1} \subset \mathcal{U}$ such that

$$\int_U |f(x+iy) - f_n(x+iy)|^2 \, dxdy < \frac{\pi}{n}, \qquad n \geq 1.$$

If f and f_n have series expansions given by

$$f(z) = \sum_{m=1}^{\infty} a_m z^m \qquad \text{and} \qquad f_n(z) = \sum_{m=1}^{\infty} a_{n,m} z^m, \qquad z \in U,$$

with $a_1 = a_{n,1} = 1$, using Lemma 5.2.2 we obtain

$$\sum_{m=2}^{\infty} \frac{|a_m - a_{n,m}|^2}{m+1} < \frac{1}{n}, \qquad n \geq 1.$$

For arbitrarily fixed $r \in (0,1)$ and $z_0 \in \overline{U_r}$, we obtain

$$\begin{aligned} |f(z_0) - f_n(z_0)| &\leq \sum_{m=0}^{\infty} |a_m - a_{n,m}| \, |z_0|^m \\ &\leq \sum_{m=0}^{\infty} |a_m - a_{n,m}| \, r^m \\ &\leq \left(\sum_{m=0}^{\infty} \frac{|a_m - a_{n,m}|^2}{m+1} \right)^{1/2} \left(\sum_{m=0}^{\infty} (m+1) \, r^{2m} \right)^{1/2} \\ &\leq \frac{1}{\sqrt{n}} \left(\sum_{m=0}^{\infty} (m+1) \, r^{2m} \right)^{1/2} . \end{aligned}$$

Since the series $\sum_{m=0}^{\infty} (m+1) \, r^{2m}$ converges for any $r \in (0,1)$, the above inequality shows that the sequence f_n converges uniformly to f on $\overline{U_r}$ for any $r \in (0,1)$, so f_n converges to f uniformly on compact subsets of U. Since the functions f_n are univalent in U, the limit function f is either univalent or it is identically constant in U (impossible, by the normalization condition), concluding the proof. □

5.3 Quadratic programming

In this section we recall the Karush-Kuhn-Tucker conditions, specialized for the case of quadratic programming problems, and we use them to solve two particular quadratic programming problems.

5.3.1 The Karush-Kuhn-Tucker conditions

Consider the problem of minimizing

$$f(x) = x^T Q x + cx \tag{5.3.1}$$

under the conditions

$$Ax \leq b \quad \text{and} \quad x \geq 0, \tag{5.3.2}$$

where $x \in \mathbb{R}^n$ are column vectors, $Q \in \mathcal{M}_{n \times n}$ is a symmetric matrix, $A \in \mathcal{M}_{m \times n}$, $b \in \mathcal{M}_{m \times 1}$ and $c \in \mathcal{M}_{1 \times n}$. Further, assume that a feasible solution exists and that the constraint region is bounded.

The above is a particular case of quadratic programming, and it is known that when the objective function $f(x)$ is strictly convex for all feasible points the problem has a unique local minimum which is also the global minimum (a sufficient condition which guarantees the strict convexity of the objective function f is that Q is a positive definite matrix).

The Karush-Kuhn-Tucker conditions below (specialized for the case of the above quadratic programming problem, see [47]) are necessary conditions for a global minimum. If Q is positive definite, they are also sufficient for a global minimum.

Consider the Lagrangian function L for the above quadratic programming problem:

$$L = x^T Q x + cx + \mu (Ax - b) .$$

The Karush-Kuhn-Tucker conditions are the following:

$$\begin{aligned}
\frac{\partial L}{\partial x_i} &\geq 0, \qquad i = 1, \ldots, n \\
\frac{\partial L}{\partial \mu_j} &\leq 0, \qquad j = 1, \ldots, m \\
x_i \frac{\partial L}{\partial x_i} &= 0, \qquad i = 1, \ldots, n \\
\mu (Ax - b) &= 0 \\
x_i &\geq 0, \qquad i = 1, \ldots, n \\
\mu_j &\geq 0, \qquad j = 1, \ldots, m
\end{aligned}$$

5.3.2 A particular quadratic programming problem

Consider the problem of finding

$$\inf \sum_{n=2}^{\infty} \frac{(x_n - a_n)^2}{n+1} \tag{5.3.3}$$

where $(a_n)_{n\geq 2}$ is a given non-negative sequence of real numbers, and the infimum is taken over all non-negative sequences $(x_n)_{n\geq 2}$ of real numbers satisfying

$$\sum_{n=2}^{\infty} n x_n \leq 1. \tag{5.3.4}$$

In this case the objective function $f(x) = \sum_{n=2}^{\infty} \frac{(x_n - a_n)^2}{n+1}$ is a quadratic function and we have the only constrained inequality $Ax \leq 1$ ($m = 1$ in the Karush-Kuhn-Tucker conditions), where $A = \begin{pmatrix} 2 & 3 & 4 & \ldots \end{pmatrix}$, and the Lagrangian is given in this case by

$$L = \sum_{n=2}^{\infty} \frac{(x_n - a_n)^2}{n+1} + \mu \left(\sum_{n=2}^{\infty} n x_n - 1 \right).$$

The Karush-Kuhn-Tucker conditions (assuming the same conditions can be applied to infinite instead of finite number of variables x_i – see Remark 5.3.2 following the proof of Theorem 5.3.1) are in this case

$$\begin{aligned}
2\frac{x_n - a_n}{n+1} + n\mu &\geq 0, \qquad n \geq 2 && (5.3.5)\\
\sum_{n=2}^{\infty} n x_n - 1 &\leq 0 && (5.3.6)\\
x_n \left(2\frac{x_n - a_n}{n+1} + n\mu \right) &= 0, \qquad n \geq 2 && (5.3.7)\\
\mu \left(\sum_{n=2}^{\infty} n x_n - 1 \right) &= 0 && (5.3.8)\\
x_n &\geq 0, \qquad n \geq 2 && (5.3.9)\\
\mu &\geq 0 && (5.3.10)
\end{aligned}$$

From (5.3.8) it can be seen that either $\mu = 0$ or $\sum_{n=2}^{\infty} n x_n = 1$, and we distinguish the following cases.

If $\sum_{n=2}^{\infty} n a_n \leq 1$, the infimum in (5.3.3) is readily seen to be 0, attained for $\mu = 0$ and $x_n = a_n$ for all $n \geq 2$.

If $\sum_{n=2}^{\infty} n a_n > 1$, we first note that in this case we must have $\mu \neq 0$. This is so for otherwise from (5.3.7) we obtain $x_n = 0$ or $x_n = a_n$, and since from condition (5.3.5) we have $x_n \geq a_n$, it follows that $x_n = a_n$ for all $n \geq 2$. However, this contradicts (5.3.6), since $\sum_{n=2}^{\infty} nx_n = \sum_{n=2}^{\infty} na_n > 1$.

The above system becomes in this case

$$2\frac{x_n - a_n}{n+1} + n\mu \geq 0, \qquad n \geq 2 \tag{5.3.11}$$

$$x_n \left(2\frac{x_n - a_n}{n+1} + n\mu \right) = 0, \qquad n \geq 2 \tag{5.3.12}$$

$$\sum_{n=2}^{\infty} nx_n = 1 \tag{5.3.13}$$

$$x_n \geq 0, \qquad n \geq 2 \tag{5.3.14}$$

$$\mu > 0 \tag{5.3.15}$$

The second equation shows that either $x_n = 0$ or $x_n = a_n - \frac{1}{2}\mu n\,(n+1)$. Denote by $\mathcal{I}$ the set of indices $n \geq 2$ for which $x_n = a_n - \frac{1}{2}\mu n\,(n+1)$, so $x_n = 0$ for $n \in \mathcal{I}^c = \{2, 3, 4, \ldots\} - \mathcal{I}$.

From (5.3.11) it follows that $\mu \geq 2\frac{a_n}{n(n+1)}$ for $n \in \mathcal{I}^c$ and from (5.3.14) it follows that $\mu \leq 2\frac{a_n}{n(n+1)}$ for $n \in \mathcal{I}$.

From (5.3.13) we obtain

$$1 = \sum_{n=2}^{\infty} nx_n = \sum_{n \in \mathcal{I}} n \left(a_n - \frac{1}{2}\mu n\,(n+1) \right) = \sum_{n \in \mathcal{I}} na_n - \frac{\mu}{2} \sum_{n \in \mathcal{I}} n^2\,(n+1)\,,$$

so we must have

$$\mu = 2\frac{\sum_{n \in \mathcal{I}} na_n - 1}{\sum_{n \in \mathcal{I}} n^2\,(n+1)} > 0. \tag{5.3.16}$$

In particular, the above shows that the set of indices $\mathcal{I}$ must be finite (otherwise, since the series $\sum_{n \in \mathcal{I}} n^2\,(n+1) = \infty$ diverges, we obtain $\mu = 0$, contradicting (5.3.15)).

In order to find the value of x_n, it remains to find the set of indices $\mathcal{I}$ (the last equality gives then the value $x_n = a_n - \frac{1}{2}\mu n\,(n+1)$ for $n \in \mathcal{I}$ and $x_n = 0$ for $n \in \mathcal{I}^c$).

To do this, recall that μ given by (5.3.16) must satisfy

$$\mu \leq \frac{2a_n}{n\,(n+1)}, \qquad n \in \mathcal{I} \tag{5.3.17}$$

and

$$\mu \geq \frac{2a_n}{n\,(n+1)}, \qquad n \in \mathcal{I}^c. \tag{5.3.18}$$

Assuming the series $\sum_{n=2}^{\infty} na_n < \infty$ converges, it follows that the sequence $(na_n)_{n \geq 2}$ converges to 0, so the sequence $\left(\frac{2a_n}{n(n+1)} \right)_{n \geq 2}$ also converges to 0. There exists therefore a non-increasing rearrangement of the sequence $\left(\frac{2a_n}{n(n+1)} \right)_{n \geq 2}$, that is, there exists a permutation $i_2 < i_3 < \ldots$ of $\{2, 3, \ldots\}$ so that so that $\left(\frac{2a_{i_n}}{i_n(i_n+1)} \right)_{n \geq 2}$ is a non-increasing sequence (also convergent to 0).

We will prove the following.

Theorem 5.3.1. *If $\sum_{n=2}^{\infty} na_n > 1$ is a convergent series, there exists an integer $N \geq 2$ such that the minimum of the quadratic problem (5.3.3) – (5.3.4) is attained for the sequence $(x_n)_{n \geq 2}$ given by*

$$x_n = \begin{cases} a_n - \frac{1}{2}\mu_N n\,(n+1)\,, & n \in \mathcal{I} \\ 0, & n \in \mathcal{I}^c \end{cases},$$

where $\mu_N = 2\frac{\sum_{n \in \mathcal{I}} na_n - 1}{\sum_{n \in \mathcal{I}} n^2(n+1)}$, $\mathcal{I} = \{i_2, \ldots, i_N\}$ and $i_2 < i_3 < \ldots$ is a permutation of $\{2, 3, \ldots\}$ such that $\left(\frac{2a_n}{n(n+1)} \right)_{n \geq 2}$ is a non-increasing sequence.

Moreover, N can be taken to be equal to

$$N = \inf\left\{n \geq 2 : \alpha_{i_{n+1}} \leq \mu_n \leq \alpha_{i_n}\right\},$$

where $\alpha_n = \frac{2a_n}{n(n+1)}$ *and* $\mu_n = 2\frac{\sum_{m=2}^{n} i_m a_{i_m} - 1}{\sum_{m=2}^{k} i_m^2 (i_m+1)}$, $n \geq 2$.

Proof. Since $\sum_{n=2}^{\infty} i_n a_{i_n} = \sum_{n=2}^{\infty} n a_n > 1$, there exists an integer $n_0 \geq 2$ such that $\sum_{n=2}^{n_0} i_n a_{i_n} > 1$, and assume that $n_0 \geq 2$ is the smallest index with this property.

Note that if $n_0 = 2$, then $\mu_2 = 2\frac{i_2 a_{i_2} - 1}{i_2^2 \cdot (i_2+1)} \leq \alpha_{i_2}$.

Also note that if $n_0 > 2$, then by the choice of n_0 we have $\mu_{n_0-1} = 2\frac{\sum_{n=2}^{n_0-1} i_n a_{i_n} - 1}{\sum_{n=2}^{n_0-1} i_n^2 (i_n+1)} < 0 \leq \alpha_{n_0}$, so $\mu_{n_0} = \frac{2\left(\sum_{n=2}^{n_0-1} i_n a_{i_n} - 1\right) + 2 i_{n_0} a_{i_{n_0}}}{\sum_{n=2}^{n_0-1} i_n^2 (i_n+1) + i_{n_0}^2 \left(i_{n_0}+1\right)} \leq \frac{2 i_{n_0} a_{i_{n_0}}}{i_{n_0}^2 \left(i_{n_0}+1\right)} = \alpha_{n_0}$ (we are using here the fact that if $\frac{a}{b} \leq \frac{c}{d}$ with $b, d > 0$, then $\frac{a}{b} \leq \frac{a+b}{c+d} \leq \frac{c}{d}$).

So in both cases above we obtained $\mu_{i_{n_0}} \leq \alpha_{n_0}$.

We distinguish now the following cases.

i) $\mu_{n_0} \geq \alpha_{i_{n_0+1}}$

Since the sequence $\left(\alpha_{i_n}\right)_{n\geq 2}$ is non-increasing, we have

$$\mu_{n_0} \leq \alpha_{i_{n_0}} \leq \alpha_{i_n}, \qquad n \in \{2, \ldots, n_0\}$$

and

$$\mu_{n_0} \geq \alpha_{i_{n_0+1}} \geq \alpha_{i_n}, \qquad n \in \{n_0+1, n_0+2, \ldots\},$$

so we can chose $N = n_0$ and $I = \{i_2, \ldots, i_{n_0}\}$, concluding the proof in this case.

ii) $\mu_{n_0} < \alpha_{i_{n_0+1}}$

In this case, using again the above observation we have

$$\mu_{n_0} \leq \mu_{n_0+1} \leq \alpha_{i_{n_0+1}}.$$

Either $\mu_{n_0+1} \geq \alpha_{i_{n_0+2}}$ or $\mu_{n_0+1} < \alpha_{i_{n_0+2}}$.

If $\mu_{n_0+1} \geq \alpha_{i_{n_0+2}}$, proceeding as in part i) above, it follows that we can chose $N = i_{n_0+1}$, so the claim holds in this case.

If $\mu_{n_0+1} < \alpha_{i_{n_0+2}}$, we obtain

$$\mu_{n_0} \leq \mu_{n_0+1} \leq \mu_{n_0+2} \leq \alpha_{i_{n_0+2}}.$$

Proceeding inductively, either at some step we can find an integer $N = n_0 + k$ for which the claim holds, or

$$0 < \mu_{n_0} \leq \mu_{n_0+1} \leq \mu_{n_0+2} \leq \cdots \leq \mu_{n_0+k} \leq \alpha_{i_{n_0+k}}, \qquad k \geq 0. \tag{5.3.19}$$

However, since $\left(\alpha_{i_n}\right)_{n\geq 2}$ is a non-increasing sequence of non-negative real numbers with $\lim_{n\to\infty} \alpha_{i_n} = 0$, the inequalities in (5.3.19) cannot hold, and therefore we can always find an integer $N = n_0 + k$ for which the claim holds, concluding the proof of the theorem.

□

Remark 5.3.2. *In the argument above we have used the Karush-Kuhn-Tucker conditions for an infinite instead of a finite number of variables x_i in order to find the minimum value of the objective function in (5.3.3). The reason for which the Karush-Kuhn-Tucker can be applied to the particular quadratic programming problem (5.3.3) in this infinite-dimensional setting is the following.*

Note that for an arbitrarily fixed sequence of non-negative numbers $(a_n)_{n\geq 2}$ *and any integer* $m \geq 2$ *we have*

$$\inf \sum_{n=2}^{\infty} \frac{(x_n - a_n)^2}{n+1} \geq \inf \sum_{n=2}^{m} \frac{(x_n - a_n)^2}{n+1}, \tag{5.3.20}$$

where both infimum are taken over all non-negative sequences $(x_n)_{n\geq 2}$ *with* $\sum_{n=2}^{\infty} n x_n \leq 1$*. Since* $x_{m+1}, x_{m+2}, \ldots$ *do not appear in the expression in the second infimum, the second infimum is the same when taken over all finite sequences of non-negative numbers* $x_2, \ldots, x_m$ *with* $\sum_{n=2}^{m} n x_n \leq 1$*.*

Solving the Karush-Kuhn-Tucker conditions for this finite-dimensional problem (same calculations as above) and using the notation of Theorem 5.3.1, it follows that for $m \geq i_N$ *the second infimum in (5.3.20) is attained for the sequence* $x_2, \ldots, x_m$ *given by*

$$x_n = \begin{cases} a_n - \frac{1}{2}\mu_N n(n+1), & n \in \mathcal{I} \\ 0, & n \in \mathcal{I}^c \cap \{2, \ldots, m\} \end{cases},$$

so

$$\inf \sum_{n=2}^{\infty} \frac{(x_n - a_n)^2}{n+1} \geq \inf \sum_{n=2}^{m} \frac{(x_n - a_n)^2}{n+1} \geq \pi \sum_{n \in \mathcal{I}_m^c} \frac{a_n^2}{n+1} + \pi \frac{\left(\sum_{n\in\mathcal{I}} n a_n - 1\right)^2}{\sum_{n\in\mathcal{I}} n^2 (n+1)},$$

where $\mathcal{I}_m^c = \{2, \ldots m\} - \mathcal{I}$ *(note that for* $m \geq i_N$ *we have* $I = \{i_2, \ldots, i_N\} \subset \{2, \ldots, m\}$*).*

Since the above inequality holds for any $m \geq i_N$*, passing to the limit with* $m \to \infty$ *we obtain*

$$\begin{aligned} \inf \sum_{n=2}^{\infty} \frac{(x_n - a_n)^2}{n+1} &\geq \lim_{m\to\infty} \pi \sum_{n \in \mathcal{I}_m^c} \frac{a_n^2}{n+1} + \pi \frac{\left(\sum_{n\in\mathcal{I}} n a_n - 1\right)^2}{\sum_{n\in\mathcal{I}} n^2 (n+1)} \\ &= \pi \sum_{n \in \mathcal{I}^c} \frac{a_n^2}{n+1} + \pi \frac{\left(\sum_{n\in\mathcal{I}} n a_n - 1\right)^2}{\sum_{n\in\mathcal{I}} n^2 (n+1)}. \end{aligned}$$

The last expression above is just the value of the on the objective function $\sum_{n=2}^{\infty} \frac{(x_n - a_n)^2}{n+1}$ *for the sequence* $(x_n)_{n\geq 2}$ *defined in Theorem 5.3.1, so the infimum of the infinite-dimensional quadratic problem (5.3.3) is attained for the sequence in the statement of Theorem 5.3.1.*

This justifies the use of the Karush-Kuhn-Tucker conditions in the infinite-dimensional quadratic problem (5.3.3), completing the argument.

5.3.3 A second quadratic programming problem

Consider the problem of finding

$$\inf \sum_{n=2}^{\infty} \frac{(x_n - a_n)^2}{n+1} \tag{5.3.21}$$

where $(a_n)_{n\geq 2}$ is a given sequence of non-negative real numbers, and the infimum is taken over all non-negative sequences $(x_n)_{n\geq 2}$ of real numbers satisfying

$$\sum_{n=2}^{\infty} n^2 x_n \leq 1. \tag{5.3.22}$$

Remark 5.3.3. *Note that without loss of generality we can reduce the above problem to the case when* $a_n > 0$ *for all* $n \geq 2$*. This is so for if we consider the set of indices* $\mathcal{P} = \{n \geq 2 : a_n > 0\}$*, then*

$$\inf \sum_{n=2}^{\infty} \frac{(x_n - a_n)^2}{n+1} = 0 \wedge \inf \sum_{n\in\mathcal{P}} \frac{(x_n - a_n)^2}{n+1},$$

where the second infimum is taken over all sequences $(x_n)_{n\in\mathcal{P}}$ *with* $\sum_{n\in\mathcal{P}} n^2 x_n \leq 1$*, and we can consider* $x_n = a_n = 0$ *for* $n \in \{2, 3, \ldots\} - \mathcal{P}$ *(recall the usual convention* $\inf \varnothing = +\infty$*).*

The above problem is a particular case of a semi-infinite quadratic problem, with objective function $f(x) = \sum_{n=2}^{\infty} \frac{(x_n - a_n)^2}{n+1}$ and only one constrained inequality $Ax \leq 1$, where $x = \begin{pmatrix} x_2 & x_3 & x_4 & \dots \end{pmatrix}'$ and $A = \begin{pmatrix} 2^2 & 3^2 & 4^2 & \dots \end{pmatrix}$. The corresponding Lagrangian is in this case given by

$$L = \sum_{n=2}^{\infty} \frac{(x_n - a_n)^2}{n+1} + \mu \left(\sum_{n=2}^{\infty} n^2 x_n - 1 \right).$$

The Karush-Kuhn-Tucker conditions (see [47] or [71]) are necessary conditions for a global minimum of a quadratic programming problem, and it is known that if in addition the objective function is strictly convex, they are also sufficient for a global minimum. Assuming for the moment that the same conditions can be applied to infinite, instead of finite number of variables x_n (see Remark 5.3.6 below), the solution of the quadratic problem (5.5.1) – (5.5.2) above is given by the Karush-Kuhn-Tucker conditions specialized for this case, that is

$$2\frac{x_n - a_n}{n+1} + n^2 \mu \geq 0, \qquad n \geq 2 \tag{5.3.23}$$

$$\sum_{n=2}^{\infty} n^2 x_n - 1 \leq 0 \tag{5.3.24}$$

$$x_n \left(2\frac{x_n - a_n}{n+1} + n^2 \mu \right) = 0, \qquad n \geq 2 \tag{5.3.25}$$

$$\mu \left(\sum_{n=2}^{\infty} n^2 x_n - 1 \right) = 0 \tag{5.3.26}$$

$$x_n \geq 0, \qquad n \geq 2 \tag{5.3.27}$$

$$\mu \geq 0 \tag{5.3.28}$$

From (5.3.26) it can be seen that either $\mu = 0$ or $\sum_{n=2}^{\infty} n^2 x_n = 1$, and we distinguish the following cases.

If $\sum_{n=2}^{\infty} n^2 a_n \leq 1$, the infimum in (5.5.1) is readily seen to be 0, attained for $\mu = 0$ and $x_n = a_n$ for all $n \geq 2$.

If $\sum_{n=2}^{\infty} n^2 a_n > 1$, we first note that in this case we must have $\mu \neq 0$. This is so for otherwise from (5.3.25) we obtain $x_n = 0$ or $x_n = a_n$, and since from condition (5.3.23) we have $x_n \geq a_n$, it follows that $x_n = a_n$ for all $n \geq 2$. However, this contradicts (5.3.24), since $\sum_{n=2}^{\infty} n^2 x_n = \sum_{n=2}^{\infty} n^2 a_n > 1$.

The above system becomes in this case

$$2\frac{x_n - a_n}{n+1} + n^2 \mu \geq 0, \qquad n \geq 2 \tag{5.3.29}$$

$$x_n \left(2\frac{x_n - a_n}{n+1} + n^2 \mu \right) = 0, \qquad n \geq 2 \tag{5.3.30}$$

$$\sum_{n=2}^{\infty} n^2 x_n = 1 \tag{5.3.31}$$

$$x_n \geq 0, \qquad n \geq 2 \tag{5.3.32}$$

$$\mu > 0 \tag{5.3.33}$$

The second equation shows that either $x_n = 0$ or $x_n = a_n - \frac{1}{2}\mu n^2 (n+1)$. Denote by $\mathcal{I}$ the set of indices $n \geq 2$ for which $x_n = a_n - \frac{1}{2}\mu n^2 (n+1)$, so $x_n = 0$ for $n \in \mathcal{I}^c = \{2, 3, 4, \dots\} - \mathcal{I}$.

From (5.3.29) it follows that $\mu \geq 2\frac{a_n}{n^2(n+1)}$ for $n \in \mathcal{I}^c$ and from (5.3.32) it follows that $\mu \leq 2\frac{a_n}{n^2(n+1)}$ for $n \in \mathcal{I}$.

From (5.3.31) we obtain

$$1 = \sum_{n=2}^{\infty} n^2 x_n = \sum_{n \in \mathcal{I}} n^2 \left(a_n - \frac{1}{2}\mu n^2 (n+1) \right) = \sum_{n \in \mathcal{I}} n^2 a_n - \frac{\mu}{2} \sum_{n \in \mathcal{I}} n^4 (n+1),$$

so we must have

$$\mu = 2\frac{\sum_{n\in\mathcal{I}} n^2 a_n - 1}{\sum_{n\in\mathcal{I}} n^4 (n+1)} > 0. \tag{5.3.34}$$

In particular, the above shows that the set of indices $\mathcal{I}$ must be finite (otherwise, since the series $\sum_{n\in\mathcal{I}} n^4 (n+1) = \infty$ diverges, we obtain $\mu = 0$, contradicting (5.3.33)).

In order to find the value of x_n, it remains to find the set of indices $\mathcal{I}$ (the last equality gives then the value $x_n = a_n - \frac{1}{2}\mu n^2 (n+1)$ for $n \in \mathcal{I}$ and $x_n = 0$ for $n \in \mathcal{I}^c$).

To do this, recall that μ given by (5.3.34) must satisfy

$$\mu \leq \frac{2a_n}{n^2 (n+1)}, \qquad n \in \mathcal{I} \tag{5.3.35}$$

and

$$\mu \geq \frac{2a_n}{n^2 (n+1)}, \qquad n \in \mathcal{I}^c. \tag{5.3.36}$$

Remark 5.3.4. *Note that if $(\alpha_n)_{n\geq 2}$ is a sequence of positive numbers with $\lim_{n\to\infty} \alpha_n = 0$, then each of the intervals $[1, +\infty)$ and $[1/(m+1), 1/m)$, $m \geq 1$, can contain only a finite number of the terms of the sequence. We can therefore find a permutation $(i_n)_{n\geq 2}$ of the indices in $\{2, 3, \ldots\}$ such that $(\alpha_{i_n})_{n\geq 2}$ is a non-increasing sequence and $\lim_{n\to\infty} \alpha_{i_n} = 0$.*

With this preparation we can now prove the following.

Theorem 5.3.5. *If $(a_n)_{n\geq 2}$ is a sequence of non-negative real numbers with*

$$\sum_{n=2}^{\infty} n^2 a_n > 1 \tag{5.3.37}$$

and

$$\lim_{n\to\infty} \frac{a_n}{n^3} = 0, \tag{5.3.38}$$

there exists an integer $N \geq 2$ such that the minimum of the quadratic problem (5.5.1) – (5.5.2) is

$$\sum_{n\in\mathcal{I}^c} \frac{a_n^2}{n+1} + \frac{\left(\sum_{n\in\mathcal{I}} (n^2 a_n) - 1\right)^2}{\sum_{n\in\mathcal{I}} (n^4 (n+1))},$$

attained for the sequence $(x_n)_{n\geq 2}$ given by

$$x_n = \begin{cases} a_n - \frac{1}{2}\mu_N n^2 (n+1), & n \in \mathcal{I} \\ 0, & n \in \mathcal{I}^c \end{cases},$$

where $\mu_N = 2\frac{\sum_{n\in\mathcal{I}}(n^2 a_n)-1}{\sum_{n\in\mathcal{I}}(n^4(n+1))}$, $\mathcal{I} = \{i_2, \ldots, i_N\}$, and $(i_n)_{n=2\ldots,|\mathcal{P}|+1}$ is a permutation of the indices in $\mathcal{P} = \{n \geq 2 : a_n > 0\}$ such that $\alpha_n = \frac{2a_{i_n}}{i_n^2 (i_n+1)}$, $n = 2, \ldots, |\mathcal{P}| + 1$, is a non-increasing sequence.

Moreover, N can be taken to be equal to

$$N = \min \{n \geq 2 : \alpha_{n+1} \leq \mu_n \leq \alpha_n\},$$

where $\mu_n = 2\frac{\sum_{m=2}^{n}(i_m^2 a_{i_m})-1}{\sum_{m=2}^{n}(i_m^4 (i_m+1))}$, $n = 2, \ldots, |\mathcal{P}| + 1$.

Proof. The discussion preceding the statement of the theorem shows that in order to prove the claim, it suffices to show that we can chose the set of indices $\mathcal{I}$ such that the relations (5.3.35) and (5.3.36) hold true (the relation (5.3.34) gives then the value of μ, and the minimum of the quadratic problem in the statement of the theorem is attained for the sequence $(x_n)_{n\geq 2}$ with $x_n = a_n - \frac{1}{2}\mu n^2 (n+1)$ for $n \in \mathcal{I}$ and $x_n = 0$ otherwise).

Consider first the case when the terms of the sequence $(a_n)_{n\geq 2}$ are positive real numbers, so $\mathcal{P}=\{2,3,\ldots\}$. Remark 5.3.4 above and the hypothesis (5.3.38) show that we can chose a permutation $(i_n)_{n\geq 2}$ of the indices in $\mathcal{P}$ such that $(\alpha_n)_{n\geq 2}$ is a non-increasing sequence.

Since $\sum_{n=2}^{\infty} i_n^2 a_{i_n} = \sum_{n=2}^{\infty} n^2 a_n > 1$, there exists an integer $n_0 \geq 2$ such that $\sum_{n=2}^{n_0} i_n^2 a_{i_n} > 1$, and assume that $n_0 \geq 2$ is the smallest index with this property.

First note that we must have $0 < \mu_{n_0} \leq \alpha_{n_0}$. This is so for if $n_0 = 2$, then

$$\mu_2 = 2\frac{i_2^2 a_{i_2} - 1}{i_2^4 \cdot (i_2+1)} \leq \frac{2a_{i_2}}{i_2^2 \cdot (i_2+1)} = \alpha_2,$$

so the claim holds in this case. If $n_0 > 2$, by the choice of n_0 we have

$$\mu_{n_0-1} = 2\frac{\sum_{n=2}^{n_0-1}\left(i_n^2 a_{i_n}\right) - 1}{\sum_{n=2}^{n_0-1}\left(i_n^4\left(i_n+1\right)\right)} \leq 0 \leq \frac{2 i_{n_0}^2 a_{i_{n_0}}}{i_{n_0}^4\left(i_{n_0}+1\right)} = \alpha_{n_0},$$

and using the observation that $\frac{a}{b} \leq \frac{c}{d}$ with $b,d>0$ implies $\frac{a+c}{b+d} \leq \frac{c}{d}$, we obtain

$$\mu_{n_0} = 2\frac{\sum_{n=2}^{n_0-1}\left(i_n^2 a_{i_n}\right) - 1 + i_{n_0}^2 a_{i_{n_0}}}{\sum_{n=2}^{n_0-1}\left(i_n^4\left(i_n+1\right)\right) + i_{n_0}^4\left(i_{n_0}+1\right)} \leq \frac{2 i_{n_0}^2 a_{i_{n_0}}}{i_{n_0}^4\left(i_{n_0}+1\right)} = \alpha_{n_0},$$

concluding the proof of the claim.

We distinguish now the following cases.

Case 1: $\mu_{n_0} \geq \alpha_{n_0+1}$.

Since the sequence $(\alpha_{i_n})_{n\geq 2}$ is non-increasing, we have

$$\mu_{n_0} \leq \alpha_{_{n_0}} \leq \alpha_{_n}, \qquad n \in \{2,\ldots,n_0\}$$

and

$$\mu_{n_0} \geq \alpha_{_{n_0+1}} \geq \alpha_n, \qquad n \in \{n_0+1, n_0+2, \ldots\},$$

so we can chose $N = n_0$ and $\mathcal{I} = \{i_2,\ldots,i_{n_0}\}$, concluding the proof in this case.

Case 2: $\mu_{n_0} < \alpha_{n_0+1}$.

In this case, using again the above observation we have

$$\mu_{n_0} \leq \mu_{n_0+1} \leq \alpha_{n_0+1},$$

and either $\mu_{n_0+1} \geq \alpha_{n_0+2}$ or $\mu_{n_0+1} < \alpha_{n_0+2}$.

If $\mu_{n_0+1} \geq \alpha_{n_0+2}$, proceeding as in Case 1 above, we can choose $N = n_0+1$, so the claim holds in this case.

If $\mu_{n_0+1} < \alpha_{n_0+2}$, we obtain

$$\mu_{n_0} \leq \mu_{n_0+1} \leq \mu_{n_0+2} \leq \alpha_{_{n_0+2}}.$$

Proceeding inductively, either at some step we can find an integer $N = n_0 + k$ for which the claim holds, or

$$0 < \mu_{n_0} \leq \mu_{n_0+1} \leq \mu_{n_0+2} \leq \cdots \leq \mu_{n_0+k} \leq \alpha_{n_0+k}, \qquad k \geq 0. \tag{5.3.39}$$

However, since $(\alpha_n)_{n\geq 2}$ is a non-increasing sequence of positive real numbers with $\lim_{n\to\infty}\alpha_n = 0$, the inequalities in (5.3.39) cannot hold for every $k \geq 0$. Therefore we can always find an integer $N = n_0 + k$ for which the claim holds, concluding the proof of the theorem in the case when $(a_n)_{n\geq 2}$ is a sequence of positive real numbers.

Consider now the general case, when $(a_n)_{n\geq 2}$ is a sequence of non-negative real numbers.

If the set $\mathcal{P}$ is infinite, applying the proof above to the sequence $(a_n)_{n\in\mathcal{P}}$ of positive real numbers (note that $\sum_{n\in\mathcal{P}} n^2 a_n = \sum_{n\geq 2} n^2 a_n > 1$ and $\lim_{\mathcal{P}\ni n\to\infty}\frac{a_n}{n^3} = \lim_{n\to\infty}\frac{a_n}{n^3} = 0$) and using Remark 5.3.3, it follows that the claim holds in this case (for $n \in \{2,3,\ldots\} - \mathcal{P} \subset \mathcal{I}^c$ we have $x_n = a_n = 0$).

If the set $\mathcal{P}$ is finite, by the hypothesis (5.3.37) it follows that $\mathcal{P}$ cannot be empty, so $|\mathcal{P}| = p$ for some $p \geq 1$. If $(i_n)_{n=2}^{p+1}$ is a permutation of the indices in $\mathcal{P}$ such that $\alpha_n = \frac{2a_{i_n}}{i_n^2(i_n+1)}$, $n = 2, \ldots, p+1$, is a non-increasing sequence, proceeding as in the proof above, either we can find an index N for which the claim holds, or

$$0 < \mu_{n_0} \leq \mu_{n_0+1} \leq \cdots \leq \mu_{p+1} \leq \alpha_{p+1}.$$

In the later case we can chose $N = p + 1$ and $\mathcal{I} = \{i_2, \ldots, i_{p+1}\}$, and obtain

$$\mu_N \leq \frac{2a_n}{n^2(n+1)}, \qquad n \in \mathcal{I}, \tag{5.3.40}$$

and

$$\mu_N \geq \frac{2a_n}{n^2(n+1)} = 0, \qquad n \in \mathcal{I}^c, \tag{5.3.41}$$

so the claim also holds in this case, concluding the proof of the theorem. □

Remark 5.3.6. *In the argument above we have used the Karush-Kuhn-Tucker conditions for an infinite instead of a finite number of variables x_n in order to find the minimum value of the objective function in (5.5.1). The reason for which the Karush-Kuhn-Tucker can be applied to the particular quadratic programming problem (5.5.1) – (5.5.2) is the following.*

Note that for an arbitrarily fixed sequence of non-negative numbers $(a_n)_{n\geq 2}$ and any integer $m \geq 2$ we have

$$\inf \sum_{n=2}^{\infty} \frac{(x_n - a_n)^2}{n+1} \geq \inf \sum_{n=2}^{m} \frac{(x_n - a_n)^2}{n+1}, \tag{5.3.42}$$

where both infima are taken over all non-negative sequences $(x_n)_{n\geq 2}$ of real numbers with $\sum_{n=2}^{\infty} n^2 x_n \leq 1$. Since $x_{m+1}, x_{m+2}, \ldots$ do not appear in the expression in the second infimum, the second infimum is the same when taken over all finite sequences of non-negative numbers $x_2, \ldots, x_m$ with $\sum_{n=2}^{m} n^2 x_n \leq 1$.

Solving the Karush-Kuhn-Tucker conditions for this finite-dimensional problem (the calculations are the same as in the proof above) and using the notation of Theorem 5.3.5, it follows that for $m \geq \max\{i_2, \ldots, i_N\}$ the second infimum in (5.3.42) is attained for the sequence $x_2, \ldots, x_m$ given by

$$x_n = \begin{cases} a_n - \frac{1}{2}\mu_N n(n+1), & n \in \mathcal{I} \\ 0, & n \in \mathcal{I}^c \cap \{2, \ldots, m\} \end{cases},$$

so

$$\inf \sum_{n=2}^{\infty} \frac{(x_n - a_n)^2}{n+1} \geq \inf \sum_{n=2}^{m} \frac{(x_n - a_n)^2}{n+1} = \sum_{n\in\mathcal{I}_m^c} \frac{a_n^2}{n+1} + \frac{\left(\sum_{n\in\mathcal{I}}(n^2 a_n) - 1\right)^2}{\sum_{n\in\mathcal{I}}(n^4(n+1))},$$

where $\mathcal{I}_m^c = \{2, \ldots m\} - \mathcal{I}$ (note that for $m \geq \max\{i_2, \ldots, i_N\}$ we have $\mathcal{I} = \{i_2, \ldots, i_N\} \subset \{2, \ldots, m\}$).

Since the above inequality holds for any $m \geq \max\{i_2, \ldots, i_N\}$, passing to the limit with $m \to \infty$ we obtain

$$\begin{aligned} \inf \sum_{n=2}^{\infty} \frac{(x_n - a_n)^2}{n+1} &\geq \lim_{m\to\infty} \sum_{n\in\mathcal{I}_m^c} \frac{a_n^2}{n+1} + \frac{\left(\sum_{n\in\mathcal{I}}(n^2 a_n) - 1\right)^2}{\sum_{n\in\mathcal{I}}(n^4(n+1))} \\ &= \sum_{n\in\mathcal{I}^c} \frac{a_n^2}{n+1} + \frac{\left(\sum_{n\in\mathcal{I}}(n^2 a_n) - 1\right)^2}{\sum_{n\in\mathcal{I}}(n^4(n+1))}. \end{aligned}$$

The last expression above is just the value of the objective function $\sum_{n=2}^{\infty} \frac{(x_n - a_n)^2}{n+1}$ for the sequence $(x_n)_{n\geq 2}$ defined in Theorem 5.3.5, so the infimum of the quadratic problem (5.5.1) – (5.5.2) is attained for the sequence in the statement of Theorem 5.3.5.

This justifies the use of the Karush-Kuhn-Tucker conditions for the quadratic problem (5.5.1) – (5.5.2) with an infinite number of variables x_n, completing the argument.

5.4 Approximation of analytic functions by starlike functions

Using the results from the previous section, we will determine $\operatorname{dist}(f, \mathcal{S}^*)$ for a given function $f \in \mathcal{A}$, that is we will find

$$\operatorname{dist}(f, \mathcal{S}^*) = \inf_{g \in \mathcal{S}^*} \left(\int_U |f(x+iy) - g(x+iy)|^2 \, dxdy \right)^{1/2},$$

and we will determine the extremal function $g \in \mathcal{S}^*$ for which the minimum is attained.

In view of Lemma 5.2.2, if $f \in \mathcal{A}$ has the series expansion $f(z) = z + \sum_{n=2}^{\infty} a_n z^n$, $z \in U$, this amounts to finding

$$\operatorname{dist}(f, \mathcal{S}^*) = \left(\pi \inf \sum_{n=2}^{\infty} \frac{|a_n - b_n|^2}{n+1} \right)^{1/2},$$

where the infimum is taken over all sequences $(b_n)_{n \geq 1}$ in $\mathbb{C}$ satisfying

$$\sum_{n=2}^{\infty} n |b_n| \leq 1.$$

The main result is the following.

Theorem 5.4.1. *Consider $f \in \mathcal{A}$ with series expansion given by*

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n, \qquad z \in U, \tag{5.4.1}$$

and assume that the series $\sum_{n=2}^{\infty} n |a_n|$ converges.

If $\sum_{n=2}^{\infty} n |a_n| \leq 1$ then $\operatorname{dist}(f, \mathcal{S}^*) = 0$ (attained for $g = f \in \mathcal{S}^* \subset S$), and if $\sum_{n=2}^{\infty} n |a_n| > 1$ we have

$$\operatorname{dist}(f, \mathcal{S}^*) = \left(\pi \sum_{n \in \mathcal{I}^c} \frac{|a_n|^2}{n+1} + \pi \frac{\left(\sum_{n \in \mathcal{I}} n |a_n| - 1 \right)^2}{\sum_{n \in \mathcal{I}} n^2 (n+1)} \right)^{1/2}, \tag{5.4.2}$$

where N and $\mathcal{I} = \{i_2, \ldots, i_N\}$ are given by Theorem 5.3.1 with $|a_n|$ instead of a_n, and the minimum value of $\operatorname{dist}(f, \mathcal{S}^*) = \inf_{g \in \mathcal{S}^*} \left(\int_U |f(x+iy) - g(x+iy)|^2 \, dxdy \right)^{1/2}$ is attained for the function $g(z) = z + \sum_{n=2}^{\infty} b_n z^n \in \mathcal{S}^* \subset \mathcal{S}$, where

$$b_n = \begin{cases} \left(|a_n| - \frac{\sum_{m \in \mathcal{I}} m |a_m| - 1}{\sum_{m \in \mathcal{I}} m^2 (m+1)} n(n+1) \right) e^{i \arg a_n}, & n \in \mathcal{I} \\ 0, & n \in \mathcal{I}^c \end{cases}. \tag{5.4.3}$$

Proof. The claim is obvious if $\sum_{n=2}^{\infty} n |a_n| \leq 1$, so assume that $\sum_{n=2}^{\infty} n |a_n| > 1$.

Using Lemma 5.2.2 and the triangle inequality we obtain

$$\begin{aligned} \operatorname{dist}(f, \mathcal{S}^*) &= \inf_{g \in \mathcal{S}^*} \left(\int_U |f(x+iy) - g(x+iy)|^2 \, dxdy \right)^{1/2} \\ &= \left(\pi \inf \sum_{n=2}^{\infty} \frac{|a_n - b_n|^2}{n+1} \right)^{1/2} \\ &\geq \left(\pi \inf \sum_{n=2}^{\infty} \frac{(|a_n| - |b_n|)^2}{n+1} \right)^{1/2} \\ &= \left(\pi \inf \sum_{n=2}^{\infty} \frac{(|a_n| - x_n)^2}{n+1} \right)^{1/2} \end{aligned}$$

where the second and the third infimum are taken over all sequences $(b_n)_{n\geq 2}$ of complex numbers satisfying $\sum_{n=2}^{\infty} n\,|b_n| \leq 1$, and the last infimum is taken over all non-negative sequences $(x_n)_{n\geq 2}$ of real numbers satisfying

$$\sum_{n=2}^{\infty} n x_n \leq 1.$$

Using Theorem 5.3.1 with $|a_n|$ instead of a_n, we obtain that the last infimum above is attained for the sequence $(x_n)_{n\geq 2}$ given by

$$x_n = \begin{cases} a_n - \frac{\sum_{m\in\mathcal{I}} m|a_m|-1}{\sum_{m\in\mathcal{I}} m^2(m+1)} n\,(n+1)\,, & n\in\mathcal{I} \\ 0, & n\in\mathcal{I}^c \end{cases}.$$

It follows that $\operatorname{dist}(f,\mathcal{S}^*) = \left(\pi \inf \sum_{n=2}^{\infty} \frac{|a_n-b_n|^2}{n+1}\right)^{1/2}$ is attained for the sequence $(b_n)_{n\geq 2}$ of complex numbers with $|b_n| = x_n$ and $\arg b_n = \arg a_n$, that is for $b_n = x_n e^{i \arg a_n}$.

Denoting $g(z) = z + \sum_{n=2}^{\infty} b_n z^n \in \mathcal{S}^*$ we have

$$\begin{aligned} \operatorname{dist}(f,\mathcal{S}^*) &= \left(\int_U |f(x+iy) - g(x+iy)|^2\,dxdy\right)^{1/2} \\ &= \left(\pi \sum_{n=2}^{\infty} \frac{|a_n-b_n|^2}{n+1}\right)^{1/2} \\ &= \left(\pi \sum_{n=2}^{\infty} \frac{(|a_n|-|b_n|)^2}{n+1}\right)^{1/2} \\ &= \left(\pi \sum_{n\in\mathcal{I}^c} \frac{|a_n|^2}{n+1} + \pi \frac{\left(\sum_{n\in\mathcal{I}} n\,|a_n| - 1\right)^2}{\sum_{n\in\mathcal{I}} n^2\,(n+1)}\right)^{1/2}, \end{aligned}$$

concluding the proof. □

As an example, consider the following.

Example 5.4.2. *Consider the function $f_a : U \to \mathbb{C}$ defined by $f_a(z) = z + az^2$, where $a\in\mathbb{C}$ is a constant.*

If $|a| \leq \frac{1}{2}$, then $f_a\in\mathcal{S}^*$ and $\operatorname{dist}(f_a,\mathcal{S}^*) = 0$.

If $2\,|a| > 1$, from Theorem 5.4.1 we obtain that $i_n = n$ for $n\geq 2$, $N = 2$ and $\mathcal{I} = \{i_2\} = \{2\}$, so $\operatorname{dist}(f_a,\mathcal{S}^*) = \sqrt{\frac{\pi}{3}}\left(|a| - \frac{1}{2}\right)$ is attained for the function $g_a(z) = z + \frac{a}{2|a|} z \in \mathcal{S}^* \subset \mathcal{S}$.

Figure 5.4 shows a comparison between the image domains $f_a(U)$ for $a = 0.5$, 0.75, and 1. Note that in all cases the minimum of $\operatorname{dist}(f_a,\mathcal{S}^*)$ is attained for $f_{0.5}$, that $f_{0.5}$ is starlike, and that $f_{0.75}$ and f_1 are not univalent.

As another example, consider the following.

Example 5.4.3. *Consider the function $f : U \to \mathbb{C}$ defined by $f(z) = z + az^2 + bz^3$, where $a, b\in\mathbb{C}$ are constants.*

If $2\,|a| + 3\,|b| \leq 1$ then $f\in\mathcal{S}^*$ and $\operatorname{dist}(f,\mathcal{S}^*) = 0$.

If $2\,|a| + 3\,|b| > 1$, we distinguish the following cases.

a) *If $2\,|a| \geq |b|$, then $i_n = n$ for all $n\geq 2$.*

We distinguish the following subcases.

i) *If $2\,|a| - 1 \geq |b|$, from Theorem 5.4.1 it follows that $N = 2$ and $\mathcal{I} = \{i_2\} = \{2\}$, so $\operatorname{dist}(f,\mathcal{S}^*) = \left(\pi \frac{|b|^2}{4} + \pi \frac{(2|a|-1)^2}{12}\right)^{1/2}$, attained for the function $g(z) = z + \frac{1}{2}\frac{a}{|a|} z^2$.*

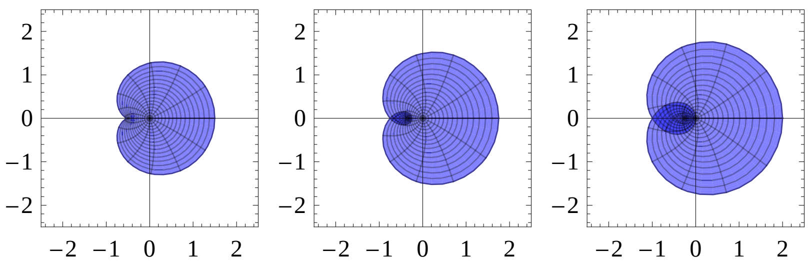

Figure 5.1: The image of the unit disk under f_a, for $a = 0.5$ (left), $a = 0.75$ (center) and $a = 1$ (right).

ii) *If* $2|a|-1 < |b|$, *then* $N = 3$ *and* $\mathcal{I} = \{i_2, i_3\} = \{2, 3\}$, *so* $\operatorname{dist}(f, \mathcal{S}^*) = \frac{\sqrt{\pi}}{4\sqrt{3}}(2|a| + 3|b| - 1)$, *attained for the function* $g(z) = z + \frac{6|a|-3|b|+1}{8} e^{i \arg a} z^2 + \frac{|b|-2|a|+1}{4} e^{i \arg b} z^3$.

b) *If* $2|a| < |b|$, *then* $i_2 = 3$, $i_3 = 2$ *and* $i_n = n$ *for* $n \geq 4$.

We distinguish the following subcases.

i) *If* $3|b| - 1 \geq 6|a|$, *from Theorem 5.4.1 it follows that* $N = 2$ *and* $\mathcal{I} = \{i_2\} = \{3\}$, *so* $\operatorname{dist}(f, \mathcal{S}^*) = \left(\pi\frac{|a|^2}{3} + \pi\frac{(3|b|-1)^2}{36}\right)^{1/2}$, *attained for the function* $g(z) = z + \frac{1}{3}\frac{b}{|b|}z^3$.

ii) *If* $3|b|-1 < 6|a|$, *then* $N = 3$ *and* $\mathcal{I} = \{i_2, i_3\} = \{2, 3\}$, *so* $\operatorname{dist}(f, \mathcal{S}^*) = \frac{\sqrt{\pi}}{4\sqrt{3}}(2|a| + 3|b| - 1)$, *attained for the function* $g(z) = z + \left(\frac{6|a|-3|b|+1}{8}\right) e^{i \arg a} z^2 + \left(\frac{|b|-2|a|+1}{4}\right) e^{i \arg b} z^3$.

The above cases can be summarized as follows. If $f(z) = z + az^2 + bz^3$ *with* $2|a| + 3|b| > 1$, *then*

$$\operatorname{dist}(f, \mathcal{S}^*) = \begin{cases} \frac{\sqrt{\pi}}{2\sqrt{3}}\left((2|a|-1)^2 + 3|b|^2\right)^{1/2}, & \text{if } |b| \leq 2|a| - 1 \\ \frac{\sqrt{\pi}}{4\sqrt{3}}(2|a| + 3|b| - 1), & \text{if } 2|a| - 1 < |b| < 2|a| + \frac{1}{3} \\ \frac{\sqrt{\pi}}{6}\left(12|a|^2 + (3|b|-1)^2\right)^{1/2}, & \text{if } |b| > 2|a| + \frac{1}{3} \end{cases},$$

attained for $g(z) = z + \frac{1}{2}\frac{a}{|a|}z^2$, $g(z) = z + \left(\frac{6|a|-3|b|+1}{8}\right) e^{i \arg a} z^2 + \left(\frac{|b|-2|a|+1}{4}\right) e^{i \arg b} z^3$, *respectively for* $g(z) = z + \frac{1}{3}\frac{b}{|b|}z^3$.

We conclude with the remark that the hypotheses on the convergence of the series $\sum_{n=2}^{\infty} n|a_n|$ in Theorem 5.4.1, respectively the convergence of the series $\sum_{n=2}^{\infty} na_n$ in Theorem 5.3.1, are not essential for the validity of these theorems. Reviewing the proofs of Theorem 5.4.1 and Theorem 5.3.1 it can be seen that these hypotheses were only used in order to show that the sequence $\left(\frac{2|a_n|}{n(n+1)}\right)_{n\geq 2}$ (respectively $\left(\frac{2a_n}{n(n+1)}\right)_{n\geq 2}$) admit a non-increasing rearrangement convergent to 0. So we can substitute these hypotheses for example by the weaker hypotheses $\lim_{n\to\infty}\frac{|a_n|}{n^2} = 0$ (respectively $\lim_{n\to\infty}\frac{a_n}{n^2} = 0$), or by requiring that the sequence $(|a_n|)_{n\geq 2}$ (respectively $(a_n)_{n\geq 2}$) is bounded.

5.5 Approximation of analytic functions by convex functions

In this section we give a method for constructing the best approximation of an analytic function in the subclass $\mathcal{K}^* \subset \mathcal{K}$ of convex functions, in the sense of the L^2 norm.

Lemma 5.2.2, and the definitions of the class $\mathcal{K}^*$ and of $\operatorname{dist}(f, \mathcal{K}^*)$ lead us to consider the problem of finding

$$\inf \sum_{n=2}^{\infty} \frac{(x_n - a_n)^2}{n+1} \tag{5.5.1}$$

where $(a_n)_{n\geq 2}$ is a given sequence of non-negative real numbers, and the infimum is taken over all non-negative sequences $(x_n)_{n\geq 2}$ of real numbers satisfying

$$\sum_{n=2}^{\infty} n^2 x_n \leq 1. \tag{5.5.2}$$

As an application of Theorem 5.3.5, for a given normed analytic function $f \in \mathcal{A}$ we will find

$$\operatorname{dist}(f, \mathcal{K}^*) = \inf_{g\in\mathcal{K}^*} \left(\int_U |f(x+iy) - g(x+iy)|^2 \, dxdy \right)^{1/2},$$

and we will determine the extremal function $g \in \mathcal{K}^*$ for which the minimum is attained.

The main result is the following.

Theorem 5.5.1. *Consider $f \in \mathcal{A}$ with series expansion given by*

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n, \qquad z \in U, \tag{5.5.3}$$

and assume that $\lim_{n\to\infty} \frac{|a_n|}{n^3} = 0$.

If $\sum_{n=2}^{\infty} \left(n^2 |a_n|\right) \leq 1$ then $\operatorname{dist}(f, \mathcal{K}^*) = 0$ (attained for $g = f \in \mathcal{K}^* \subset \mathcal{K}$), and if $\sum_{n=2}^{\infty} \left(n^2 |a_n|\right) > 1$ we have

$$\operatorname{dist}(f, \mathcal{K}^*) = \left(\pi \sum_{n\in\mathcal{I}^c} \frac{|a_n|^2}{n+1} + \pi \frac{\left(\sum_{n\in\mathcal{I}} \left(n^2 |a_n|\right) - 1\right)^2}{\sum_{n\in\mathcal{I}} \left(n^4 (n+1)\right)} \right)^{1/2}, \tag{5.5.4}$$

where N and $\mathcal{I} = \{i_2, \ldots, i_N\}$ are given by Theorem 5.3.5 with $|a_n|$ instead of a_n.

Moreover, the minimum value of $\operatorname{dist}(f, \mathcal{K}^*)$ is attained for the function $g(z) = z + \sum_{n=2}^{\infty} b_n z^n \in \mathcal{K}^* \subset \mathcal{K}$, where

$$b_n = \begin{cases} \left(|a_n| - \frac{\sum_{m\in\mathcal{I}}\left(m^2|a_m|\right)-1}{\sum_{m\in\mathcal{I}}\left(m^4(m+1)\right)} n^2 (n+1) \right) e^{i \arg a_n}, & n \in \mathcal{I} \\ 0, & n \in \mathcal{I}^c \end{cases}. \tag{5.5.5}$$

Proof. The claim is obvious if $\sum_{n=2}^{\infty} n^2 |a_n| \leq 1$, so assume that $\sum_{n=2}^{\infty} n^2 |a_n| > 1$.

Using Lemma 5.2.2 and the triangle inequality we obtain

$$\begin{aligned} \operatorname{dist}(f, \mathcal{K}^*) &= \inf_{g\in\mathcal{K}^*} \left(\int_U |f(x+iy) - g(x+iy)|^2 \, dxdy \right)^{1/2} \\ &= \left(\pi \inf \sum_{n=2}^{\infty} \frac{|a_n - b_n|^2}{n+1} \right)^{1/2} \\ &\geq \left(\pi \inf \sum_{n=2}^{\infty} \frac{(|a_n| - |b_n|)^2}{n+1} \right)^{1/2} \\ &= \left(\pi \inf \sum_{n=2}^{\infty} \frac{(|a_n| - x_n)^2}{n+1} \right)^{1/2}, \end{aligned}$$

where the second and the third infimum are taken over all sequences $(b_n)_{n\geq 2}$ of complex numbers satisfying $\sum_{n=2}^{\infty} n^2 |b_n| \leq 1$, and the last infimum is taken over all non-negative sequences $(x_n)_{n\geq 1}$ of real numbers satisfying $\sum_{n=2}^{\infty} n^2 x_n \leq 1$.

Applying Theorem 5.3.5 with $|a_n|$ instead of a_n, we obtain that the last infimum above is attained for the sequence $(x_n)_{n\geq 2}$ given by

$$x_n = \begin{cases} |a_n| - \frac{\sum_{m\in\mathcal{I}}(m^2|a_m|)-1}{\sum_{m\in\mathcal{I}}(m^4(m+1))} n^2(n+1), & n \in \mathcal{I} \\ 0, & n \in \mathcal{I}^c \end{cases}.$$

Observing that the triangle inequality $|a_n - b_n| \leq (|a_n| - |b_n|)^2$ becomes an equality if $\arg a_n = \arg b_n$, it follows that $\operatorname{dist}(f, \mathcal{K}^*) = \left(\pi \inf \sum_{n=1}^{\infty} \frac{|a_n - b_n|^2}{n+1}\right)^{1/2}$ is attained for the sequence $(b_n)_{n\geq 2}$ of complex numbers with $|b_n| = x_n$ and $\arg b_n = \arg a_n$, that is for $b_n = x_n e^{i \arg a_n}$, $n \geq 2$ (if $a_n = 0$, from the proof of Theorem 5.3.5 we have $n \in \mathcal{I}^c$, so $x_n = 0$ and $b_n = x_n e^{i\arg a_n} = 0$ is unambiguously defined).

Since $b_n = 0$ for $n \in \mathcal{I}^c$ and $|b_n| = |x_n| = x_n \geq 0$ for $n \in \mathcal{I}$, we obtain

$$\begin{aligned} \sum_{n=2}^{\infty} (n^2|b_n|) &= \sum_{n\in\mathcal{I}} \left[n^2 \left(|a_n| - \frac{\sum_{m\in\mathcal{I}}(m^2|a_m|) - 1}{\sum_{m\in\mathcal{I}}(m^4(m+1))} n^2(n+1) \right) \right] \\ &= \sum_{n\in\mathcal{I}} (n^2|a_n|) - \frac{\sum_{m\in\mathcal{I}}(m^2|a_m|) - 1}{\sum_{m\in\mathcal{I}}(m^4(m+1))} \sum_{n\in\mathcal{I}} (n^4(n+1)) \\ &= 1, \end{aligned}$$

so $g(z) = z + \sum_{n=2}^{\infty} b_n z^n \in \mathcal{K}^*$, and

$$\begin{aligned} \left(\int_U |f(x+iy) - g(x+iy)|^2 \, dxdy \right)^{1/2} &= \left(\pi \sum_{n=2}^{\infty} \frac{|a_n - b_n|^2}{n+1} \right)^{1/2} \\ &= \left(\pi \sum_{n=2}^{\infty} \frac{(|a_n| - |b_n|)^2}{n+1} \right)^{1/2} \\ &= \left(\pi \sum_{n\in\mathcal{I}^c} \frac{|a_n|^2}{n+1} + \pi \frac{\left(\sum_{n\in\mathcal{I}} n^2|a_n| - 1\right)^2}{\sum_{n\in\mathcal{I}} n^4(n+1)} \right)^{1/2} \\ &= \operatorname{dist}(f, \mathcal{K}^*) \end{aligned}$$

as needed, concluding the proof. □

As an application of the previous theorem, consider the following.

Example 5.5.2. *Let $f_a : U \to \mathbb{C}$ be defined by $f_a(z) = z + az^2$, where $a \in \mathbb{C}$ is a constant.*

If $4|a| \leq 1$, then $f_a \in \mathcal{K}^*$ and $\operatorname{dist}(f_a, \mathcal{K}^*) = 0$.

If $4|a| > 1$, from Theorem 5.5.1 we obtain $\mathcal{P} = \{2\}$, $i_2 = 2$, $N = 2$, and $\mathcal{I} = \{i_2\} = \{2\}$, so $\operatorname{dist}(f_a, \mathcal{K}^*) = \frac{(4|a|-1)\sqrt{\pi}}{4\sqrt{3}}$ is attained for the function $g_a(z) = z + \frac{1}{4} e^{i\arg a} z^2 \in \mathcal{K}^* \subset \mathcal{K}$.

As another example, consider the following.

Example 5.5.3. *Let $f_{a,b} : U \to \mathbb{C}$ be defined by $f_{a,b}(z) = z + az^2 + bz^3$, where $a, b \in \mathbb{C}$ are constants.*

If $4|a| + 9|b| \leq 1$ then $f_{a,b} \in \mathcal{K}^*$ and $\operatorname{dist}(f_{a,b}, \mathcal{K}^*) = 0$.

If $4|a| + 9|b| > 1$, applying Theorem 5.5.1 it follows that if $3|a| \geq |b| + \frac{3}{4}$ we have $N = 2$ and $\mathcal{I} = \{i_2\} = \{2\}$, if $|b| \geq 3|a| + \frac{1}{9}$ we have $N = 2$ and $\mathcal{I} = \{i_2\} = \{3\}$, and in the rest of the cases we have $N = 3$ and $\mathcal{I} = \{i_2, i_3\} = \{2, 3\}$. We obtain

$$\operatorname{dist}(f_{a,b}, \mathcal{K}^*) = \begin{cases} \frac{\sqrt{\pi}}{4\sqrt{3}} \left((4|a| - 1)^2 + 12|b|^2 \right)^{1/2}, & \text{if } 3|a| \geq |b| + \frac{3}{4} \\ \frac{\sqrt{\pi}}{18} \left(108|a|^2 + (9|b| - 1)^2 \right)^{1/2}, & \text{if } |b| \geq 3|a| + \frac{1}{9} \\ \frac{\sqrt{\pi}}{2\sqrt{93}} |4|a| + 9|b| - 1|, & \text{otherwise} \end{cases},$$

attained for the function $g_{a,b} : U \to \mathbb{C}$ *(belonging to* $\mathcal{K}^* \subset \mathcal{K}$*), defined by*

$$g_{a,b}(z) = \begin{cases} z + \frac{1}{4} e^{i \arg a} z^2, & \text{if } 3\,|a| \geq |b| + \frac{3}{4} \\ z + \frac{1}{9} e^{i \arg b} z^3, & \text{if } |b| \geq 3\,|a| + \frac{1}{9} \\ z + \frac{27|a| - 9|b| + 1}{31} e^{i \arg a} z^2 + \frac{-12|a| + 4|b| + 3}{31} e^{i \arg b} z^3, & \text{otherwise} \end{cases}, \quad z \in U.$$

Figure 5.2 below shows a comparison of the images of the unit disk under the function $f_{a,b}$ *and under its best convex approximation function* $g_{a,b}$ *for some values of* a *and* b *in the three cases above.*

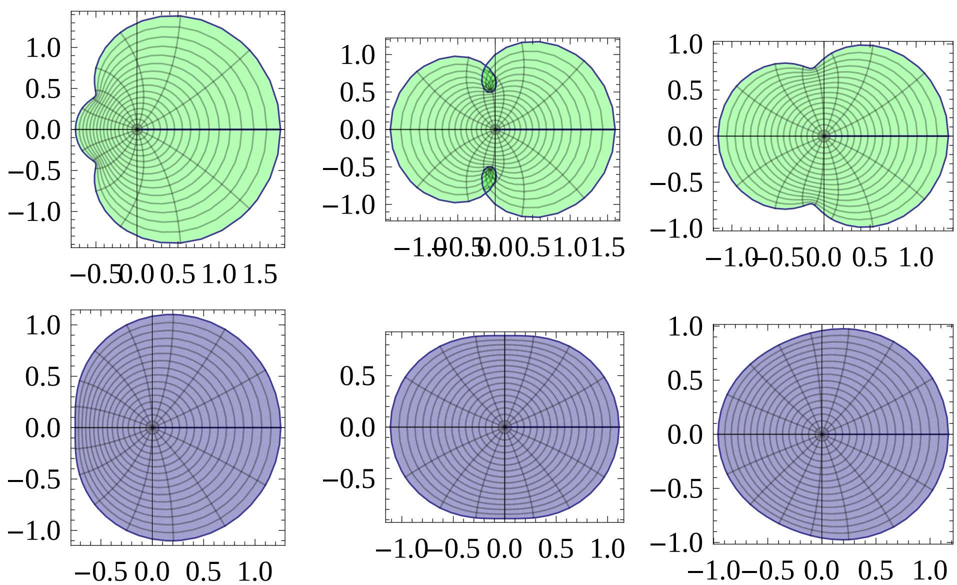

Figure 5.2: The image of the unit disk under $f_{a,b}$ (top row) and $g_{a,b}$ (bottom row), for $(a, b) = (0.5, 0.25)$ (left), $(a, b) = (0.1, 0.5)$ (center) and $(a, b) = (0.25, 0.5)$ (right).

We conclude with the observation that for $f \in \mathcal{A}$ for which $\operatorname{dist}(f, \mathcal{K}^*)$ is not too large, the best convex approximation of f given by Theorem 5.5.1 is in general a good approximation of f (see Figure 5.2), suitable for both practical problems and numerical implementation.

Chapter 6

Neighborhoods of univalent functions

In this chapter we consider the problem of studying the perturbations of a given univalent function. As a measure of the (non)univalency of a function we introduce the constant $K(f, D)$ associated with a function $f : D \subset \mathbb{C} \to \mathbb{C}$ analytic in a domain D, and we use it in order to show that a small perturbation of a univalent function is again a univalent function. As a consequence, a univalent function has a neighborhood consisting entirely of univalent functions.

As applications of the main result, we derive a corollary which is shown to be equivalent to the classical Noshiro-Warschawski-Wolff univalence criterion, and we present an application in terms of Taylor series.

6.1 Introduction

It is known that if $f : D \to \mathbb{C}$ is a univalent map in a domain D, then $f' \neq 0$ in D. The non-vanishing of the derivative of an analytic function (local univalence) is not in general sufficient to insure the univalence of the function, as it can be seen by considering for example the exponential function $f(z) = e^z$ defined in the upper half-plane.

The classical Noshiro-Warschawski-Wolff univalence criterion gives a partial converse of the above result, as follows:

Theorem 6.1.1. *If $f : D \to \mathbb{C}$ is analytic in the convex domain D and*

$$\operatorname{Re} f'(z) > 0, \qquad z \in D,$$

then f is univalent in D.

In this chapter we introduce the constant $K(f, D)$ associated with a function $f : D \to \mathbb{C}$ analytic in a domain D, which is a measure of the "degree of univalence" of f (see Proposition 6.2.1 and the remark following it).

Using the constant $K(f, D)$ thus introduced, in Theorem 6.2.4 we obtain a sufficient condition for univalence, which shows that a small perturbation of a univalent function is again univalent. As a theoretical consequence of this result, it follows that a univalent function has a neighborhood consisting entirely of univalent functions (see Remark 6.2.8).

The Theorem 6.2.4 is sharp, in the sense that we cannot replace the upper bound appearing in the hypothesis of this theorem by a larger one, as shown in Example 6.2.9.

For the particular choice of a linear function in Theorem 6.2.4, we obtain a simple sufficient condition for univalence (Corollary 6.2.6), which is shown to be equivalent to the Noshiro-Warschawski-Wolff univalence criterion. The main result in Theorem 6.2.4 can be viewed therefore as a generalization of this classical result, in which the linear function is replaced by a general univalent function.

The chapter concludes with another application of the main result in the case of analytic functions defined in the unit disk. Thus, in Theorem 6.2.11 and the corollary following it, we obtain sufficient conditions for the univalence of an analytic function defined in the unit disk in terms of the coefficients of its Taylor series representation, which might be of independent interest.

6.2 Main results

We denote by $U_r = \{z \in \mathbb{C} : |z| < r\}$ the open disk of radius $r > 0$ centered at the origin and we let $U = U_1$. The class of functions $f : D \to \mathbb{C}$ analytic in the domain D will be denoted by $\mathcal{A}(D)$.

Given a function $f : D \to \mathbb{C}$ analytic in the domain D we introduce the constant $K(f, D)$ defined as follows:

$$K(f,D) = \inf_{\substack{a,b\in D \\ a\neq b}} \left| \frac{f(a) - f(b)}{a - b} \right|. \tag{6.2.1}$$

It is immediate from the definition that if the function f is not univalent in D then $K(f, D) = 0$. The constant $K(f, D)$ characterizes the univalence of the function f in D in the following sense:

Proposition 6.2.1. *Let $f : D \to \mathbb{C}$ be an analytic function in the domain D. If $K(f, D) > 0$ then f is univalent in D.*

Conversely, if f is univalent in D and $\Omega \subset \overline{\Omega} \subset D$ is a bounded domain strictly contained in D, then $K(f, \Omega) > 0$.

Proof. See [70]. □

Remark 6.2.2. *Note that the converse in the above proposition may not hold for $\Omega = D$ without the additional hypothesis, as shown in the example below.*

In order to have the equivalence

$$f \text{ univalent in } D \iff K(f, D) > 0,$$

one needs additional hypotheses, which guarantee the existence of a continuous extension of f, f' to $\overline{D}$, such that f is injective on $\overline{D}$ and $f' \neq 0$ in $\overline{D}$.

For example, in the case $D = U$, if the boundary of the image domain $f(U)$ is a Jordan curve of class $C^{1,\alpha}$ $(0 < \alpha < 1)$, by Carathéodory theorem the function f has a continuous injective extension to $\overline{D}$, and also, by Kelogg-Warschawski theorem, the function f' has continuous extension to $\overline{D}$, with $f' \neq 0$ in $\overline{D}$ (see for example [76], p. 24 and pp. 48 – 49). Following the proof above with Ω replaced by U, we obtain $K(f, U) > 0$, and therefore in this case we have

$$f \text{ univalent in } U \iff K(f, U) > 0.$$

Example 6.2.3. *Let $D = U - [0, 1]$ be the unit disk with a slit along the positive real axis. Since D is simply connected, there exists a conformal map $f : U \to D$ between the unit disk U and D (see Figure 6.1 below). The map f has a continuous extension to $\overline{U}$, and without loss of generality we may assume that there exists $\theta \in (0, 2\pi)$ such that $f\left(e^{i\theta}\right) = f\left(e^{-i\theta}\right) \in (0, 1)$.*

The function f is univalent in U, but $K(f, U) = 0$ since

$$K(f,U) \leq \lim_{\substack{a\to e^{i\theta} \\ b\to e^{-i\theta}}} \left| \frac{f(a) - f(b)}{a - b} \right| = \left| \frac{f\left(e^{i\theta}\right) - f\left(e^{-i\theta}\right)}{e^{i\theta} - e^{-i\theta}} \right| = 0.$$

The main result is contained in the following:

Theorem 6.2.4. *Let $f : D \to \mathbb{C}$ be a non-constant analytic function in the convex domain D. If there exists an analytic function $g : D \to \mathbb{C}$ univalent in D such that*

$$|f'(z) - g'(z)| \leq K(g, D), \qquad z \in D, \tag{6.2.2}$$

then the function f is also univalent in D.

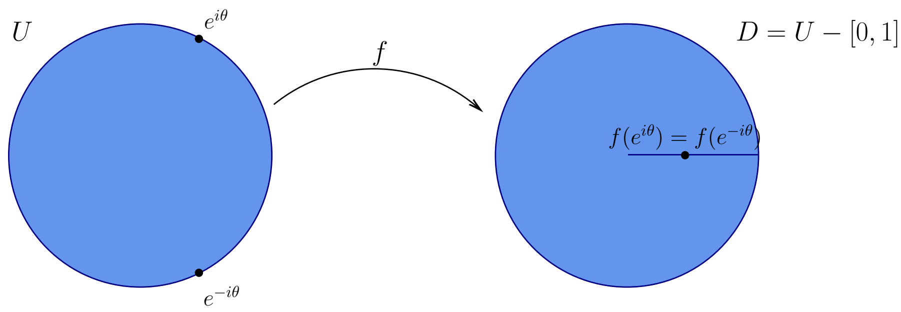

Figure 6.1: An example of a univalent function $f : U \to \mathbb{C}$ for which $K(f, U) = 0$.

Proof. Assuming that f is not univalent in D, there exist distinct points $z_{1,2} \in D$ such that $f(z_1) = f(z_2)$. Integrating the derivative of $f - g$ along the line segment $[z_1, z_2] \subset D$ and using the hypothesis (6.2.2) we obtain

$$\begin{aligned} |g(z_2) - g(z_1)| &= |(f(z_2) - g(z_2)) - (f(z_1) - g(z_1))| \\ &= \left| \int_{[z_1,z_2]} f'(z) - g'(z)\, dz \right| \\ &\le \int_{[z_1,z_2]} |f'(z) - g'(z)|\, |dz| \\ &\le \int_{[z_1,z_2]} K(g, D)\, |dz| \\ &= K(g, D)\, |z_1 - z_2|\,. \end{aligned}$$

Since the points $z_{1,2}$ are assumed to be distinct, from the definition of the constant $K(g, D)$ we obtain equivalently

$$\left| \frac{g(z_2) - g(z_1)}{z_2 - z_1} \right| \le K(g, D) = \inf_{\substack{a,b \in D \\ a \ne b}} \left| \frac{g(a) - g(b)}{a - b} \right| \le \left| \frac{g(z_2) - g(z_1)}{z_2 - z_1} \right|, \tag{6.2.3}$$

and therefore

$$K(g, D) = \inf_{\substack{a,b \in D \\ a \ne b}} \left| \frac{g(a) - g(b)}{a - b} \right| = \left| \frac{g(z_2) - g(z_1)}{z_2 - z_1} \right|. \tag{6.2.4}$$

Consider now the auxiliary function $G : D - \{z_2\} \to \mathbb{C}$ defined by

$$G(z) = \frac{g(z) - g(z_2)}{z - z_2}, \qquad z \in D - \{z_2\}, \tag{6.2.5}$$

and note that since g is analytic in D, G is also analytic in $D - \{z_2\}$ and moreover the limit

$$\lim_{z \to z_2} G(z) = \lim_{z \to z_2} \frac{g(z) - g(z_2)}{z - z_2} = g'(z_2) \tag{6.2.6}$$

exists and it is finite. The function G can be therefore extended by continuity to an analytic function in D, denoted also by G.

Since

$$\inf_{z \in D} |G(z)| = \inf_{\substack{z \in D \\ z \ne z_2}} |G(z)| = \inf_{\substack{z \in D \\ z \ne z_2}} \left| \frac{g(z) - g(z_2)}{z - z_2} \right| \ge \inf_{\substack{a,b \in D \\ a \ne b}} \left| \frac{g(a) - g(b)}{a - b} \right| = K(g, D),$$

combining with (6.2.4) we obtain that

$$\inf_{z\in D}|G(z)| \geq K(g,D) = \left|\frac{g(z_2)-g(z_1)}{z_2-z_1}\right| = |G(z_1)| \geq \inf_{z\in D}|G(z)|,$$

which shows that minimum value of the modulus of G in D is attained at z_1:

$$\inf_{z\in D}|G(z)| = |G(z_1)|.$$

However, since the function g is univalent in D, from the definition of G it follows that $G(z) \neq 0$ for any $z \in D - \{z_2\}$, and also $G(z_2) = g'(z_2) \neq 0$, and therefore the function G does not vanish in D. Applying the maximum modulus principle to the analytic function $1/G$ it follows that $|G|$ must be constant in D, and therefore G is constant in D.

It follows that

$$g(z) = g(z_2) + c(z - z_2), \qquad z \in D, \tag{6.2.7}$$

for a certain constant $c \in \mathbb{C}$ (from the definition of G it can be seen that the constant c can be written in the form $c = g'(z_2)e^{i\theta}$, for some $\theta \in \mathbb{R}$).

The relation (6.2.7) shows that g is a linear function, and therefore the constant $K(g,D)$ becomes in this case

$$\begin{aligned}
K(g,D) &= \inf_{\substack{a,b\in D\\ a\neq b}}\left|\frac{g(a)-g(b)}{a-b}\right| \\
&= \inf_{\substack{a,b\in D\\ a\neq b}}\left|\frac{(g(z_2)+c(a-z_2))-(g(z_2)+c(b-z_2))}{a-b}\right| \\
&= \inf_{\substack{a,b\in D\\ a\neq b}}\left|\frac{c(a-b)}{a-b}\right| \\
&= |c|.
\end{aligned}$$

The hypothesis (6.2.2) of the theorem can be written therefore as follows

$$|f'(z) - c| \leq |c|, \qquad z \in D,$$

which shows that either f is linear in D (and thus univalent, since f is assumed to be non-constant in D), or the following strict inequality holds

$$|f'(z) - c| < |c|, \qquad z \in D.$$

Repeating the proof above with $g(z) \equiv cz$ we obtain

$$\begin{aligned}
|cz_2 - cz_1| &= |(f(z_2) - cz_2) - (f(z_1) - cz_1)| \\
&= \left|\int_{[z_1,z_2]} f'(z) - cdz\right| \\
&\leq \int_{[z_1,z_2]} |f'(z) - c|\,|dz| \\
&< |c|\,|z_2 - z_1|,
\end{aligned}$$

a contradiction.

The contradiction obtained shows that the function f is univalent in D, concluding the proof of the theorem. □

In the particular case $D = U$, from the previous thereom we obtain immediately the following sufficient criterion for univalence in the unit disk:

Theorem 6.2.5. *Let $f : U \to \mathbb{C}$ be a non-constant analytic function in the unit disk. If there exists an analytic function $g : U \to \mathbb{C}$ univalent in U such that*

$$|f'(z) - g'(z)| \leq K(g, U), \qquad z \in U, \tag{6.2.8}$$

then the function f is also univalent in U.

As a corollary of Theorem 6.2.4 we obtain the following:

Corollary 6.2.6. *If $f : D \to \mathbb{C}$ is non-constant and analytic in the convex domain D and there exists $c > 0$ such that*

$$|f'(z) - c| \leq c, \qquad z \in D, \tag{6.2.9}$$

then f is univalent in D.

Proof. Follows from Theorem 6.2.4 by considering the univalent function $g : D \to \mathbb{C}$ defined by $g(z) = cz$, for which we have

$$K(g, D) = \inf_{\substack{a,b \in D \\ a \neq b}} \left| \frac{g(a) - g(b)}{a - b} \right| = \inf_{\substack{a,b \in D \\ a \neq b}} \left| \frac{ca - cb}{a - b} \right| = c.$$

□

Remark 6.2.7. *Let us note that the previous corollary can also be obtained as a direct consequence of the classical Noshiro-Warschawski-Wolff univalence criterion, since the hypothesis (6.2.9) implies the hypothesis*

$$\operatorname{Re} f'(z) > 0, \qquad z \in D. \tag{6.2.10}$$

of this theorem (the fact that the above inequality is a strict inequality follows from the maximum principle, the function f being assumed to be non-constant in D).

Conversely, the Noshiro-Warschawski-Wolff univalence criterion follows from the previous corollary. To see this, without loss of generality we may assume $0 \in D$, and in order to prove the univalence of f, it suffices to prove the univalence of f in $D_r = rD \subset D$, for an arbitrarily fixed $r \in (0, 1)$.

If the condition (6.2.10) holds, there exists $c = c(r) > 0$ such that

$$f'(D_r) \subset \{w \in \mathbb{C} : |w - c| < c\},$$

or equivalent

$$|f'(z) - c| < c, \qquad z \in D_r.$$

Applying Corollary 6.2.6 to the restriction of of f to D_r, it follows that the function f is univalent in D_r. Since $r \in (0, 1)$ was arbitrarily fixed, it follows that f is univalent in U, concluding the proof of the claim.

The remark above shows that Corollary 6.2.6 and the Noshiro-Warschawski-Wolff univalence criterion are equivalent, and therefore Theorem 6.2.4 is a generalization of it. The Noshiro-Warschawski-Wolff univalence criterion can be viewed therefore as a particular case of Theorem 6.2.4, corresponding to the choice of a linear function g.

Remark 6.2.8. *Fixing an arbitrarily univalent function $g : U \to \mathbb{C}$ for which $K(g, U) \neq 0$ (see Remark 6.2.2 above), Theorem 6.2.5 shows that a whole neighborhood*

$$V(g) = \{f \in \mathcal{A} : ||f' - g'|| \leq K(g, U)\}$$

of g consists entirely of univalent functions in U ($||\cdot||$ denotes here the supremum norm in the space $\mathcal{A}_0 = \{f \in \mathcal{A} : f(0) = 0\}$ of normalized analytic functions). Loosely stated, Theorem 6.2.5 shows that an univalent function has a neighborhood consisting entirely of univalent functions.

The hypotheses of Theorem 6.2.4 and Theorem 6.2.5 are sharp, in the sense that we cannot replace the right side of the inequalities (6.2.2), respectively (6.2.8), by larger constants, as can be seen from the following example.

Example 6.2.9. *Consider the function $f : U \to \mathbb{C}$ defined by $f(z) = z + az^2$, $z \in U$, where $a \in \mathbb{C}$ is a parameter.*

Using Theorem 6.2.5 above with $g(z) \equiv z$, for which $K(g, U) = 1$, we obtain that the function f is univalent in U if

$$|2az| \leq 1, \qquad z \in U,$$

that is if $|2a| \leq 1$.

This result is sharp, since the function f is univalent iff $|a| \leq \frac{1}{2}$, as it can be checked by direct computation.

The univalence of the function f in the previous example can also be obtained by using the Noshiro-Warschawski-Wolff univalence criterion (for $|a| \leq 1/2$ we have $\operatorname{Re} f'(z) > 0$ for any $z \in U$). The next example shows that we may still use Theorem 6.2.5 also in situations when the Noshiro-Warschawski-Wolff univalence criterion cannot be applied:

Example 6.2.10. *Consider the linear map $g : U \to \mathbb{C}$ defined by $g(z) = \frac{z}{1-z}$. The function g is univalent in U and we have*

$$K(g, U) = \inf_{\substack{a,b \in U \\ a \neq b}} \left| \frac{g(a) - g(b)}{a - b} \right| = \inf_{\substack{a,b \in U \\ a \neq b}} \left| \frac{\frac{a}{1-a} - \frac{b}{1-b}}{a - b} \right| = \inf_{\substack{a,b \in U \\ a \neq b}} \frac{1}{|1-a|\,|1-b|} = \frac{1}{4}.$$

The function $f : U \to \mathbb{C}$ defined by $f(z) = \frac{z^2}{8} + \frac{z}{1-z}$ is analytic in U and satisfies

$$|f'(z) - g'(z)| = \left| \frac{z}{4} \right| < \frac{1}{4} = K(g, U), \qquad z \in U,$$

and therefore by Theorem 6.2.5 it follows that f is univalent in the unit disk.

The univalence of f does not follow however by the Noshiro-Warschawski-Wolff univalence criterion since $\operatorname{Re} f'(z)$ takes (arbitrarily small) negative values for $z \in U$ sufficiently close to 1.

As another application of Theorem 6.2.5, in the next result we show that by perturbing the coefficients of the Taylor series of an univalent function, the resulting function is also univalent. More precisely, we have the following:

Theorem 6.2.11. *Let $g : U \to \mathbb{C}$ be an analytic univalent function with Taylor series representation*

$$g(z) = \sum_{n=0}^{\infty} b_n z^n, \qquad z \in U. \tag{6.2.11}$$

If the coefficients $a_0, a_1, \ldots \in \mathbb{C}$ satisfy the inequality

$$\sum_{n=1}^{\infty} n\,|a_n - b_n| < K(g, U) \tag{6.2.12}$$

then the function $f : U \to \mathbb{C}$ defined by

$$f(z) = \sum_{n=0}^{\infty} a_n z^n, \qquad z \in U, \tag{6.2.13}$$

is analytic and univalent in U.

Proof. Since g is univalent in U, the radius of convergence of the Taylor series (6.2.11) is at least 1, hence

$$\limsup \sqrt[n]{|b_n|} \leq 1,$$

and therefore given $\varepsilon > 0$ we have $|b_n| \leq 1 + \varepsilon$ for all n sufficiently large.

Using the hypothesis (6.2.12) we obtain

$$\limsup \sqrt[n]{|a_n|} \leq \limsup \sqrt[n]{|b_n| + |a_n - b_n|} \leq \limsup \sqrt[n]{1 + \varepsilon + \frac{K(g,U)}{n}} = 1,$$

and therefore the radius of convergence of the series in (6.2.13) is at least 1, thus the function f is well defined by (6.2.13) and it is analytic in U.

Since

$$\begin{aligned} |f'(z) - g'(z)| &= \left| \sum_{n=0}^{\infty} n a_n z^{n-1} - \sum_{n=0}^{\infty} n b_n z^{n-1} \right| \\ &\leq \sum_{n=1}^{\infty} n |a_n - b_n| |z|^{n-1} \\ &\leq \sum_{n=1}^{\infty} n |a_n - b_n| \\ &< K(g,U), \end{aligned}$$

for any $z \in U$, by Theorem 6.2.5 follows that f is univalent in U, concluding the proof. □

Using a comparison with the generalized harmonic series, from the above we can obtain the following:

Corollary 6.2.12. *Let $g : U \to \mathbb{C}$ be an analytic univalent function with Taylor series representation*

$$g(z) = \sum_{n=0}^{\infty} b_n z^n, \qquad z \in U. \tag{6.2.14}$$

If the coefficients $a_0, a_1, \ldots \in \mathbb{C}$ satisfy the inequality

$$|a_n - b_n| < K(g,U) \frac{\zeta(p)}{n^{p+1}}, \qquad n = 1, 2, \ldots, \tag{6.2.15}$$

for some $p > 1$ (ζ denotes the Riemann zeta function), then the function $f : U \to \mathbb{C}$ defined by

$$f(z) = \sum_{n=0}^{\infty} a_n z^n, \qquad z \in U, \tag{6.2.16}$$

is analytic and univalent in U.

Example 6.2.13. *Considering the function $g(z) = \frac{z}{1-z} = \sum_{n=1}^{\infty} z^n$ defined in Example 6.2.10, which is analytic and univalent in U and has $K(g,U) = 1$, from the previous theorem it follows that the function $f : U \to \mathbb{C}$ defined by $f(z) = \sum_{n=0}^{\infty} a_n z^n$ is analytic and univalent in U if the coefficients a_n satisfy the inequality*

$$\sum_{n=1}^{\infty} n |a_n - 1| < 1.$$

Using for example the fact that $\zeta(2) = \frac{\pi^2}{6} \approx 1.645$, from the previous corollary it follows that the function f is also analytic and univalent in U if the coefficients a_n satisfy the inequality

$$|a_n - 1| \leq \frac{\pi^2}{6n^3} \approx \frac{1.645}{n^3}, \qquad n = 1, 2, \ldots$$

Chapter 7

Achievements and plans for further career development

7.1 Scientific and professional achievements

Mihai N. Pascu, the author of the present Habilitation Thesis graduated from Faculty of Sciences of *Transilvania* University of Braşov in 1995, with a B. S. in the field of Mathematics and Computer Science, GPA 10.00. One year later, in 1996, we also received a Master degree in "Statistics, Probability and Systems reliability" from the same institution, also with a GPA of 10.00.

Between 1995 - 2001 the author was a Master, respectively a Ph. D. student of the University of Connecticut at Storrs, period of time in which he both prepared for this specializations, and taught various classes to undergraduate students in the Department of Mathematics. His activity was recognized with *Louis J. de Luca Award* memorial award for excellence in teaching.

In 1998 he received a second Master degree in *Mathematics* from the University of Storrs, Connecticut, USA. In 2001, under the guidance of Prof. Richard F. Bass (associated advisors William Abikoff and Evarist Gine), he received a Ph. D. in Mathematics from the same university.

In 2002, he received a second Ph. D. in mathematics from *Transilvania* University of Braşov (advisor Prof. Gabriel V. Orman).

Between 2001 - 2004 the author obtain a visiting research assistant professor position in the Department of Mathematics at Purdue University, where he taught classes to undergraduate and graduate students of this university, and conducted scientific research, especially with Rodrigo Bañuelos. As a member of this institution, he won by competition a National Science Foundation - Division of Mathematical Sciences research grant (2002 - 2004) in the area of Stochastic Processes (Brownian motion).

In 2004, the author left Purdue University and returned to the Faculty of Mathematics and Computer Science, *Transilvania* University of Braşov, as an Assistant Professor. He then became Associate Professor in 2007, and he acted at this institution in this position since then.

Between 2010 - 2012, under the supervision of Prof. L. Beznea, the author completed a post-doctoral research program at the partener institution "Simion Stoilow" Institute of Mathematics of the Romanian Academy, of the contract POSDRU/89/1.5/S/62988 of "Costin C. Kiriţescu" National Institute of Economic Research.

The author published 5 books (3 monographs and 2 proceedings), in the area of Stochastic Processes and Complex Analysis, as follows.

1. M. N. Pascu and N. R. Pascu, *Probleme şi soluţii în Analiza complexă*, Transilvania University Press, Braşov, 2011, ISBN 978-973-598-924-8.

2. M. N. Pascu, *Calcul stochastic, mişcare Browniană şi aplicaţii*, Transilvania University Press, Braşov, 2010, ISBN 978-973-598-749-7.

3. M. N. Pascu, G. V. Orman (editors), *Proceedings of the 23rd Scientific Session "Mathematics and Its Applications"*, Transilvania University of Braşov, Brasov, May 8-9, 2009, Transilvania University Press, Braşov, 2009, ISSN 1843-6994.

4. M. N. Pascu, *Brownian motion and Applications*, Transilvania University Press, Braşov, 2006, ISBN 973-635-828-3.

5. M. N. Pascu, S. Owa (editors), *Proceedings of the International Symposium "Complex Function Theory and Applications"*, Transilvania University of Braşov, Braşov, September 1 - 5, 2006, Transilvania University Press, Brasov, 2006, ISBN 973-635-827-5

The author won by competition (as director/coordinator) several research grants, as follows:

1. *Stochastic Analysis and Parameter Estimation in Systems with memory* (grant coordinator at *Transilvania* University of Braşov, partener of ASE Bucureşti), grant CNCSIS - PNII-ID-PCCE-2011-2-0015, 2012 - 2015

2. *Mişcarea Browniană şi aplicaţii: proprietăţi de monotonie şi extrem*, grant CNCSIS - PNII ID_209, 2007 - 2010

3. *Proprietăţi de monotonie, principii de extrem şi aplicaţii*, grant CNCSIS - AT 61, 2006 - 2007

4. *Reflecting Brownian Motion in Convex Domains*, NSF - DMS #0203961, 2002 - 2004

and he was a member of several other research grants:

5. *Randomness, Geometric Problems and Functional Inequalities*, director I. Popescu, grant CNCSIS PN-II-RU-TE-2011-3-0259, 2012 - 2015.

6. *Analiză complexă şi domenii conexe*, director G. Sălăgean, contract CEEX (2-CEx06-11-10/25.07.06), 2006 - 2008.

7. *Contribuţii la teoria geometrică a funcţiilor analitice*, director N. N. Pascu, grant CNCSIS - A, 2004 - 2006.

8. *Contribuţii la teoria geometrică a funcţiilor univalente*, director N. N. Pascu, grant CNCSIS - A, 2001 - 2003.

As a director of these grants, the author supervised (with the accord of the Ph. D. advisor) the research activity of several Ph. D. students. As a result, two of these students obtained their Ph.D. under his guidance: Maria Gageonea (2009) and Alina Nicolaie (2012). This proves the ability of the author to advise Ph.D. students in the future. Mrs. Gageonea in now teaching at University of Connecticut at Storrs (USA), and Ms. Nicolaie completed a post-doctoral program at the Institute of Statistics, Biostatistics and Actuarial Sciences of the Catholic University of Louvain (Belgium), an is now a Visiting Scholar at University of Michigan School of Public Health (Belgium).

The author was an organizer of several prestigious conferences, as follows.

1. Co-organizer (with K. Burdzy, University of Washington, SUA) of the special session *Probability and its Relation to Other Fields of Mathematics*, in the Joint International Meeting of the American Mathematical Society and Romanian Mathematical Society, "1 Decembrie 1918" University of Alba Iulia, 27 - 30 June 2013.

2. Co-organizer (with Fr. Russo, Universite Paris 13, France) of the special session *Processus stochastiques* in the 10th Colloque Franco-Roumain de Mathèmatiques Appliquèes, Universite de Poitiers (France), 26 - 31 August 2010.

3. Co-organizer (with C. Tudor, Universite de Lille, Frana) of the Workshop on Mathematics " PDEs and Stochastic Processes", *Transilvania* University of Braşov, 10 November 2012.

4. Co-organizer (with S. Owa, Kinki University, Japan), of the conference *International Symposium Complex Function Theory and Applications*, *Transilvania* University of Braşov, 1 - 5 September 2006.

5. Co-organizer of several workshops in the area of tochastic processes (IMAR - Bucureşti, *Transilvania* University of Braşov, University of Piteşti, etc)

The author published 38 scientific research papers, of which 14 in the last 5 years (of these, 6 papers were published in ISI journals). In recognition for his research contributions in the area of Brownian motions, the author received in 2013 the "Dimitrie Pompeiu" prize of the Romanian Academy.

Other activities and professional memberships:

- member of the American Mathematical Society and of the Romanian Mathematical Society;
- reviewer for several mathematics journals (Annals of Mathematics, Proceedings of the London Mathematical Society, Electronic Communications in Probability, Potential Analysis, Mathematical Reports, etc);
- CNCSIS reviewer for research grants in mathematics;
- visiting professor (Department of Mathematics, François Rabelais University of Tours, France, 14 - 28 May, 2006, Institute of Mathematics, Budapest University of Technology and Economics, Budapest, March 1 - May 31, 2012), etc;
- organizer of several international conferences;
- invited speaker at several prestigious conferences.

7.2 Open problems and future plans

We begin by presenting some open problems related to the research presented in the previous chapters.

As indicated in Chapter 1 (Section 1.5), a very interesting problem which drew the attention of many researchers in both Analysis and Probability is the Hot Spots conjecture, which is still open in its full generality.

Although it is widely believed that the conjecture is true, and there are recent advances in this research area (see for example the Polymath Project web page or the recent papers [59], [42], [44], [77], or [18], to mention just a few), a proof of it is still missing. This suggests that new tools for approaching the conjecture are needed. To attack the problem, perhaps a first thing to do is to try to solve the conjecture in the case of acute triangles (the conjecture is known to be true for obtuse triangles, see [7]), and understand better the role played by the acute/obtuse angles in the proof. From the point of view of couplings of Brownian motions, the reason for which the conjecture can be proved just for obtuse triangles (and not for acute ones) is that the mirror coupling preserves the left/right starting position of the Brownian motions, and it does not do so in the case of acute triangles. An idea that might lead to a resolution of this problem is that instead of using the mirror/synchronous couplings, to construct a new fixed-distance coupling (similar to the one constructed in Chapter 3 in the case of complete manifolds, but for the case of reflecting Brownian motion in polygonal domains, viewed as a two-sided flat manifold), and to use its properties in order to derive the validity of the conjecture.

In Section 2.3 we presented a resolution of the Laugesen-Morpurgo conjecture for unit ball in $\mathbb{R}^n$ ($n \geq 1$), and from this result we derived in particular the validity of the Hot Spots conjecture in the case of the unit ball (see [66]). One challenging and interesting open problem would be to try to formulate and then to prove a corresponding Laugesen-Morpurgo conjecture for general convex domains. That is, for a given smooth convex domain $D \subset \mathbb{R}^n$ to find a family of curves along which the diagonal $p_D(t, x, x)$ of the Neumann heat kernel for D increases for all times $t > 0$.

In turn, as in [66], this would also give a resolution of the Hot Spots conjecture. Note that our results in Section 2.3 identify (in the case of the unit ball) the family of curves along which the diagonal of the Neumann heat kernel is monotone with the diameters of the ball.

In Section 1.6 we studied the Brownian motion with killing and reflection in a domain D (Neumann conditions on a part of the boundary and Dirichlet conditions on the remaining one), and we found sufficient condition under which the lifetime of Brownian motion is monotone on a certain family of curves. We showed that if the domain is convex, the two parts of the boundary meet at acute angles, and one of them is a line segment of an arc of a circle, then the lifetime of Brownian motion is monotone along hyperbolic geodesics in the domain. One interesting problem would be to extend this result beyond the conditions we found in [10]. In turn, this and additional information on the nodal line of the second Neumann eigenfunction of the Laplacian could give a resolution of the Hot Spots conjecture.

In Chapter 4 we showed that the Maximum modulus principle for analytic functions can be extended to certain classes of non-analytic functions. Interesting open problems here are the parallel extension of the known theory of analytic functions to these classes of functions (i.e. Schwarz lemma, univalence criteria, coefficient inequalities, aso), and the problem of finding more general classes of functions and operators for which the maximum principle holds.

In Chapter 5 we introduced a method for obtaining the best univalent approximation of analytic functions in certain subclasses of analytic functions. The method used the fact that the corresponding classes of functions can be described by certain inequalities on the coefficients of the Taylor series development of the functions. A possible line of research for extending these results is to find conditions on the coefficients of the Taylor series development which guarantee the univalence of the function, and then to use the same ideas in order to find the best univalent approximation of an analytic function in these newly defined classes of functions.

In Chapter 6 we introduced the constant $K(f, D)$ which characterizes the univalency of the analytic function $f : D \to \mathbb{C}$, and using it we derived sufficient conditions for univalence. The method seems to be a powerful one, since a corollary of our main result (Corollary 6.2.6) was shown to be equivalent to the classical Noshiro-Warschawski-Wolff univalence criterion. One interesting line of research would be continue the study of this constant, and to try to derive new sufficient conditions for univalence involving it.

Regarding the future plans for professional growth and development, we have in mind the following activities:

- leading the research activity of Ph. D. students at the host institution, thus contributing to the awareness, growth, and development of Mathematics;
- organizing a strong research seminar at the host institution, for the benefit of both senior and young researchers, where they can find a stimulating research environment and a place for disseminating their most recent research findings;
- obtaining a full professorship in Mathematics and contributing in this position, as a member of the Romanian Mathematical School, to its international recognition and prestige through active research and publications in the field of Mathematics.

Bibliography

[1] Marc Arnaudon, Koléhè Abdoulaye Coulibaly, and Anton Thalmaier. Horizontal diffusion in C^1 path space. In *Séminaire de Probabilités XLIII*, volume 2006 of *Lecture Notes in Math.*, pages 73–94. Springer, Berlin, 2011.

[2] Marc Arnaudon, Anton Thalmaier, and Feng-Yu Wang. Harnack inequality and heat kernel estimates on manifolds with curvature unbounded below. *Bull. Sci. Math.*, 130(3):223–233, 2006.

[3] Rami Atar. Invariant wedges for a two-point reflecting Brownian motion and the "hot spots" problem. *Electron. J. Probab.*, 6:no. 18, 19 pp. (electronic), 2001.

[4] Rami Atar and Krzysztof Burdzy. On Neumann eigenfunctions in lip domains. *J. Amer. Math. Soc.*, 17(2):243–265 (electronic), 2004.

[5] Rami Atar and Krzysztof Burdzy. Mirror couplings and Neumann eigenfunctions. *Indiana Univ. Math. J.*, 57(3):1317–1351, 2008.

[6] Catherine Bandle. *Isoperimetric inequalities and applications*, volume 7 of *Monographs and Studies in Mathematics*. Pitman (Advanced Publishing Program), Boston, Mass.-London, 1980.

[7] Rodrigo Bañuelos and Krzysztof Burdzy. On the "hot spots" conjecture of J. Rauch. *J. Funct. Anal.*, 164(1):1–33, 1999.

[8] Rodrigo Bañuelos, Tadeusz Kulczycki, and Bartlomiej Siudeja. Neumann bessel heat kernel monotonicity. *Potential Anal.*, 30(1):65–83, 2009.

[9] Rodrigo Bañuelos and Michael Pang. An inequality for potentials and the "hot-spots" conjecture. *Indiana Univ. Math. J.*, 53(1):35–47, 2004.

[10] Rodrigo Bañuelos, Michael Pang, and **Mihai N. Pascu**. Brownian motion with killing and reflection and the "hot-spots" problem. *Probab. Theory Related Fields*, 130(1):56–68, 2004.

[11] Richard F. Bass. *Probabilistic techniques in analysis.* Probability and its Applications (New York). Springer-Verlag, New York, 1995.

[12] Richard F. Bass and Krzysztof Burdzy. On domain monotonicity of the Neumann heat kernel. *J. Funct. Anal.*, 116(1):215–224, 1993.

[13] Richard F. Bass and Krzysztof Burdzy. Fiber Brownian motion and the "hot spots" problem. *Duke Math. J.*, 105(1):25–58, 2000.

[14] Richard F. Bass and Elton P. Hsu. Pathwise uniqueness for reflecting Brownian motion in Euclidean domains. *Probab. Theory Related Fields*, 117(2):183–200, 2000.

[15] Richard F. Bass and Pei Hsu. Some potential theory for reflecting Brownian motion in Hölder and Lipschitz domains. *Ann. Probab.*, 19(2):486–508, 1991.

[16] Itai Benjamini, Krzysztof Burdzy, and Zhen-Qing Chen. Shy couplings. *Probab. Theory Related Fields*, 137(3-4):345–377, 2007.

[17] Maury Bramson, Krzysztof Burdzy, and Wilfrid Kendall. Shy couplings, CAT(0) spaces, and the Lion and Man. *Ann. Probab.*, 41(2):744–784, 2013.

[18] Lorenzo Brasco, Rolando Magnanini, and Paolo Salani. The location of the hot spot in a grounded convex conductor. *Indiana Univ. Math. J.*, 60(2):633–659, 2011.

[19] David R. Brillinger. A particle migrating randomly on a sphere. *J. Theoret. Probab.*, 10(2):429–443, 1997. Dedicated to Murray Rosenblatt.

[20] Krzysztof Burdzy. The hot spots problem in planar domains with one hole. *Duke Math. J.*, 129(3):481–502, 2005.

[21] Krzysztof Burdzy. Neumann eigenfunctions and Brownian couplings. In *Potential theory in Matsue*, volume 44 of *Adv. Stud. Pure Math.*, pages 11–23. Math. Soc. Japan, Tokyo, 2006.

[22] Krzysztof Burdzy and Zen-Qing Chen. Weak convergence of reflecting Brownian motions. *Electron. Comm. Probab.*, 3:29–33 (electronic), 1998.

[23] Krzysztof Burdzy and Zhen-Qing Chen. Coalescence of synchronous couplings. *Probab. Theory Related Fields*, 123(4):553–578, 2002.

[24] Krzysztof Burdzy and Zhen-Qing Chen. Discrete approximations to reflected Brownian motion. *Ann. Probab.*, 36(2):698–727, 2008.

[25] Krzysztof Burdzy and Wilfrid S. Kendall. Efficient Markovian couplings: examples and counterexamples. *Ann. Appl. Probab.*, 10(2):362–409, 2000.

[26] Krzysztof Burdzy and Wendelin Werner. A counterexample to the "hot spots" conjecture. *Ann. of Math. (2)*, 149(1):309–317, 1999.

[27] René A. Carmona and Wei An Zheng. Reflecting Brownian motions and comparison theorems for Neumann heat kernels. *J. Funct. Anal.*, 123(1):109–128, 1994.

[28] Isaac Chavel. *Eigenvalues in Riemannian geometry*, volume 115 of *Pure and Applied Mathematics*. Academic Press, Inc., Orlando, FL, 1984. Including a chapter by Burton Randol, With an appendix by Jozef Dodziuk.

[29] Isaac Chavel. Heat diffusion in insulated convex domains. *J. London Math. Soc. (2)*, 34(3):473–478, 1986.

[30] Jeff Cheeger and David G. Ebin. *Comparison theorems in Riemannian geometry*. North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York, 1975. North-Holland Mathematical Library, Vol. 9.

[31] Hee Je Cho and Seong-Hun Paeng. Ollivier's Ricci curvature and the coloring of graphs. *European J. Combin.*, 34(5):916–922, 2013.

[32] Michael Cranston. Gradient estimates on manifolds using coupling. *J. Funct. Anal.*, 99(1):110–124, 1991.

[33] Manfredo P. do Carmo. *Riemannian Geometry*. Birkhäuser, Boston, MA, 1992.

[34] Peter L. Duren. *Univalent functions*, volume 259 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1983.

[35] Hans J. Engelbert and Wolfgang Schmidt. On solutions of one-dimensional stochastic differential equations without drift. *Z. Wahrsch. Verw. Gebiete*, 68(3):287–314, 1985.

[36] Maria E. Gageonea, Shigeyoshi Owa, Radu N. Pascu, and **Mihai N. Pascu**. A maximum modulus principle for non-analytic functions defined in the unit disk. *Appl. Math. Comput.*, 187(1):163–169, 2007.

[37] Maria E. Gageonea, **Mihai N. Pascu**, and Nicolae R. Pascu. A maximum modulus principle for a class of non-analytic functions defined in the unit disk. *Mathematica*, 49(72)(2):169–174, 2007.

[38] Ian Graham and Gabriela Kohr. *Geometric function theory in one and higher dimensions*, volume 255 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, Inc., New York, 2003.

[39] Elton P. Hsu. A domain monotonicity property of the Neumann heat kernel. *Osaka J. Math.*, 31(1):215–223, 1994.

[40] Elton P. Hsu. *Stochastic analysis on manifolds*, volume 38 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.

[41] Marius Iosifescu. *Finite Markov processes and their applications*. John Wiley & Sons, Ltd., Chichester; Editura Tehnică, Bucharest, 1980. Wiley Series in Probability and Mathematical Statistics.

[42] Kazuhiro Ishige. Movement of hot spots on the exterior domain of a ball under the Dirichlet boundary condition. *Adv. Differential Equations*, 12(10):1135–1166, 2007.

[43] Kazuhiro Ishige and Yoshitsugu Kabeya. Hot spots for the heat equation with a rapidly decaying negative potential. *Adv. Differential Equations*, 14(7-8):643–662, 2009.

[44] Kazuhiro Ishige and Yoshitsugu Kabeya. L^p norms of nonnegative Schrödinger heat semigroup and the large time behavior of hot spots. *J. Funct. Anal.*, 262(6):2695–2733, 2012.

[45] Kiyosi Itô and Henry P. McKean, Jr. *Diffusion processes and their sample paths*. Springer-Verlag, Berlin-New York, 1974. Second printing, corrected, Die Grundlehren der mathematischen Wissenschaften, Band 125.

[46] Ioan M. James. *The topology of Stiefel manifolds*. Cambridge University Press, Cambridge-New York-Melbourne, 1976. London Mathematical Society Lecture Note Series, No. 24.

[47] Paul A. Jensen and Jonathan F. Bard. *Operations Research Models and Methods*. Wiley, 2003.

[48] David Jerison and Nikolai Nadirashvili. The "hot spots" conjecture for domains with two axes of symmetry. *J. Amer. Math. Soc.*, 13(4):741–772, 2000.

[49] Bernhard Kawohl. *Rearrangements and convexity of level sets in PDE*, volume 1150 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1985.

[50] Wilfrid S. Kendall. Nonnegative Ricci curvature and the Brownian coupling property. *Stochastics*, 19(1-2):111–129, 1986.

[51] Wilfrid S. Kendall. Coupled Brownian motions and partial domain monotonicity for the Neumann heat kernel. *J. Funct. Anal.*, 86(2):226–236, 1989.

[52] Wilfrid S. Kendall. Brownian couplings, convexity, and shy-ness. *Electron. Commun. Probab.*, 14:66–80, 2009.

[53] Richard Snyder Laugesen and Carlo Morpurgo. Extremals for eigenvalues of Laplacians under conformal mapping. *J. Funct. Anal.*, 155(1):64–108, 1998.

[54] Yong Lin, Linyuan Lu, and Shing-Tung Yau. Ricci curvature of graphs. *Tohoku Math. J. (2)*, 63(4):605–627, 2011.

[55] Torgny Lindvall and Leonard C. G. Rogers. Coupling of multidimensional diffusions by reflection. *Ann. Probab.*, 14(3):860–872, 1986.

[56] John E. Littlewood. *Littlewood's miscellany*. Cambridge University Press, Cambridge, 1986. Edited and with a foreword by Béla Bollobás.

[57] John Lott and Cédric Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math. (2)*, 169(3):903–991, 2009.

[58] Yasuhito Miyamoto. The "hot spots" conjecture for a certain class of planar convex domains. *J. Math. Phys.*, 50(10):103530, 7, 2009.

[59] Yasuhito Miyamoto. A planar convex domain with many isolated "hot spots" on the boundary. *Jpn. J. Ind. Appl. Math.*, 30(1):145–164, 2013.

[60] Nikolai S. Nadirashvili. Multiplicity of eigenvalues of the Neumann problem. *Dokl. Akad. Nauk SSSR*, 286(6):1303–1305, 1986.

[61] Yann Ollivier. Ricci curvature of Markov chains on metric spaces. *J. Funct. Anal.*, 256(3):810–864, 2009.

[62] Gabriel V. Orman, editor. *Monotonicity properties of reflecting Brownian motion*, volume 21 of *Scientific Session on Mathematics and its Applications*. Transilvania University Press, June 2007.

[63] **Mihai N. Pascu**. *Probabilistic approaches to eigenvalue problems.* ProQuest LLC, Ann Arbor, MI, 2001. Thesis (Ph.D.)–University of Connecticut.

[64] **Mihai N. Pascu**. Scaling coupling of reflecting Brownian motions and the hot spots problem. *Trans. Amer. Math. Soc.*, 354(11):4681–4702 (electronic), 2002.

[65] **Mihai N. Pascu**. Mirror coupling of reflecting Brownian motion and an application to Chavel's conjecture. *Electron. J. Probab.*, 16:No. 18, 504–530, 2011.

[66] **Mihai N. Pascu** and Maria E. Gageonea. Monotonicity properties of the Neumann heat kernel in the ball. *J. Funct. Anal.*, 260(2):490–500, 2011.

[67] **Mihai N. Pascu** and Alina Nicolaie. On a discrete version of the Laugesen-Morpurgo conjecture. *Statist. Probab. Lett.*, 79(6):797–806, 2009.

[68] **Mihai N. Pascu** and Nicolae R. Pascu. Brownian motion on the circle and applications. *Bull. Transilv. Univ. Braşov Ser. III*, 1(50):469–478, 2008.

[69] **Mihai N. Pascu** and Nicolae R. Pascu. A closer look at the solutions of a degenerate stochastic differential equation. *Bull. Transilv. Univ. Braşov Ser. III*, 4(53)(1):59–66, 2011.

[70] **Mihai N. Pascu** and Nicolae R. Pascu. Neighbourhoods of univalent functions. *Bull. Aust. Math. Soc.*, 83(2):210–219, 2011.

[71] **Mihai N. Pascu** and Nicolae R. Pascu. Starlike approximations of analytic functions. *Appl. Math. Comput.*, 218(12):6825–6832, 2012.

[72] **Mihai N. Pascu** and Nicolae R. Pascu. Convex approximations of analytic functions. *Appl. Math. Comput.*, 232:559–567, 2014.

[73] **Mihai N. Pascu** and Ionel Popescu. Shy and fixed-distance couplings of brownian motions on manifolds. *To appear (preprint available on Arxiv), 29 pp.*

[74] Nicolae N. Pascu. *Funcţii Univalente*. Editura Universităţii "Transilvania" Braşov, 2002.

[75] Christian Pommerenke. *Univalent functions*. Vandenhoeck & Ruprecht, Göttingen, 1975. With a chapter on quadratic differentials by Gerd Jensen, Studia Mathematica/Mathematische Lehrbücher, Band XXV.

[76] Christian Pommerenke. *Boundary behaviour of conformal maps*, volume 299 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1992.

[77] Huo-Jun Ruan. The "hot spots" conjecture for the Sierpinski gasket. *Nonlinear Anal.*, 75(2):469–476, 2012.

[78] Huo-Jun Ruan and Yong-Wen Zheng. The "hot spots" conjecture on the level-3 Sierpinski gasket. *Nonlinear Anal.*, 81:101–109, 2013.

[79] Daniel W. Stroock. On the growth of stochastic integrals. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 18:340–344, 1971.

[80] Daniel W. Stroock. *Probability theory, an analytic view*. Cambridge University Press, Cambridge, 1993.

[81] Daniel W. Stroock. *An introduction to the analysis of paths on a Riemannian manifold*, volume 74 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2000.

[82] Karl-Theodor Sturm. On the geometry of metric measure spaces. I. *Acta Math.*, 196(1):65–131, 2006.

[83] Karl-Theodor Sturm. On the geometry of metric measure spaces. II. *Acta Math.*, 196(1):133–177, 2006.

[84] Max K. von Renesse and Karl-Theodor Sturm. Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.*, 58(7):923–940, 2005.

[85] Feng Yu Wang. Application of coupling methods to the Neumann eigenvalue problem. *Probab. Theory Related Fields*, 98(3):299–306, 1994.

[86] Peng-fei Yang. The hot spots conjecture on a class of domains in R^n with $n \geqslant 3$. *Acta Math. Appl. Sin. Engl. Ser.*, 27(4):639–646, 2011.