REPORTS@SCM

AN ELECTRONIC JOURNAL OF THE SOCIETAT CATALANA DE MATEMÀTIQUES

Volume 3, num. 1, 2017

ISSN (electronic edition): 2385 - 4227

http://reportsascm.iec.cat

Institut d'Estudis Catalans

AN ELECTRONIC JOURNAL OF THE
SOCIETAT CATALANA DE MATEMÀTIQUES

Volume 3, number 1

December 2017

http://reportsascm.iec.cat
ISSN electronic edition: 2385 - 4227

Societat Catalana de Matemàtiques

Institut d'Estudis Catalans

Focus and Scope

Reports@SCM is a non-profit electronic research journal on Mathematics published by the Societat Catalana de Matemàtiques (SCM) which originated from the desire of helping students and young researchers in their first steps into the world of research publication.

Reports@SCM publishes short papers (maximum 10 pages) in all areas of pure mathematics, applied mathematics, and mathematical statistics, including also mathematical physics, theoretical computer science, and any application to science or technology where mathematics plays a central role. To be considered for publication in Reports@SCM an article must be written in English (with an abstract in Catalan), be mathematically correct, and contain some original interesting contribution. All submissions will follow a peer review process before being accepted for publication.

Research announcements containing preliminary results of a larger project are also welcome. In this case, authors are free to publish in the future any extended version of the paper elsewhere, with the only condition of making an appropriate citation to Reports@SCM.

We especially welcome contributions from researchers at the initial period of their academic careers, such as Master or PhD students. We wish to give special attention to the authors during the whole editorial process. We shall take special care of maintaining a reasonably short average time between the reception of a paper and its acceptance, and between its acceptance and its publication.

All manuscripts submitted will be **peer reviewed** by at least one reviewer. Final decisions on the acceptance of manuscripts are taken by the editorial board, based on the reviewer's opinion.

Edited by Societat Catalana de Matemàtiques, Institut d'Estudis Catalans (IEC)
Carrer del Carme 47, 08001 Barcelona.
http://scm.iec.cat Telèfon: (+34) 93 324 85 83
scm@iec.cat Fax: (+34) 93 270 11 80

Style revision by Enric Ventura.

Institut d'Estudis Catalans
http://www.iec.cat
informacio@iec.cat

http://reportsascm.iec.cat
ISSN electronic edition: 2385-4227

Reports@SCM
Volume 3, number 1, 2017

Table of Contents

The wave equation for stiff strings and piano tuning
Xavier Gracia and Tomas Sanz-Perela 1

Explicit bounds for growth of sets in non-abelian groups
Alberto Espuny Díaz 17

Compactification of a diagonal action on the product of CAT (-1) spaces
Teresa García 27

REPORTS@SCM
AN ELECTRONIC JOURNAL OF THE
SOCIETAT CATALANA DE MATEMÀTIQUES

The wave equation for stiff strings and piano tuning

***Xavier Gràcia**
Universitat Politècnica de Catalunya
xavier.gracia@upc.edu

Tomás Sanz-Perela
Universitat Politècnica de Catalunya
tomassanzperela@gmail.com

∗Corresponding author

Resum *(CAT)*
Estudiem l'equació de les ones per a una corda amb rigidesa. Resolem l'equació i n'enunciem un teorema d'unicitat amb condicions de contorn adequades. Per a una corda punxada calculem l'espectre, que és lleugerament inharmònic. Per tant, l'habitual escala de 12 divisions iguals de l'octava justa no és la millor elecció per afinar instruments com ara el piano. Basant-nos en la teoria de la dissonància, proporcionem una manera d'afinar el piano a fi de millorar-ne la consonància. Una bona solució s'obté afinant una nota i la seva quinta tot minimitzant els seus batecs.

Abstract *(ENG)*
We study the wave equation for a string with stiffness. We solve the equation and provide a uniqueness theorem with suitable boundary conditions. For a pinned string we compute the spectrum, which is slightly inharmonic. Therefore, the widespread scale of 12 equal divisions of the just octave is not the best choice to tune instruments like the piano. Basing on the theory of dissonance, we provide a way to tune the piano in order to improve its consonance. A good solution is obtained by tuning a note and its fifth by minimizing their beats.

Keywords: *wave equation, vibrating string, stiffness, inharmonic spectrum, musical scale, dissonance.*
MSC (2010): *00A65, 35G16, 35L05, 74K05.*
Received: *February 4th, 2016.*
Accepted: *June 23th, 2016.*

Acknowledgement
X.G. acknowledges partial financial support by Generalitat de Catalunya project 2014–SGR–634. The authors wish to thank the anonymous referee whose comments have greatly improved the paper.

Societat Catalana de Matemàtiques
Institut d'Estudis Catalans

1. Introduction

The problem of finding appropriate scales for playing music has worried musical theorists and instrument makers for centuries. There is a close relationship between the theory of musical scales and the frequency spectra of musical instruments; indeed, the harmonic spectrum of most instruments has lead to the present day tempered scale, with 12 *equal divisions of the octave* (12-edo). However, piano strings have some degree of stiffness, which implies that their spectrum is slightly inharmonic, and this explains why the tuning of the piano is actually "stretched", with octaves slightly larger than should; see [6, p. 389]. The purpose of this paper is to perform an accurate mathematical study of the wave equation of a string with stiffness, and what does it imply to the choice of a scale. Throughout the paper we will assume an elementary knowledge of acoustics (physical and perceptive properties of the sound) and of music theory (intervals and the tempered scale); for the benefit of the reader, we have collected some of these notions in an appendix.

The classical wave equation, $u_{tt} = c^2 u_{xx}$, models a perfectly elastic string. If we want to take the stiffness into account, we need to modify the equation. The simplest way to do this consists in adding a term coming from the Euler–Bernoulli beam equation, which is used to model the deflection of rigid bars. The result is a fourth order PDE of the form $u_{tt} = c^2 u_{xx} - M^2 u_{xxxx}$ that has been seldom studied in the acoustics literature [6, 10], and often sketchily. This is why we have found it convenient to perform a more detailed and self-contained study with rather elementary techniques; see Section 2. We compute the explicit form of the solutions, which turns out to be the same as in the non-stiff case, except by the fact that the frequency spectrum is no longer harmonic, but of the form $f_n = n\, f_\circ \sqrt{1 + Bn^2}$, with $n \geq 1$, where B is a constant depending on the physical parameters of the string. We also show the existence and uniqueness for the PDE with appropriate boundary conditions.

For piano strings the value of the inharmonicity parameter B is about 10^{-3}. This means that its spectral frequencies slightly deviate from the harmonic ones (it has a "stretched harmonic spectrum"). Though small, this deviation is of great importance for the consonance of the intervals between notes, because the human ear is very sensitive to frequency differences.

The auditory perception qualifies some musical sounds as *consonant* ("pleasant"), whereas others are *dissonant*. As it is explained in more detail in Section 3 and in the Appendix, there is a close relationship between dissonance, spectrum and scale: the choice of the notes used to play music aims to achieve the best possible consonance, and this consonance depends directly on the spectrum of the sounds. Therefore, the fact that stiff strings have a slightly inharmonic spectrum leads to reconsider the exact tuning of the notes we play with them. A tool to perform a systematic study of this problem is the *dissonance curve* of a spectrum. Basing on experimental results by Plomp and Levelt [11], one can define a function to measure the dissonance of two notes as a function of their frequency ratio, and draw a dissonance curve, which depends strongly on the spectrum. The local minima of this curve indicate possible good choices for the notes of the scale as far as the consonance of its intervals is concerned, [15].

We apply this approach to the string with stiffness. Its spectrum is given by $f_n = n\, f_\circ \sqrt{1 + Bn^2}$ (see Section 2) and therefore the ratio between the first and the second partials is not the just octave 2:1, but a "stretched octave" $2\sqrt{1+4B}/\sqrt{1+B}$. So, we wonder if there exist scales for this spectrum that could possibly be more "consonant" than the usual 12-edo scale. Aiming to preserve the freedom to modulate to any tonality, we look for a scale with equal steps, or, equivalently, equal divisions of a certain interval, as for instance a stretched octave. In Section 4 we use the dissonance curve of the stretched spectrum to study this problem in two different ways. One is based in the coincidence of some partials. The other one minimizes a weighted mean of the dissonance. As a result, we obtain that a good solution is to tune the fifth by making the second partial of the higher note to coincide with the third partial of the fundamental note.

The paper is organised as follows. In Section 2 we study the modelling of a string with stiffness: we give an explicit solution of the equation when the boundary conditions are those of a pinned string, we present a rigorous derivation of its spectrum, and we state a uniqueness theorem. In Section 3 we recall some facts about the theory of dissonance and how to draw dissonance curves, and we obtain the dissonance curve of the string with stiffness. In Section 4 we study several proposals to tune the piano, either based in the coincidence of partials or the minimization of the mean dissonance. Section 5 is devoted to conclusions. Finally, an appendix gathers some basic concepts of acoustics and music theory.

2. The wave equation for the string with stiffness

It is well known [14, 16] that the motion of a vibrating string (for instance, a violin string) can be represented by the solutions of the problem

$$\begin{cases} \dfrac{\partial^2 u}{\partial t^2} = c^2 \dfrac{\partial^2 u}{\partial x^2} & x \in (0, L),\ t > 0 \\ u(0,t) = u(L,t) = 0 & t \geq 0 \\ u(x,0) = \phi(x) & x \in [0, L] \\ \partial_t u(x,0) = \psi(x) & x \in [0, L], \end{cases}$$

where $u(x,t)$ represents the transversal displacement of the string of length L (represented by the interval $[0, L]$) from its equilibrium position, and $\phi(x)$ and $\psi(x)$ are, respectively, the initial shape and velocity of the string. The boundary conditions $u(0,t) = u(L,t) = 0$ for $t \geq 0$ mean that the string has fixed ends and $c^2 = \tau/\rho$, with τ the tension of the string and ρ its linear density. The value c is the velocity of the travelling waves along the string.

The solution of this equation can be computed using the method of separation of variables, obtaining

$$u(x,t) = \sum_{n=1}^{\infty} \left[a_n \cos(2\pi f_n t) + b_n \sin(2\pi f_n t)\right] \sin\left(\frac{n\pi}{L} x\right), \quad f_n = \frac{nc}{2L},$$

where the coefficients a_n and b_n are obtained from the Fourier coefficients of the initial conditions ϕ, ψ. For the convergence and smoothness of this series some regularity conditions are required on ϕ, ψ; see, for instance, [14].

The model of the wave equation is a good approximation for instruments like the guitar, whose strings are almost perfectly flexible. However, when we want to model the motion of the piano strings, which have greater stiffness, the classical wave equation is not good enough. For this reason, a term describing the resistance against bending is added to it (see [6, p. 64]), obtaining the following equation:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} - \frac{ESK^2}{\rho} \frac{\partial^4 u}{\partial x^4}, \tag{1}$$

where S is the cross-sectional area of the string, E is Young's modulus of its material, ρ is its linear density and K is the radius of gyration, which is $K = R/2$ for a cylindrical shape of radius R; see [6, p. 58].

The added term is the same that appears in the beam equation (also called Euler–Bernoulli equation), modelling the motion of a vibrating beam under the hypotheses of no shear stress nor rotational inertia; a deduction of this equation can be found, for instance, in [6, p. 58], or at the end of [17, Chapter 2]. One can view (1) as the generalization of a PDE for a vibrating material: the first term is due to the elasticity of the material (its capacity to return to the initial position after a deformation) and the second one, due to the resistance against bending. If the first term is zero, the material is not elastic and we get the beam equation. On the contrary, if the second term is zero, the material is not rigid and we get the wave equation.

The string with stiffness based on the Euler–Bernouilli model is the most widely used model. Nevertheless, there exist other equations that can model the vibration of rigid materials (and in particular piano strings). Prominent among them is the Timoshenko beam model, which takes into account shear stress and rotational inertia [7]. The description of the motion of piano strings using this model has been thoroughly discussed recently in the thesis [3]. It is shown there that the frequencies of the string based upon the Timoshenko beam model behave as the ones based on the Euler–Bernouilli model for the lower partials; the Timoshenko model provides a better description for higher partials, a region where their contribution to dissonance is negligible. Therefore, the Euler–Bernouilli model is enough for our purposes.

2.1 Solving the equation

Equation (1) was studied in [5], where the author guesses the form of some solutions with separate variables. Besides that article, only a few references in the acoustics literature deal with the string with stiffness, and they merely give approximate solutions of the spectrum, without further justification. A recent study of this equation is in [3], where the exact formula for the frequency of the partials is found using Fourier transform, though the equation is not actually solved. So we have found it convenient to perform a detailed study: by following the standard method of separation of variables, we give a solution of the initial value problem with appropriate boundary conditions, obtaining also the formula for the frequencies. Uniqueness of the solution is studied in the following section.

We start by looking for a solution to equation (1) of the form $u(x,t) = X(x)T(t)$. We have

$$XT'' = c^2X''T - \frac{ESK^2}{\rho}X^{(4)}T \quad \Longrightarrow \quad \frac{T''}{T} = c^2\frac{X''}{X} - \frac{ESK^2}{\rho}\frac{X^{(4)}}{X}\,. \tag{2}$$

As the left-hand side of the equation depends only on t and the right-hand side, only on x, (2) has to be a non positive constant, called $-\omega^2$ (non positive because we are looking for periodic solutions in time):

$$\frac{T''}{T} = c^2\frac{X''}{X} - \frac{ESK^2}{\rho}\frac{X^{(4)}}{X} = -\omega^2.$$

If we look at the time equation, we have an ODE which is easy to solve: $T_\omega(t) = A\cos\omega t + B\sin\omega t$.

We look now for the solutions of the ODE for X:

$$\frac{ESK^2}{\rho}X^{(4)} - c^2X'' - \omega^2X = 0\,. \tag{3}$$

We divide the equation by ESK^2/ρ and define $a := c^2\rho/ESK^2$ and $b := \rho\omega^2/ESK^2$. After that, (3) becomes $X^{(4)} - aX'' - bX = 0$, whose solutions are of the form

$$C_1\cosh k_1x + C_2\sinh k_1x + C_3\cos k_2x + C_4\sin k_2x\,,$$

with $k_1 = \sqrt{(a+\sqrt{a^2+4b})/2}$ and $k_2 = \sqrt{(-a+\sqrt{a^2+4b})/2}$. We introduce again two convenient constants:

$$B := \pi^2 \frac{ESK^2}{\tau L^2} \quad \text{and} \quad f_\circ := \frac{c}{2L}.$$

In this way, using the definition of a and b, we obtain the following relations between k_1, k_2 and ω:

$$k_1^2 = \frac{\pi^2}{2BL^2}\left[\sqrt{1+\frac{\omega^2 B}{f_\circ^2\pi^2}}+1\right] \quad \text{and} \quad k_2^2 = \frac{\pi^2}{2BL^2}\left[\sqrt{1+\frac{\omega^2 B}{f_\circ^2\pi^2}}-1\right]. \tag{4}$$

We want to find the possible values of k_1 and k_2, that will determine the possible values of ω. In order to do it, we will impose the boundary conditions, but now, as the equation is of 4-th order, we need 4 boundary conditions, two more apart from the Dirichlet boundary conditions on both ends of the string. We will consider two cases:

- $X' = 0$ at the ends; this case appears when the string is clamped at the ends.
- $X'' = 0$ at the ends; this happens when the string is pinned at the ends, since there is no moment.

The first case, $X = X' = 0$ at the ends of the string, leads to an equation that can be solved numerically, but it is not possible to get a closed formula for the spectrum of frequencies; see [5] for more details and for an approximate formula. The second case, $X = X'' = 0$ at the ends of the string, is easier to solve and will lead us to a formula for the frequencies of the partials. In the case of the piano, this second option seems to be closer to reality, because the strings are supported on a bridge. From now on, we will focus in this case.

Pinned boundary conditions. We are going to solve the problem with the condition $X = X'' = 0$ at the ends of the string. Consider a general solution of (3), say

$$X(x) = C_1 \cosh k_1 x + C_2 \sinh k_1 x + C_3 \cos k_2 x + C_4 \sin k_2 x. \tag{5}$$

We want to find the possible values of k_1 and k_2 that make (5) satisfy (non-trivially) the boundary conditions. Let us impose these boundary conditions at the string ends, $x = 0, L$.

For $x = 0$, we obtain:

$$X(0) = C_1 + C_3 = 0 \qquad \text{and} \qquad X''(0) = C_1 k_1^2 - C_3 k_2^2 = 0\,.$$

From the first equation we get $-C_3 = C_1$ and, replacing it in the second one, we arrive to the equation $C_1(k_1^2 + k_2^2) = 0$, which implies $C_1 = C_3 = 0$.

Now we impose the boundary conditions at $x = L$ to $X(x) = C_2 \sinh k_1 x + C_4 \sin k_2 x$, obtaining:

$$X(L) = C_2 \sinh k_1 L + C_4 \sin k_2 L = 0 \qquad \text{and} \qquad X''(L) = C_2 k_1^2 \sinh k_1 L - C_4 k_2^2 \sin k_2 L = 0.$$

Multiplying the first equation by k_2^2 and adding it to the second one, we get $C_2(k_1^2 + k_2^2)\sinh k_1 L = 0$. As the last two factors are different from zero, we conclude that $C_2 = 0$.

Finally, we have $C_4 \sin k_2 L = 0$. As we want nontrivial solutions, we need $C_4 \neq 0$ and, thus, $k_2 L = n\pi$ for $n \geq 1$. From this relation and (4) we obtain

$$\left(\frac{n\pi}{L}\right)^2 = \frac{\pi^2}{2BL^2}\left[\sqrt{1+\frac{\omega_n^2 B}{f_\circ^2\pi^2}}-1\right]$$

and, isolating ω_n, we obtain the possible frequencies:

$$f_n = \frac{\omega_n}{2\pi} = n\, f_\circ \sqrt{1 + Bn^2} \quad \text{with} \quad n = 1, 2, \ldots \tag{6}$$

Thus, for each of the ω_n, the solution of (3) satisfying the boundary conditions is a multiple of

$$X_n(x) = \sin\left(\frac{n\pi}{L}x\right) \quad \text{with} \quad n = 1, 2, \ldots$$

Remarkably, these are the same modes of vibration as in the case without stiffness: the difference only shows up in the frequencies of vibration.

To conclude, we can write the general solution of the PDE (1), with boundary conditions $u = 0$ and $u_{xx} = 0$ at the ends of the string, as:

$$u(x,t) = \sum_{n=1}^{\infty} \Big[a_n \cos(2\pi f_n t) + b_n \sin(2\pi f_n t)\Big] \sin\left(\frac{n\pi}{L}x\right), \qquad f_n = n\, f_\circ \sqrt{1 + Bn^2},$$

where a_n, b_n are obtained from initial conditions in the same way as in the ideal case.

As we can see, the spectrum is no longer harmonic, but it is 'stretched' from the harmonic one due to the factor $\sqrt{1 + Bn^2}$. For a cylindrical string of radius R the value of B is $B = \pi^3 E R^4 / 4\tau L^2$; its typical values for a piano string are about 10^{-3}.

Notice that the constant $f_\circ = c/2L$ would be the fundamental frequency of the string if it did not have stiffnes ($B = 0$); in this case, we would recover the frequency spectrum of the ideal string. When $B > 0$, the fundamental frequency is $f_1 = f_\circ \sqrt{1 + B}$, higher than $f_\circ$.

2.2 Uniqueness of solutions

We prove now a theorem of uniqueness of solutions for the wave equation with stiffness. This can be seen a particular case of the results of semigroup theory for evolution problems with monotone operators (see [2, 4]), but in this case we provide an elementary proof, similar to the uniqueness theorem for the wave equation, which can be found for instance in [14]. We will use the notation $\partial_\xi^n := \partial^n / \partial \xi^n$ when necessary.

Lemma 2.1. *Let* $u(x,t) \in \mathcal{C}^4([0,L] \times [0,\infty))$ *satisfying*

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} - M^2 \frac{\partial^4 u}{\partial x^4}, \tag{7}$$

for $x \in (0,L)$ *and* $t > 0$*. In any of the two following cases*

$$\begin{cases} u(0,t) = u(L,t) = 0 & t \geq 0 \\ \partial_x u(0,t) = \partial_x u(L,t) = 0 & t \geq 0 \end{cases} \quad \text{or} \quad \begin{cases} u(0,t) = u(L,t) = 0 & t \geq 0 \\ \partial_x^2 u(0,t) = \partial_x^2 u(L,t) = 0 & t \geq 0 \end{cases},$$

the quantity

$$\mathcal{E}(u) = \frac{1}{2} \int_0^L \left((\partial_t u)^2 + c^2 (\partial_x u)^2 + M^2 (\partial_x^2 u)^2 \right) \mathrm{d}x$$

is constant in time.

Proof. We just need to show that the derivative with respect to t of $\mathcal{E}(u)$ is zero:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{E}(u)) &= \int_0^L \Big(\partial_t u\,\partial_t^2 u + c^2\partial_x u\,\partial_t\partial_x u + M^2\partial_x^2 u\,\partial_t\partial_x^2 u\Big)\mathrm{d}x \qquad \text{(using (7))}\\
&= \int_0^L \Big(\partial_t u(c^2\partial_x^2 u - M^2\partial_x^4 u) + c^2\partial_x u\,\partial_t\partial_x u + M^2\partial_x^2 u\,\partial_t\partial_x^2 u\Big)\mathrm{d}x\\
&= c^2\int_0^L \Big(\partial_t u\,\partial_x^2 u + \partial_x u\,\partial_x\partial_t u\Big)\mathrm{d}x + M^2\int_0^L \Big(\partial_x^2 u\,\partial_x^2\partial_t u - \partial_t u\,\partial_x^4 u\Big)\mathrm{d}x =: c^2 I_c + M^2 I_S\,.
\end{aligned}$$

Now,

$$I_c = \int_0^L \Big(\partial_t u\,\partial_x^2 u + \partial_x u\,\partial_x\partial_t u\Big)\mathrm{d}x \overset{\text{parts}}{=} \int_0^L \Big(\partial_t u\,\partial_x^2 u - \partial_x^2 u\,\partial_t u\Big)\mathrm{d}x + \big[\partial_x u\,\partial_t u\big]_0^L = \big[\partial_x u\,\partial_t u\big]_0^L,$$

which is zero due to the Dirichlet boundary condition ($u = 0$ at $0, L$ for all $t \geq 0$ implies $\partial_t u = 0$ for all $t \geq 0$). Similarly, we have

$$\begin{aligned}
I_S &= \int_0^L \Big(\partial_x^2 u\,\partial_x^2\partial_t u - \partial_t u\,\partial_x^4 u\Big)\mathrm{d}x \overset{\text{parts}}{=} \int_0^L \Big(-\partial_x^3 u\,\partial_x\partial_t u - \partial_t u\,\partial_x^4 u\Big)\mathrm{d}x + \big[\partial_x^2 u\,\partial_x\partial_t u\big]_0^L\\
&\overset{\text{parts}}{=} \int_0^L \Big(\partial_x^4 u\,\partial_t u - \partial_t u\,\partial_x^4 u\Big)\mathrm{d}x + \big[\partial_x^2 u\,\partial_x\partial_t u\big]_0^L - \big[\partial_x^3 u\,\partial_t u\big]_0^L = \big[\partial_x^2 u\,\partial_x\partial_t u\big]_0^L - \big[\partial_x^3 u\,\partial_t u\big]_0^L,
\end{aligned}$$

which is again zero due to the boundary conditions.

Therefore, we get $\mathrm{d}(\mathcal{E}(u))/\mathrm{d}t = 0$. □

Thanks to this result, we can now prove the next theorem. Indeed, in its formulation we include a source term and nonhomogeneous boundary conditions, so it is slightly more general than the problem of the stiff string we are studying.

Theorem 2.2. *There exists at most one solution $u \in \mathcal{C}^4([0, L] \times [0, \infty))$ of the problem*

$$\begin{cases}
\dfrac{\partial^2 u}{\partial t^2} = c^2\dfrac{\partial^2 u}{\partial x^2} - M^2\dfrac{\partial^4 u}{\partial x^4} + f(x,t) & x \in (0, L),\ t > 0\\
u(0,t) = u(L,t) = \delta(t) & t \geq 0\\
\partial_x^2 u(0,t) = \partial_x^2 u(L,t) = \mu(t) & t \geq 0\\
u(x,0) = \phi(x) & x \in [0, L]\\
\partial_t u(x,0) = \psi(x) & x \in [0, L]\,.
\end{cases} \tag{8}$$

This is also true if, instead of conditions on $\partial_x^2 u$, we put conditions on $\partial_x u$: $\partial_x u(0,t) = \partial_x u(L,t) = \eta(t)$.

Proof. Let u_1 and u_2 be two solutions of the problem (8). Since the PDE is linear, $u := u_1 - u_2$ solves the homogeneous problem

$$\begin{cases}
\dfrac{\partial^2 u}{\partial t^2} = c^2\dfrac{\partial^2 u}{\partial x^2} - M^2\dfrac{\partial^4 u}{\partial x^4} & x \in (0, L),\ t > 0\\
u(0,t) = u(L,t) = 0 & t \geq 0\\
\partial_x^2 u(0,t) = \partial_x^2 u(L,t) = 0 & t \geq 0\\
u(x,0) = 0 & x \in [0, L]\\
\partial_t u(x,0) = 0 & x \in [0, L]\,.
\end{cases} \tag{9}$$

By lemma 2.1, the non negative quantity $\mathcal{E}(u)$ is constant in t. But at time $t = 0$,

$$\mathcal{E}(u)\big|_{t=0} = \frac{1}{2}\int_0^L \Big((\partial_t u)^2 + c^2(\partial_x u)^2 + M^2(\partial_x^2 u)^2\Big)\Big|_{t=0} \mathrm{d}x = 0$$

so, $\mathcal{E}(u) \equiv 0$ for all time. Therefore,

$$(\partial_t u)^2 + c^2(\partial_x u)^2 + M^2(\partial_x^2 u)^2 = 0 \quad \Longrightarrow \quad \partial_t u = 0\,,\ \partial_x u = 0 \ \ (\text{and } \partial_x^2 u = 0)\,.$$

Since all partial derivatives of first order of u are zero, u is a constant function. Finally, at $t = 0$, $u = 0$, so $u \equiv 0$ for all time $t \geq 0$. Therefore, $u_1 = u_2$.

The proof for boundary conditions on $\partial_x u$ is completely analogous. □

3. Scales, spectrum and dissonance curves

As we mentioned in the introduction, if a spectrum is given, by analysing its dissonance curve, one can try to find appropriate scales for it. In this section we provide some details for this analysis, following [15].

In 1965, Plomp–Levelt [11] performed an experiment aiming to measure the dissonance of two pure tones in terms of their distance; this dissonance was evaluated by many subjects, and as a result they concluded that the maximal degree of dissonance is attained at roughly 1/4 of the critical bandwidth, a concept from psychoacoustics that had been introduced and studied some years before (see also [12]). Except for low frequencies, the width of a critical band corresponds to an interval around a minor third.

Plomp and Levelt claimed that this could be extrapolated to complex tones, so that the dissonance of a sound could be computed as the sum of the dissonances of all the pairs of its partials. More specifically, by taking a *harmonic spectrum* with 6 partials (and equal loudnesses), they obtained a dissonance curve similar to Helmholtz's, showing points of local minimal dissonance for frequency ratios α equal to 1:1, 2:1, 3:2, 5:3, 4:3, 6:5, 5:4, and a maximum of dissonance near the semitone interval. Figure 1 shows some of Plomp and Levelt's results.

More recently, Sethares [15] did a systematic study of the dissonance curve of several spectra, and showed the close relationship between spectrum and scales; his work includes the synthesis of artificial spectra adapted to play music in exotic scales, while still retaining some degree of consonance.

We want to apply this procedure to the spectrum of the string with stiffness. For this, we need a specific expression of a function modelling the dissonance. Following [15], given two pure tones of frequencies $f_1 \leq f_2$ (expressed in Hz) and loudnesses ℓ_1, ℓ_2 then the dissonance (in an arbitrary scale) can be expressed as $d(f_1, f_2, \ell_1, \ell_2) = \min(\ell_1, \ell_2)\left(e^{-b_1 s (f_2 - f_1)} - e^{-b_2 s (f_2 - f_1)}\right)$, where $s = x^*/(s_1 f_1 + s_2)$, and the parameters are $b_1 = 3.5$, $b_2 = 5.7$, $x^* = 0.24$, $s_1 = 0.021$ and $s_2 = 19$. The graph of this function reproduces the shape obtained by Plomp and Levelt, Figure 1 (left); dissonance is measured in an arbitrary scale, therefore usually we will normalize its expression so that it takes values between 0 and 1. The preceding expressions and numbers give just a possible model for the dissonance of two tones; other models (see for instance [1]) can be used and, qualitatively, the results are the same.

Then, if $\mathcal{F}$ is a spectrum with frequencies $f_1 < \cdots < f_n$ and loudnesses $\ell_1, \ldots, \ell_n$, the dissonance of $\mathcal{F}$ is defined as the sum of the dissonances of all the pairs of partials, $d_{\mathcal{F}} = \sum_{i<j} d(f_i, f_j, \ell_i, \ell_j)$. Finally, the dissonance function of a given spectrum $\mathcal{F}$ is the function that yields the dissonance of two tones as a

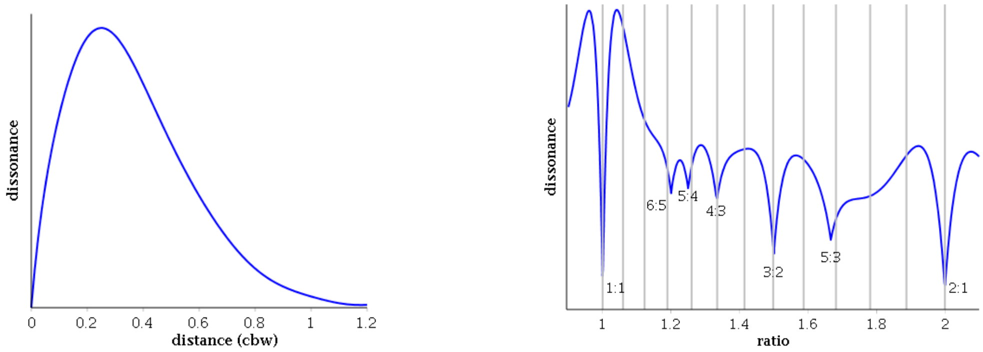

Figure 1: Graphics of Plomp and Levelt's results [11]. *Left:* Dissonance of two pure tones as a function of their distance measured in critical bandwidths; extrapolated from experimental data. *Right:* Theoretical model of the dissonance of two harmonic tones as a function of their frequency ratio (the vertical lines mark the steps of the 12-edo scale).

function of the ratio α of their fundamental frequencies: $D_{\mathcal{F}}(\alpha) = d_{\mathcal{F}\cup\alpha\mathcal{F}}$, where we denote by $\alpha\mathcal{F}$ the spectrum $\mathcal{F}$ with its frequencies scaled by the factor α, and by $\mathcal{F}\cup\alpha\mathcal{F}$ the union of both spectra.

The graph of the function $D_{\mathcal{F}}$ is the *dissonance curve* of the given spectrum, and its analysis can help us to find an appropriate scale (and conversely: given an arbitrary scale, is there an appropriate spectrum for it?). Nevertheless, this is not so immediate, and these results do not tell us how to construct a scale. For instance, consider the harmonic spectrum and its dissonance curve as in Figure 1 (right). From a reference note —a C, say— one can form a just scale by adding other notes coinciding with the local minima of the dissonance curve: G (3:2), F (4:3), A (5:3), E (5:4), etc. Notice, however, that from each of these new notes we should consider again the dissonance curve in relation with the notes already chosen. This analysis is simpler when we use an equal-step scale, like 12-edo, because the relative positions of the notes are the same. In the same figure we see the abscissas of the 12-edo scale; it is clear that local minima of dissonance are attained near points that are at a distance of 3, 4, 5, 7, 9 and 12 steps from any given note.

Now let us apply this procedure to the piano. As we have already noted in the preceding sections, its strings have a certain degree of stiffness, and, according to (6), their spectrum is given by $f_n = n\,f_\circ\sqrt{1+Bn^2}$, for $n \geq 1$. We can draw its dissonance curve and we observe that, for small $B > 0$, the local minima of dissonance are slightly shifted to the right with respect to those of the harmonic spectrum; see Figure 2.

Notice, in particular, that the octave and the fifth (the most important intervals of Western music) of the usual 12-edo scale are noticiably flatter than the "optimal" octave and fifth deduced from the stretched spectrum, i.e., the corresponding intervals where this spectrum has a local minimum of dissonance. Therefore the 12-edo scale seems not to be the best choice to play music as far as dissonance is concerned. This fact makes us wonder which is the "best" tuning for the piano, i.e., a tuning that fits better with the minima of the dissonance curve. We give an answer to this question in the next section.

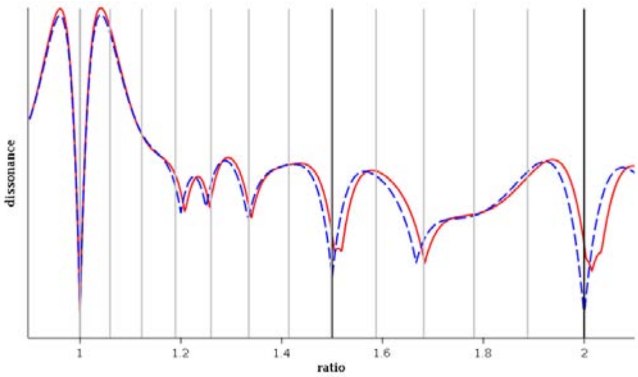

Figure 2: Comparison betwwen the dissonance curves of the harmonic spectrum (dashed) and the stretched spectrum (solid) of a string with stiffness. The grey vertical lines show the steps of the usual 12-edo scale; the black ones show the just fifth and octave. (We have used $B = 0.0013$ for the sake of clarity.)

4. Proposals for the piano tuning

We have just seen that, due to the stiffness of the strings, the spectrum of the piano is slightly stretched, and therefore the minima of the dissonance curve do not coincide with the notes of the usual 12-edo scale, not to say other tunings like just intonation. Now our goal is to find a scale that preserves, as much as possible, the consonance of the main intervals of music. We will restrict our search to scales with equal steps, because we want to preserve the freedom to modulate to arbitrary tonalities —this is especially important for piano music. So, if r is the frequency ratio of the step of the scale, and f is the frequency of its fundamental note, the frequencies of all the notes are $f, r\,f, r^2 f, r^3 f, \ldots$

We will follow two procedures. The first one is based on the *coincidence of a couple of partials*: then their beats disappear and we avoid their dissonance, as it is explained in the Appendix. We will explore three possible choices for the step and see what do they imply for the dissonance curve. The second one is to define an *average dissonance* as a function of the step and try to minimize it.

It should be remarked that in this study we assume that the stiffness parameter B is the same for all the strings. This is approximately true in the middle third of the keyboard [5]. For the lower third of the keyboard, the stiffness parameter of the strings is very low, indeed they are manufactured in a special way, so that a different analysis would be required; besides this, the overall dissonance in this region is high. For the upper third of the keyboard, the upper partials are weak (and even become rapidly inaudible), so that their effect on the dissonance can be neglected.

In all calculations we will use the obtained formula for the partials, $f_n = n\, f_\circ \sqrt{1+Bn^2}$ $(n \geq 1)$, as well as the expresions of the dissonance functions defined in Section 3. We will consider $B \in [0.0004, 0.002]$, see [5], but we will also see that we recover the results for the harmonic case when $B \to 0$.

4.1 Coincidence of partials

Here, our strategy to construct a scale *close to 12-edo* is as follows:

(i) We consider a fixed note of fundamental frequency f_1. Suppose we have already fixed a second note,

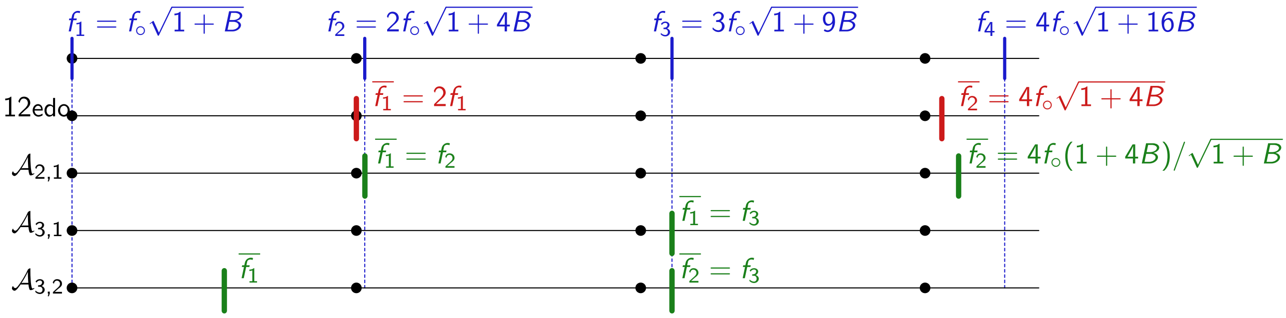

Figure 3: The partials of the stretched spectrum are represented in the upper line; they are compared with the harmonic ones (black nodes). In the lower lines the partials of four tunings are shown: the usual just octave (12-edo), the octave of $\mathcal{A}_{2,1}$, the twelfth of $\mathcal{A}_{3,1}$, and the fifth of $\mathcal{A}_{3,2}$.

of fundamental frequency $\overline{f}_1$. Then we divide the interval $\overline{f}_1 : f_1$ in p equal parts, thus obtaining a step whose frequency ratio is $r = (\overline{f}_1/f_1)^{1/p}$. We will choose the number of parts p that makes the step r to be the closest possible to the frequency ratio of the 12-edo semitone, $2^{1/12}$.

(ii) So we have to properly choose the second note $\overline{f}_1$. We base this choice upon the spectrum of the notes, $(f_i)_{i\geq 1}$ and $(\overline{f}_i)_{i\geq 1}$. In our particular case, we seek the coincidence of some partials. So, we define the tuning $\mathcal{A}_{m,n}$ as the one obtained by letting the m-th partial f_m of the first note to coincide with the n-th partial $\overline{f}_n$ of the second note. Equating $f_m = \overline{f}_n$ determines $\overline{f}_n$ and therefore $\overline{f}_1$.

(iii) Finally, if $r = r_{m,n}$ is the step of $\mathcal{A}_{m,n}$, the notes of the scale are $r^k f_1$, for integer values of k.

We have noticed before that the coincidence of some partials does not necessarily imply consonance. However, this analysis is meaningful because it can be directly applied to actual tuning, since it is easy to tune an interval by letting beats disappear; moreover, we will see later that one of our proposals will be especially good in terms of dissonance.

As the octave and the fifth are the most important intervals in music, three natural tunings can be considered:

(i) $\mathcal{A}_{2,1}$: the second partial f_2 of the first note coincides with the first partial $\overline{f}_1$ of the second note (we try to minimize the beats of the octave).

(ii) $\mathcal{A}_{3,1}$: the third partial f_3 of the first note coincides with the first partial $\overline{f}_1$ of the second note (we try to minimize the beats of the twelfth).

(iii) $\mathcal{A}_{3,2}$: the third partial f_3 of the first note coincides with the second partial $\overline{f}_2$ of the second note (we try to minimize the beats of the fifth).

In Figure 3 we show a schematic description of these tunings. Once we have tuned our interval $\overline{f}_1 : f_1$, we divide it in p equal parts: 12 for the octave, 19 for the twelfth, and 7 for the fifth. These steps are the semitones of the corresponding tuning. Their frequency ratio is given by $(\overline{f}_1/f_1)^{1/p}$; in our cases, they are:

$$r_{2,1} = 2^{1/12}\left(\frac{1+4B}{1+B}\right)^{1/24}, \quad r_{3,1} = 3^{1/19}\left(\frac{1+9B}{1+B}\right)^{1/38}, \quad \text{and} \quad r_{3,2} = \left(\frac{3}{2}\right)^{1/7}\left(\frac{1+9B}{1+4B}\right)^{1/14}.$$

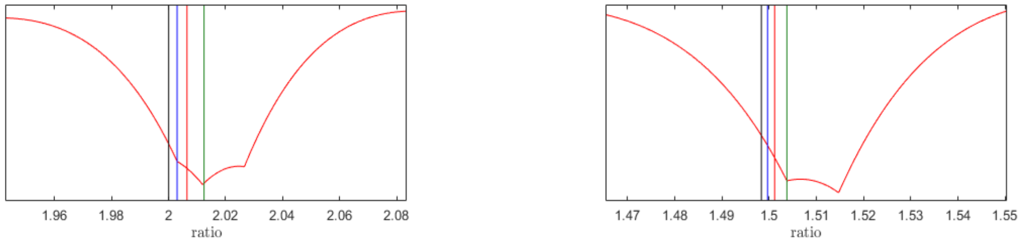

Figure 4: *Left:* The dissonance curve of the stretched spectrum ($B = 0.001$) near the just octave; vertical lines show the octave generated by different steps r: from left to right, $r = 2^{1/12}$ (12-edo), $r = r_{2,1}$, $r = r_{3,1}$, and $r = r_{3,2}$. *Right:* The same near the just fifth.

Now, a given choice of the step r defines a scale, and we can locate its notes in the dissonance curve. The idea is that, by stretching the gap between the notes (vertical lines in Figure 2), possibly the new notes will fit better the minima of the dissonance curve of the stretched spectrum (solid line in the figure).

One of our main goals is to tune the octave, the most important interval in music. Therefore, we analyze in particular how the new octaves generated by these steps (the ratio obtained by r^{12}, for each of the chosen values of r) fit the minimum of the dissonance curve near the frequency ratio 2. The same can be done with the fifth by analyzing the ratios r^7 near the ratio 3:2 in the dissonance curve. The results are shown in Figure 4.

From that figure it appears that the tuning $\mathcal{A}_{3,2}$ fits better than the others the minimum of dissonance at the octave and also at the fifth. For the octave this may seem paradoxical because $\mathcal{A}_{2,1}$ was set to tune the octave ad hoc, but actually this tuning only makes the dissonance caused by a single pair of partials to disappear, whereas other partials may give rise to higher dissonance. This suggests also that our study should consider all the other intervals, because we are not controlling their dissonance. In the next section we make a proposal to deal with this.

4.2 Minimization of dissonance

The preceding analysis can be completed by performing a general study of the dissonance of all the intervals of the scale. We would like to find the semitone r which minimizes (in some sense) the total dissonance of the scale.

In the most general setting, we could define the mean dissonance of a scale as a weighted sum of the dissonances of all couples of notes. The weighting is necessary because not all intervals are equally used in music, and different intervals have different musical roles; therefore their consonances are not equally important.

For an equal step scale it is enough to consider the dissonances of all the notes with respect to a given one, that is, the dissonances between the fundamental note of the scale (with frequency f_1) and the others (with frequencies $r^k f_1$). More specifically, given a semitone r, we define the *mean dissonance* of the equal step scale generated by r as a weighted average of the dissonances of all the intervals from the fundamental note of the scale within the range of an octave, that is:

$$D_{\mathrm{m}}(r) := \sum_{k=1}^{12} w_k \, D_{\mathcal{F}}(r^k),$$

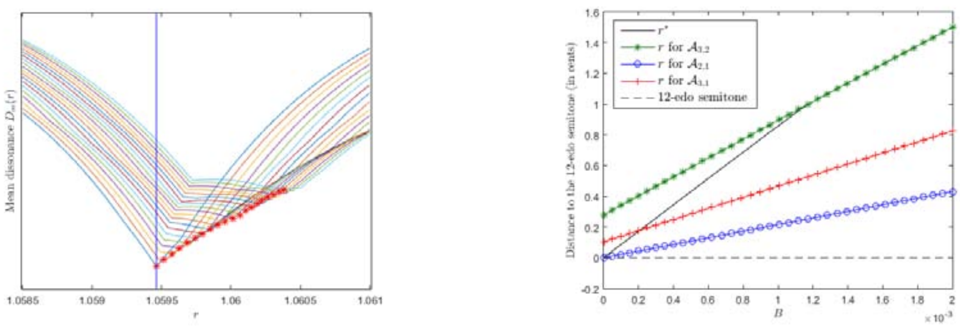

Figure 5: *Left:* Plot of $D_\mathrm{m}(r)$ for different values of $B \in [0, 0.002]$. For each B the marked point corresponds to the minimum r^*. The vertical line represents the 12-edo semitone. *Right:* Distance to the 12-edo semitone of the three tunings and the optimal semitone (in cents).

where $\mathcal{F}$ is the spectrum of the fundamental note f_1, including frequencies and loudnesses, and $w = (w_k)$ is a vector of weights. In the following calculations we have used $w = (1, 1, 4, 4, 5, 2, 6, 4, 4, 2, 1, 10)$ to give preeminence to the octave, the fifth, etc. Intervals larger than an octave could be considered in the sum; we omit them because their effect on the dissonance is small and we have to cut the sum at some point.

If we minimize the function $D_\mathrm{m}(r)$, numerically, on the interval $r \in [1.0585, 1.061]$, near the 12-edo semitone $2^{1/12}$, we find different values for the minimum point r^* depending on B. The results are shown in Figure 5 (left).

We want to compare this optimal semitone with the semitones of the three tuning proposals considered before. For each B, we compute the distance of these four semitones to the 12-edo semitone; this is shown in Figure 5 (right). As we can see, the semitone of $\mathcal{A}_{3,2}$ approximates the optimal semitone: (i) better than the 12-edo semitone if B is higher than 0.00025; (ii) better than the other proposals if B is higher than 0.0005; and (iii) coincides with the optimal semitone if B is higher than 0.001. Notice also that the optimal semitone coincides with the 12-edo semitone when $B = 0$.

The graphics in Figure 5 have been computed with Matlab. We have used a spectrum of 6 partials, the value $f_\circ = 440$ Hz, the weighting vector as defined before, and the loudnesses inversely proportional to the number of the partial. Nevertheless, we have also checked that the results obtained are quite similar if we use the same loudness for all partials, the same weights for all intervals, or 7 partials instead of 6.

5. Conclusions

We have studied the spectrum of strings with stiffness modelled according to the Euler–Bernouilli model. Although this was already known, we have done a mathematically rigorous derivation of it using elementary techniques. We have applied this result and the theory of dissonance to study the tuning of the piano with a scale of equal steps. We have followed two approaches: one is to define a scale based on the coincidence of some specific partials; the other one is to define an average dissonance of a scale and trying to minimize it as a function of the step. It appears that a good solution is to tune a note and its fifth by forcing their 3rd and 2nd partials, respectively, to coincide.

Appendix: sound and music

In this appendix we summarize some basic information about sound and music. This can be found in many books, as for instance [1, 13].

Sound, pitch, spectrum. *Sound* is both an oscillation of the air pressure, and also the auditory sensation it creates. Besides duration, sound has three main perceptive attributes: loudness, pitch and timbre. These are related to physical attributes: amplitude, frequency and spectrum. However, these relations are by no means simple.

Let us consider the *pitch*, a quality that allows sounds to be ordered from lower to higher pitches. A *pure tone* of frequency ν and amplitude A is described by a sinusoid $A\sin(2\pi\nu t)$, and its pitch can be identified with the frequency. A musical sound is usually a superposition of pure tones (the *partials*) of several frequencies and amplitudes; these constitute the *spectrum* of the sound. For instance, most wind and string instruments have *harmonic spectrum*, i.e., their spectral frequencies are integer multiples of a *fundamental frequency* f_1, that is, $f_n = n f_1$, with $n \geq 1$. Such a sound is perceived to have a pitch identified with frequency f_1. However, not every sound can be attributed a pitch; some musical instruments, for instance most drums, have indefinite pitch.

Intervals, octave, semitone, cents. The difference between two pitches is called *interval*. The pitch perception obeys two fundamental rules. One is the logarithmic correspondence: the interval from two pitches of frequencies ν_1, ν_2 only depends on their *frequency ratio* $\nu_2 : \nu_1$. The other one is the octave equivalence: two pitches an *octave* apart (ratio 2:1) are musically equivalent.

One can measure intervals in the multiplicative scale by their frequency ratio, or in the additive scale by their size expressed in octaves, for instance. A frequency ratio of r corresponds to $\log_2 r$ octaves. Other important intervals are the *semitone*, which is $1/12$ of an octave (therefore its ratio is $2^{1/12}$), and the *cent*, which is $1/100$ of a semitone.

The human ear is exceedingly sensitive to pitch perception. The *difference limen* (or just noticeable difference) between two tones can be, depending on the frequency and intensity, as small as 10 cents. It can be much smaller if both sounds are played together. Therefore it is of the greater importance to correctly tune a musical instrument.

Notes, scales, 12-edo. In some instruments (e.g. the violin) the player can play virtually any pitch within its playing range. This is not true for other instruments (e.g. the piano, or most wind instruments), where only a finite set of pitches is directly playable. A selection of pitches to play music is called a *scale*, and its elements are the *notes* of the scale. The construction of these scales is one of the fundamental problems in music theory. Notice that if we have chosen a scale on a theoretical basis, then we have to adjust or to *tune* the pitches of the notes of the instrument to the pitches of the scale; therefore one frequently says *tuning system* to mean a scale.

From ancient times it is known that two similar strings sounding together are more pleasant when their fundamental frequencies are in a ratio of *small integers*. These intervals are called *just*, and, in addition to the octave, the most important ones are the fifth (ratio 3:2), the fourth (4:3) and the major and minor thirds (5:4 and 6:5). (These names have a historic origin, of course.) So one would look for scales whose notes define such intervals. But, of course, other intervals will appear, and maybe they will be not so pleasant. Moreover, the evolution of the musical language during the last centuries has added more requirements to the scales, and as a result the problem of defining a scale does not have a universal optimal solution. What is more, from antiquity to modern times, *dozens* of scales have been proposed and put into practice, [9].

Among all of these scales, there is one that is pervasive in Western music since 19-th century. It is the so-called equal temperament, and consists of 12 equal divisions of the octave (12-edo). The explanation for this choice is that the 12-edo scale yields an excellent approximation of the just fifth ($2^{7/12} \approx 3:2$) and the just fourth, but also acceptable approximations of the just thirds.

It is worth noting that, for instance, the same name "fifth" is applied to two intervals that are indeed different: the just fifth and the 12-edo fifth. This is usual: the traditional name of an interval applies to all the intervals that have the same musical function regardless of their exact tuning. The same happens with the notes: A4 has nowadays a "standard pitch" of 440 Hz, but it is usual to tune this note to 442 Hz, for instance. In past times its values were much more diverse.

Due to the logarithmic correspondence, from the viewpoint of music theory, to define a scale one can fix the pitch ν_0 of a *fundamental note* with some degree of arbitrariness; what is really important are the intervals $r_k = \nu_k : \nu_0$ between this note and the other ones, ν_k. From these intervals, and the fundamental note, the other notes can be reconstructed as $\nu_k = r_k\, \nu_0$. Alternatively, one can define a scale by giving the *steps* between consecutive notes, $\nu_k : \nu_{k-1}$; for instance, the 12-edo scale has equal steps of ratio $2^{1/12}$.

Beats, dissonance, consonance. Using trigonometric identities it is easily proved that the superposition of two pure tones $A \sin(2\pi\nu_1 t)$ and $A \sin(2\pi\nu_2 t)$ can be expressed as $2\,A \cos\left(2\pi \frac{\nu_1-\nu_2}{2}\, t\right) \sin\left(2\pi \frac{\nu_1+\nu_2}{2}\, t\right)$. If the frequency difference $\nu_1 - \nu_2$ is small (about less than 10–15 Hz), this is perceived as a sound of frequency $\nu = \frac{\nu_1+\nu_2}{2}$ with slowly fluctuating amplitude; these are the *beats*. If the frequency difference is somewhat bigger, one perceives some *roughness*. When the difference is even higher, then one perceives two separate tones [12, pp. 37–40]. This roughness gives rise to the notion of *sensory dissonance*; this is the only notion of dissonance we are concerned about, though there are others (see, for example, [11] and [15, Ch. 5]).

Now, consider two (or more) complex tones sounding together: they have many partials that may be close in frequencies. In the middle of the 19-th century, H. Helmholtz described the dissonance as the roughness produced by close partials, and the *consonance* as the exceptional condition where this roughness almost disappears. By computing the beats of the partials of two harmonic tones, Helmholtz showed that the aforementioned just intervals (octave, fifth, fourth, thirds) are more consonant than others [8, p. 193], in good agreement with music theory. This result can be easily understood: just notice that, if the fundamental frequencies $f, \tilde{f}$ of two harmonic tones are in a ratio of small integers $\tilde{f} : f = \ell : k$, then the ℓ-th partial of the first tone will coincide with the k-th partial of the second one; a slight change of $\tilde{f}$ would lead to close but different partials, and therefore to some roughness.

References

[1] D.J. Benson, "Music, a mathematical offering", Cambridge University Press, 2008.

[2] H. Brezis, "Functional analysis, sobolev spaces and partial differential equations", Springer, 2011.

[3] J. Chabassier, "Modélisation et simulation numérique d'un piano par modèles physiques", Ph.D. thesis, École Polytechnique ParisTech, 2012.

[4] L.C. Evans, "Partial differential equations", AMS Press, 1998.

[5] H. Fletcher, "Normal vibration frequencies of

a stiff piano string", *J. Acoust. Soc. Am.* **36** (1964), 203–209.

[6] N.H. Fletcher and T.D. Rossing, "The physics of musical instruments" (2nd ed), Springer, 1998.

[7] S.M. Han, H. Benaroya, and T. Wei,. "Dynamics of transversely vibrating beams using four engineering theories", *J. Sound Vibr.* **225** (1999), 935–988.

[8] H. Helmholtz, "On the sensations of tone" (3rd ed), Longmans (original in German, 1877, 4th ed), 1895.

[9] M. Lindley, "Temperaments", *Grove Music Online*, 2015.

[10] P.M. Morse and K.U. Ingard, "Theoretical acoustics", McGraw-Hill, 1968.

[11] R. Plomp and W.J.M. Levelt, "Tonal consonance and critical bandwidth", *J. Acoust. Soc. Am.* **38** (1965), 548–560.

[12] J.G. Roederer, "The physics and psychophysics of music", Springer, 2008.

[13] T.D. Rossing, F.R. Moore, and P.A. Wheeler, "The science of sound" (3rd ed), Addison-Wesley, 2002.

[14] S. Salsa, "Partial differential equations in action. From modelling to theory", Springer, 2008.

[15] W.A. Sethares, "Tuning, timbre, spectrum, scale" (2nd ed), Springer, 2005.

[16] W.A. Strauss, "Partial differential equations. An introduction", Wiley, 2008.

[17] A.N. Tikhonov and A.A. Samarskii, "Equations of mathematical physics". Pergamon, 1963.

REPORTS@SCM

AN ELECTRONIC JOURNAL OF THE
SOCIETAT CATALANA DE MATEMÀTIQUES

Explicit bounds for growth of sets in non-abelian groups

*Alberto Espuny Díaz
Universitat Politècnica de Catalunya
alberto.espuny@estudiant.upc.edu
*Corresponding author

Resum *(CAT)*
Les desigualtats de Plünnecke–Ruzsa proporcionen fites superiors del creixement de les sumes iterades d'un conjunt en un grup abelià. Aquesta mena de resultats han estat estesos recentment per Petridis i per Tao al cas no abelià. El resultat principal d'aquest treball és una demostració de la desigualtat de Plünnecke–Ruzsa pel cas no abelià que no fa servir el mètode de Petridis. També s'obtenen resultats anàlegs pel cas del producte iterat de dos conjunts diferents.

Abstract *(ENG)*
The Plünnecke–Ruzsa inequalities give upper bounds for the growth of iterated sumsets in an abelian group. These inequalities have been recently extended to the non-abelian case by Petridis and by Tao. The main result in this work is a proof of the non-abelian Plünecke–Ruzsa inequalities which makes no use of the method introduced by Petridis. Analogous inequalities for iterated products of two distinct sets are also obtained.

Keywords: *Additive combinatorics, combinatorial number theory, growth in groups.*
MSC (2010): *Primary 11B13, 11B30, 11P70. Secondary 20D60.*
Received: *December 2, 2015.*
Accepted: *December 31, 2016.*

Acknowledgement
The author was partially supported by grant 2015/COLAB/00069 of the Spanish Ministerio de Educación, Cultura y Deporte. He would also like to thank Profs. Juanjo Rué and Oriol Serra for their many insightful suggestions, as well as the anonymous referee, whose comments and remarks greatly improved the final version of this paper.

Societat Catalana de Matemàtiques

Institut d'Estudis Catalans

1. Introduction

The theory of set addition was initiated by Freiman [2] in the 1960's in the context of abelian groups. More recently, a lot of effort has been directed at extending this theory to the non-abelian case, as well as searching for connections between this and many other areas of mathematics, such as Lie group theory, number theory, or probability theory; see [1, 3]. In this context, the group operation is usually referred to as set multiplication, instead of set addition. The product of two sets is defined as

$$AB = \{ab \mid a \in A,\ b \in B\}.$$

One can also define iterated product sets recursively, and define the inverse of a set as the set of the inverses. Note that the inverse of a set has the same size as the set itself.

One of the most basic and important problems in this setting is bounding the growth of iterated product sets. Some trivial tight bounds can be found, but a more interesting problem arises when bounding iterated product sets, given a bound for the base product set $|AA| \leq \alpha|A|$. In this sense, the first important result is Plünnecke's inequality, which was first proved in Plünnecke [5] in the late 1960's, and has become one of the most commonly used tools in additive combinatorics. His proof is based on a graph-theoretic method, using what he called *commutative graphs* in order to obtain the following result.

Theorem 1.1 (Plünnecke's inequality)**.** *Let j, h be two non-negative integers such that $j < h$, and let A and B be finite sets in a commutative group. Assume that $|A| = m$ and $|AB^j| = \alpha m$. Then, there exists a non-empty set $X \subseteq A$ such that $|XB^h| \leq \alpha^{\frac{h}{j}}|X|$.*

The proof is simple in its technical parts, but long and arduous. Variations of this proof have been used to prove some more general results.

Ruzsa [6] rediscovered Plünnecke's work by providing a proof based on Menger's theorem on graph theory, and used some of his own techniques to obtain some more general results. One of the most important states as follows.

Theorem 1.2 (Plünnecke–Ruzsa inequality)**.** *Let A and B be finite sets in a commutative group, and j be a positive integer. Assume that $|BA^j| \leq \alpha|B|$. Then, for any nonnegative integers k and l such that $j \leq \min\{k, l\}$, we have that $|A^lA^{-k}| \leq \alpha^{\frac{k+l}{j}}|B|$.*

The main limitation of these results is that they only hold in abelian groups (in fact, it is usual to write them using additive notation). Furthermore, one can find counterexamples for an extension to the non-commutative case. The only exception to this rule is when considering triple product sets; in this case, Plünnecke's graph-theoretic method can be used to obtain some results (see, for instance, Section 2). For this reason, finding statements that resemble those of Plünnecke and Ruzsa and hold in the non-commutative setting recently became an interesting problem.

Under these circumstances, Tao was the first to realise, in the late 2000's, that one has to impose a further restriction on the sets under consideration. In some sense, it is enough to see that the growth of triple products of a set is bounded, $|AAA| \leq \alpha|A|$, in order to bound all its iterated product sets. He also proved that one can weaken this condition to $|AaA| \leq \alpha|A|$ for every $a \in A$, and a bound for any product sets of A can be given. His more general theorem in this setting can be stated as follows.

Theorem 1.3 (Tao, [10])**.** *Let A be a finite set in a group. Assume that $|AaA| \leq \alpha|A|$ for every $a \in A$, and that $|AA| \leq \alpha|A|$. Then, there exists some absolute constant c such that, for any signs $\epsilon_1, \dots, \epsilon_h \in \{-1, 1\}$, $|A^{\epsilon_1}A^{\epsilon_2}\cdots A^{\epsilon_h}| \leq \alpha^{ch}|A|$.*

Setting all the signs to be equal in Theorem 1.3, one obtains the following corollary.

Corollary 1.4. *Let A be a finite set in a group such that $|AaA| \leq \alpha|A|$ for every $a \in A$, and $|AA| \leq \alpha|A|$. Then, there is an absolute constant c such that $|A^h| \leq \alpha^{ch}|A|$.*

In 2011, Petridis presented a new method to prove estimates of the growth of product sets; see [4]. With his new technique, he was able to give an elementary proof of Plünnecke's inequality and several other results. He also used his method to give a specific value to the absolute constant in Corollary 1.4.

Theorem 1.5 (Petridis)**.** *Let A be a finite set in a group. Suppose that $|AA| \leq \alpha|A|$ and $|AaA| \leq \beta|A|$ for every $a \in A$. Then, for all $h > 2$, $|A^h| \leq \alpha^{8h-17}\beta^{h-2}|A|$.*

In the statement of Tao's theorem we had $\alpha = \beta$ so, Petridis's result gives a constant $c = 9$ as an upper bound for c. The same approach using Petridis's new method serves to obtain bounds in similar settings, when considering more than one set.

In this paper, we work with the growth of sets under multiplication in the non-commutative setting, and with results similar to Theorem 1.3 and Theorem 1.5. In particular, one of the results is a weaker version of Theorem 1.5 that can be obtained without reference to Petridis's new method, and hence could have been developed before.

Theorem 1.6. *Let A be a finite set in a group such that $|AA| \leq \alpha|A|$ and $|AaA| \leq \beta|A|$, for every $a \in A$. Then, for any $h > 2$, $|A^h| \leq \alpha^{9h-19}\beta^{h-2}|A|$.*

Additionally, we use Theorem 1.5 in order to obtain estimates for the size of iterated products of two different sets. We also exploit it in order to give a specific value for the constant c in Theorem 1.3.

Theorem 1.7. *Let A be a non-empty finite set in a group such that $|AA| \leq \alpha|A|$ and $|AaA| \leq \beta|A|$, for every $a \in A$. Then, for any signs $\epsilon_1, \dots, \epsilon_h \in \{-1, 1\}$, $|A^{\epsilon_1}A^{\epsilon_2}\cdots A^{\epsilon_h}| \leq \alpha^{8h-15}\beta^{h-2}|A|$.*

The remainder of this paper is divided in the following way. In Section 2 we present the three results that are needed in order to obtain our new results. In Section 3 we use these results to present the proof of Theorem 1.6, and use it to give a specific value to the constant in Corollary 1.4. Finally, in Section 4 we use Theorem 1.5 and the tools from Section 2 to prove Theorem 1.7. Furthermore, we prove several results when considering the product of two different sets.

2. Tools

In this section we present the tools needed for the proofs of the results of Sections 3 and 4. Two of them are elementary, and their proofs are presented here for the sake of completeness. The third one is a non-commutative theorem by Ruzsa, which cannot be proved without a thorough presentation of the graph-theoretic method designed by Plünnecke. An account of its proof can be found, for example, in [9].

The first versions of the tools we present here were developed specifically for the abelian case. However, they could easily be extended to the non-commutative setting. The first of these tools is known as Ruzsa's triangle inequality.

Theorem 2.1 (Ruzsa's triangle inequality, [7]). *Let X, Y and Z be finite non-empty sets in a (not necessarily commutative) group. Then, $|X||YZ^{-1}| \leq |YX^{-1}||XZ^{-1}|$.*

Proof. The idea of the proof is to find an injection of $X \times (YZ^{-1})$ into $(YX^{-1}) \times (XZ^{-1})$. Since the sizes of these sets are $|X||YZ^{-1}|$ and $|YX^{-1}||XZ^{-1}|$, respectively, this yields the result.

Consider the following map:

$$\begin{aligned} \varphi\colon X \times (YZ^{-1}) &\longrightarrow (YX^{-1}) \times (XZ^{-1}) \\ (x, yz^{-1}) &\longmapsto (yx^{-1}, xz^{-1}). \end{aligned}$$

We would like to see that this is an injection. First, observe that an element $yz^{-1} \in YZ^{-1}$ may come from different elements $y, y' \in Y$ and $z, z' \in Z$ such that $yz^{-1} = y'z'^{-1}$. Hence, we must first fix a representation in Y, Z for each element of YZ^{-1}. We do so by defining an injection $f\colon YZ^{-1} \longrightarrow Y \times Z$ such that $f(a)_Y f(a)_Z^{-1} = a$ for every $a \in YZ^{-1}$, where $f(a)_Y$ denotes the first coordinate of $f(a)$, and $f(a)_Z$ denotes the second. Such an injection exists because of the definition of the set YZ^{-1}. For example, if we give the elements of Y some order $y_1 < y_2 < \cdots < y_k$, we could map a to the pair (y_i, z_j) such that $y_i z_j^{-1} = a$ and the index i is minimum.

Now, assume that $\varphi(x, a) = \varphi(x', a')$. Then, $f(a)_Y x^{-1} = f(a')_Y x'^{-1}$ and $xf(a)_Z^{-1} = x'f(a')_Z^{-1}$ and, multiplying these two equalities, we get that $f(a)_Y f(a)_Z^{-1} = f(a')_Y f(a')_Z^{-1}$. By definition of f, this means that $a = a'$. Substituting this in the former equations yields $x = x'$ so, φ is an injection. □

The second tool is the simplest of a group of results known as covering lemmas.

Lemma 2.2 (Ruzsa's covering lemma, [8]). *Let A and B be finite sets in a group G. Assume that $|AB| \leq \alpha|A|$. Then, there exists a non-empty set $S \subseteq B$ such that $|S| \leq \lfloor \alpha \rfloor$ and $B \subseteq A^{-1}AS$.*

Proof. The proof follows from choosing $S \subseteq B$ in the right way. Select S to be maximal subject to As_1 being disjoint with As_2 for every pair $s_1, s_2 \in S$. This is equivalent to choosing S to be maximal subject to $|AS| = |A||S|$ being true.

Now, take $b \in B$. We distinguish two possible cases: if $b \in S$ then, for any $a \in A$, we have that $b = a^{-1}ab \in A^{-1}AS$. Otherwise, $b \notin S$, and b cannot be added to S without breaking the maximality condition so, there must be an element $s \in S$ such that $Ab \cap As \neq \varnothing$; equivalently, there exist some elements $s \in S$, $a, a' \in A$ such that $ab = a's$ hence, $b = a^{-1}a's \in A^{-1}AS$. □

Finally, Ruzsa's non-commutative bound for the product set of three sets in a non-commutative setting can be stated as follows.

Theorem 2.3 (Ruzsa,[9]). *Let A, B and C be finite sets in a group G. Assume that $|AB| \leq \alpha_1|A|$ and $|CA| \leq \alpha_2|A|$. Then, there exists a set $\varnothing \neq X \subseteq A$ such that $|CXB| \leq \alpha_1\alpha_2|X|$.*

3. An explicit value for Tao's theorem

A combination of the three tools presented in the previous section can be used to give a value to the constant c in Tao's Corollary 1.4. We start by using Ruzsa's triangle inequality to prove a lemma that appeared in Petridis [4]. Let us mention that this lemma is not related to Theorem 1.5 of said paper, which is its main contribution.

Lemma 3.1 (Petridis)**.** *Let A and B be finite non-empty sets in a group. Suppose that $|AA| \leq \alpha|A|$ and $|ABA| \leq \alpha^2|A|$. Then, $|AB^{-1}BA^{-1}| \leq \alpha^6|A|$.*

Proof. In Theorem 2.1, take $X = A$ and $Y = Z = AB^{-1}$. Then, we have that

$$|A||AB^{-1}BA^{-1}| \leq |AB^{-1}A^{-1}||ABA^{-1}| = |ABA^{-1}|^2$$

since $\left(AB^{-1}A^{-1}\right)^{-1} = ABA^{-1}$ and a set and its inverse have the same cardinality. In order to bound this, take Theorem 2.1 again and consider $X = A^{-1}$, $Y = AB$ and $Z = A$. This yields

$$|A||ABA^{-1}| \leq |ABA||A^{-1}A^{-1}| = |ABA||AA| \leq \alpha^3|A|^2$$

so, $|ABA^{-1}| \leq \alpha^3|A|$. Substituting this above and dividing by $|A|$ results in the statement. □

We can use this to prove the following result.

Theorem 3.2. *Let A be a finite set in a group. Assume that $|AA| \leq \alpha|A|$ and $|AaA| \leq \beta|A|$, for every $a \in A$. Then, $|AAA| \leq \alpha^8\beta|A|$.*

Proof. We can use Theorem 2.3 setting $A = C = B$. The theorem states that there exists some set $T \subseteq A$ such that $|ATA| \leq \alpha^2|T|$.

We can now use the trivial bound $|TA| \leq |ATA| \leq \alpha^2|T|$ for the hypothesis of Lemma 2.2. Applying this covering lemma, we have that there exists a set $S \subseteq A$ of size $|S| \leq \alpha^2$ such that $A \subseteq T^{-1}TS$. Hence, we have that $AAA \subseteq AT^{-1}TSA$.

Consider Ruzsa's triangle inequality in the form of Theorem 2.1, and substitute $X = A$, $Y = AT^{-1}T$, and $Z = A^{-1}S^{-1}$ to obtain $|A||AAA| \leq |A||AT^{-1}TSA| \leq |AT^{-1}TA^{-1}||ASA|$.

Now, we can use Lemma 3.1 to bound the first of these product sets. We can do this because we have $|AA| \leq \alpha|A|$, and $|ATA| \leq \alpha^2|T| \leq \alpha^2|A|$ since $T \subseteq A$, so we have all the hypothesis needed. To bound the second one, consider

$$|ASA| = \left|\bigcup_{s\in S} AsA\right| \leq \sum_{s\in S}|AsA| \leq \sum_{s\in S}\beta|A| = |S|\beta|A| \leq \alpha^2\beta|A|.$$

Putting everything together, we have that $|A||AAA| \leq \alpha^6|A|\alpha^2\beta|A| = \alpha^8\beta|A|^2$; and dividing by $|A|$ gives the desired result. □

Now, we can use this theorem as a base case to inductively obtain bounds on the size of higher product sets.

Proof of Theorem 1.6. The proof is done by induction on h. The base case $h = 3$ has been proved in Theorem 3.2. Let us prove the general case. Assume that $h > 3$. Using Ruzsa's triangle inequality with $X = A^{-1}$, $Y = AA$ and $Z^{-1} = A^{h-2}$, we have

$$|A^h| \leq \frac{|AAA||A^{-1}A^{h-2}|}{|A|}.$$

Taking now $X = A$, $Y = A^{-1}$ and $Z^{-1} = A^{h-2}$ yields

$$|A^{-1}A^{h-2}| \le \frac{|AA||A^{h-1}|}{|A|}.$$

Putting both equations together and using Theorem 3.2, we obtain $|A^h| \le \alpha^9\beta|A^{h-1}|$, and the last term is bounded by the induction hypothesis. $\square$

In the particular case when $\beta = \alpha$, this result gives us $c = 10$ in the statement of Tao's Corollary 1.4. This constant is worse that the one obtained in Petridis [4] by one unit. However, it can be obtained without using Petridis's new method, so it is interesting by itself. Observe that Plünnecke's graph-theoretic method is necessary in order to obtain this bound, as it is needed to prove Theorem 2.3.

4. Further results

We can use Theorem 1.6 in order to get new product estimates. One may consider more sets and impose further restrictions on them. For example, we may consider the iterated product of two sets A and B with restrictions over the product sets of A, the product sets of A and B, and the size of each other. With this, we can obtain a bound for iterated product sets of two sets. As we have already observed that Theorem 1.5 gives a better bound on product sets than Theorem 1.6, we will use Petridis's result in this section in order to obtain tighter bounds.

We start with a result giving a bound for the size of the product set of A and an iterated product of B's. From this point onwards, Ruzsa's triangle inequality (i.e., Theorem 2.1) will be used repeatedly without warning, with X always being a simple set A or B, or one of their inverses.

Theorem 4.1. *Let A and B be two finite non-empty sets in a group. Assume that $|AA| \le \alpha|A|$, $|AaA| \le \beta|A|$ for every $a \in A$, $|AB| \le \delta|A|$, $|AbB| \le \varepsilon|A|$ for every $b \in B$, and $|A| \le \gamma|B|$. Then, for any $k \ge 2$,*

$$|AB^k| \le \begin{cases} \alpha^{16(k-1)}\,\beta^{2(k-1)}\,\gamma^{k-2}\,\delta^{2k-3}\,\varepsilon^{k-1}|A| & \text{if } k \text{ is even,} \\ \alpha^{16(k-1)}\,\beta^{2(k-1)}\,\gamma^{k-1}\,\delta^{2k-1}\,\varepsilon^{k-1}|A| & \text{if } k \text{ is odd.} \end{cases}$$

Proof. The proof is done by induction on k. We need two base cases in order to complete the induction.

When $k = 2$ we can apply Ruzsa's covering lemma due to the third condition on the sets. This gives us a set $S \subseteq B$ such that $|S| \le \lfloor\delta\rfloor$ and $B \subseteq A^{-1}AS$. Hence,

$$|ABB| \le |AA^{-1}ASB| \le \frac{|AA^{-1}AA^{-1}||ASB|}{|A|}.$$

The second term in this expression can be bounded as

$$|ASB| = \left|\bigcup_{s\in S} AsB\right| \le \sum_{s\in S}|AsB| \le \sum_{s\in S}\varepsilon|A| = |S|\varepsilon|A| \le \delta\varepsilon|A|. \tag{1}$$

In order to bound the first one, use Theorem 1.5 to obtain

$$|AA^{-1}AA^{-1}| \le \frac{|AAA^{-1}|^2}{|A|} \le \frac{(|AA||AAA|)^2}{|A|^3} \le \alpha^{16}\beta^2|A|. \tag{2}$$

Putting the two expressions together we have

$$|ABB| \leq \alpha^{16}\beta^2\delta\varepsilon|A|. \tag{3}$$

For $k = 3$, we use again the covering lemma with the same conditions as above and get

$$|ABBB| \leq |AA^{-1}ASBB| \leq \frac{|AA^{-1}AA^{-1}||ASBB|}{|A|},$$

as $B \subseteq A^{-1}AS$ for some $S \subseteq B$ with $|S| \leq \delta$. The first term is bounded by (2). In order to bound the second term, consider that

$$|A||ASBB| \leq |ASBA^{-1}||AB| \leq \delta|ASBA^{-1}||A|,$$
$$|B||ASBA^{-1}| \leq |ASB||B^{-1}BA^{-1}| \leq \delta\varepsilon|B^{-1}BA^{-1}||A|,$$
$$|A||B^{-1}BA^{-1}| \leq |B^{-1}A^{-1}||ABA^{-1}| \leq \delta|ABA^{-1}||A|$$

and

$$|B||ABA^{-1}| \leq |ABB||B^{-1}A^{-1}| \leq \delta|A|\alpha^{16}\beta^2\delta\varepsilon|A|,$$

where the last inequalities in each line come from (1) in the second line, (3) in the fourth, and the statement hypothesis in all the others. With this,

$$|ASBA^{-1}| \leq \delta\varepsilon\frac{|A|}{|B|}\delta\delta\frac{|A|}{|B|}\alpha^{16}\beta^2\delta\varepsilon|A| \leq \alpha^{16}\beta^2\gamma^2\delta^4\varepsilon^2|A| \tag{4}$$

and $|ABBB| \leq \alpha^{16}\beta^2\delta\alpha^{16}\beta^2\gamma^2\delta^4\varepsilon^2|A| = \alpha^{32}\beta^4\gamma^2\delta^5\varepsilon^2|A|$.

For the general case, we can use the covering lemma in the same way. We have that

$$|AB^k| \leq |AA^{-1}ASB^{k-1}| \leq \frac{|AA^{-1}AA^{-1}||ASB^{k-1}|}{|A|}.$$

The first term is, again, bounded by (2), and the second is bounded as $|A||ASB^{k-1}| \leq |ASBA^{-1}||AB^{k-2}|$. The first term is now bounded by (4), and the second one is bounded by the induction hypothesis. Putting everything together the result follows. □

Using the previous result, we can give a general bound for iterated products of A's and B's, as long as all the A's come before the B's.

Theorem 4.2. *Let A and B be two finite non-empty sets in a group, with the conditions from Theorem 4.1. Then, for any $l \geq 2$ and $k \geq 2$,*

$$|A^lB^k| \leq \begin{cases} \alpha^{8l+16k-24}\,\beta^{2l+2k-3}\,\gamma^{k-2}\,\delta^{2k-3}\,\varepsilon^{k-1}|A| & \text{if } k \text{ is even,} \\ \alpha^{8l+16k-24}\,\beta^{2l+2k-3}\,\gamma^{k-1}\,\delta^{2k-1}\,\varepsilon^{k-1}|A| & \text{if } k \text{ is odd.} \end{cases}$$

Proof. As before, we can use Ruzsa's triangle inequality (twice) to bound

$$|A^lB^k| \leq \frac{|A^{l+1}||AB^k||AA|}{|A|^2}.$$

The three terms can now be bounded using Theorem 1.5, Theorem 4.1 and the statement hypotheses, respectively, and this immediately yields the result. □

In order to complete all the bounds of product sets of three or more sets as those we have presented so far, the only remaining case is that when $l \geq 2$ and $k = 1$.

Theorem 4.3. *Let A and B be two finite non-empty sets in a group, with the conditions from Theorem 4.1. Then, for any $l \geq 2$, $|A^l B| \leq \alpha^{8(l-1)} \beta^{l-1} \delta |A|$.*

Proof. We start by proving the base case $l = 2$. Using Ruzsa's triangle inequality we have

$$|AAB| \leq \frac{|AAA^{-1}||AB|}{|A|}$$

and

$$|AAA^{-1}| \leq \frac{|AAA||AA|}{|A|}$$

so, using Theorem 1.5 and putting everything together, we get $|AAB| \leq \alpha^7 \beta \alpha \delta |A|$.

For the general case ($l > 2$), observe that

$$|A^l B| \leq \frac{|AAA^{-1}||A^{l-1}B|}{|A|}.$$

The first term can be bounded using Theorem 2.1 and Theorem 1.5 as $|AAA^{-1}| \leq \alpha^8 \beta |A|$, as before, and the second is bounded by the induction hypothesis. □

An easy corollary is obtained when taking $B = A^{-1}$ in Theorem 4.2. This would correspond to an extension of the Plünnecke–Ruzsa inequality to the non-commutative case, when $A = B$.

Corollary 4.4. *Let A be a non-empty finite set in a group such that $|AA| \leq \alpha |A|$ and $|AaA| \leq \beta |A|$, for every $a \in A$. For any $k, l \geq 2$, let $m = \min\{k, l\}$ and $n = \max\{k, l\}$. Then,*

$$|A^l A^{-k}| \leq \alpha^{8n+21m-27} \beta^{2n+3m-4} |A|.$$

Proof. Take $B = A^{-1}$. For this particular choice of sets we have $\gamma = 1$ and, in virtue of Ruzsa's triangle inequality, $|AA^{-1}| \leq \alpha^2 |A|$ and $|AaA^{-1}| \leq \alpha\beta |A|$. Substituting these into Theorem 4.2, we can write

$$|A^l A^{-k}| \leq \begin{cases} \alpha^{8l+21k-31} \beta^{2l+3k-4} |A| & \text{if } k \text{ is even,} \\ \alpha^{8l+21k-27} \beta^{2l+3k-4} |A| & \text{if } k \text{ is odd.} \end{cases}$$

First, observe that this can be written in such a way that it does not depend on the parity by taking the worst exponent for each of the coefficients. Since these coefficients are all lower-bounded by 1 for this choice of sets, this means we must take the highest exponents, which correspond to the odd case. Then, for any $k, l \geq 2$ we may write

$$|A^l A^{-k}| \leq \alpha^{8l+21k-27} \beta^{2l+3k-4} |A|.$$

Observe now that this is a symmetric result. That is, the fact that a set and its inverse have the same size means that $|A^l A^{-k}| = |A^k A^{-l}|$. Then, as the bound given by the above expression is much weaker when $k > l$, if this occurs one can use the bound for the size of the inverse set to obtain a better bound. This is what allows us to take the minimum and the maximum of k and l. □

We can use this to obtain a particular bound for Tao's Theorem 1.3 if we impose $\alpha = \beta$. In this case, we have $|A^l A^{-k}| \leq \alpha^{10l+24k-31}|A|$.

If we want to obtain a constant c with respect to $h = l + k$, as appears in the statement of Theorem 1.3, we have to consider the following. The exponent $10l + 24k - 31$, for a fixed h, is increasing with k and maximized when $l = k$ because of the possibility to take the maximum and minimum of k and l. Hence,

$$10l + 24k - 31 \leq 10\frac{h}{2} + 24\frac{h}{2} - 31 \leq 34\frac{h}{2} = 17h$$

so, we have $c = 17$ for all these different cases.

However, we can obtain a much better constant, in a more general setting, if we work with the sets A and A^{-1} from the beginning.

Proof of Theorem 1.7. We start with the base case $h = 3$. We must consider all the possible signs that can appear in the exponents. By using Theorem 1.5, we have that

$$|AAA| \leq \alpha^7\beta|A|,$$

$$|AAA^{-1}| \leq \frac{|AAA||AA|}{|A|} \leq \alpha^8\beta|A|,$$

$$|A^{-1}AA| \leq \frac{|AA||AAA|}{|A|} \leq \alpha^8\beta|A|,$$

and

$$|AA^{-1}A| \leq \frac{|AA^{-1}A^{-1}||AA|}{|A|} \leq \alpha^9\beta|A|.$$

The other four possible configurations are the inverses of these ones. Hence, in general,

$$|A^{\epsilon_1}A^{\epsilon_2}A^{\epsilon_3}| \leq \alpha^9\beta|A|.$$

For the general case, there are two different possibilities. First, assume $\epsilon_1 = \epsilon_2$. Then, take $X = A^{-\epsilon_1}$ in the statement of Ruzsa's triangle inequality to obtain

$$|A||A^{\epsilon_1}A^{\epsilon_1}A^{\epsilon_3}\cdots A^{\epsilon_h}| \leq |A^{\epsilon_1}A^{\epsilon_1}A^{\epsilon_1}||A^{-\epsilon_1}A^{\epsilon_3}\cdots A^{\epsilon_h}| \leq \alpha^7\beta|A^{-\epsilon_1}A^{\epsilon_3}\cdots A^{\epsilon_h}||A|.$$

If, on the contrary, $\epsilon_1 = -\epsilon_2$, we have

$$\begin{aligned} |A||A^{\epsilon_1}A^{\epsilon_2}A^{\epsilon_3}\cdots A^{\epsilon_h}| &\leq |A^{\epsilon_1}A^{\epsilon_1}||A^{\epsilon_2}A^{\epsilon_2}A^{\epsilon_3}\cdots A^{\epsilon_h}| \\ &\leq \alpha|A^{\epsilon_2}A^{\epsilon_2}A^{\epsilon_2}||A^{-\epsilon_2}A^{\epsilon_3}\cdots A^{\epsilon_h}| \\ &\leq \alpha^8\beta|A^{-\epsilon_2}A^{\epsilon_3}\cdots A^{\epsilon_h}||A|. \end{aligned}$$

The worst exponent is given in the second case. The two last sets can be bounded by the induction hypothesis. □

Taking $\alpha = \beta$ in the statement of Theorem 1.7, we obtain $|A^{\epsilon_1}A^{\epsilon_2}\cdots A^{\epsilon_h}| \leq \alpha^{9h-17}|A|$, for any signs $\epsilon_1, \dots, \epsilon_h \in \{-1, 1\}$. With this, we have an explicit constant value for the statement of Tao's Theorem 1.3, $c \leq 9$, so we have that the same constant working when all signs are set to be equal serves in the rest of cases as well.

References

[1] E. Breuillard, B. Green, and T. Tao, "The structure of approximate groups", *Publ. Math. Inst. Hautes Études Sci.* **116** (2012), 115–221.

[2] G.A. Freĭman, "Foundations of a structural theory of set addition", American Mathematical Society, Providence, R.I. (1973), vii+108.

[3] H.A. Helfgott, "Growth in groups: ideas and perspectives", *Bull. Amer. Math. Soc.* **52**(3) (2015), 357–413.

[4] G. Petridis, "New proofs of Plünnecke-type estimates for product sets in groups", *Combinatorica* **32**(6) (2012), 721–733.

[5] H. Plünnecke, "Eine zahlentheoretische Anwendung der Graphentheorie", *J. Reine Angew. Math.* **243** (1970), 171–183.

[6] I.Z. Ruzsa, "An application of graph theory to additive number theory", *Sci. Ser. A Math. Sci.* **3** (1989), 97–109.

[7] I.Z. Ruzsa, "Sums of finite sets", in "Number theory", (1996) Springer, New York, 281–293.

[8] I.Z. Ruzsa, "An analog of Freĭman's theorem in groups", *Astérisque* **258** (1999), 323–326.

[9] I.Z. Ruzsa, "Sumsets and structure", *Adv. Courses Math. CRM Barcelona* (2009), Birkhäuser Verlag, Basel, 87–210.

[10] T. Tao, "Product set estimates for non-commutative groups", *Combinatorica* **28**(5) (2008), 547–594.

REPORTS@SCM
AN ELECTRONIC JOURNAL OF THE
SOCIETAT CATALANA DE MATEMÀTIQUES

Compactification of a diagonal action on the product of CAT(-1) spaces

*Teresa García
Universitat Autònoma de Barcelona
tggalvez@uab.math.cat
*Corresponding author

Resum *(CAT)*
Sigui X un espai CAT(-1) propi i no compacte, i sigui Γ un subgrup discret cocompacte de les isometries de X. Compactifiquem l'acció diagonal de Γ a $X \times X$ considerant un domini de la frontera per horofuncions respecte a la mètrica del màxim.

Abstract *(ENG)*
Let X be a proper, non-compact CAT(-1) space, and Γ a discrete cocompact subgroup of the isometries of X. We compactify the diagonal action of Γ on $X \times X$ considering a domain of the horofunction boundary with respect to the maximum metric.

Keywords: *CAT(-1) space, group of isometries, compactification, horofunctions.*
MSC (2010): *Primary: 55M25, 57P10. Secondary: 55P15, 57R19, 57N15.*
Received: *June 9th, 2016.*
Accepted: *December 7th, 2016.*

Acknowledgement
This work has been supported by grants BES-2013-06571 (MINECO) and MTM2015-66165-P (MINECO/FEDER). The author thanks her advisor, Joan Porti, for his unvaluable advice.

1. Introduction

Let (X, d_X) be a proper, non-compact CAT(-1) space. As examples of such spaces one can think of complete simply connected Riemannian manifolds of negative sectional curvature, for instance the hyperbolic space, and of metric trees. For a review on CAT(-1) spaces one can consult the first section of Bourdon [3]. Here, we consider a discrete and cocompact subgroup Γ of the isometries of (X, d_X), and the diagonal action of Γ on the product space $X \times X$,

$$\begin{array}{rcl} \Gamma \times X \times X & \longrightarrow & X \times X \\ (\gamma, x, y) & \mapsto & (\gamma x, \gamma y). \end{array}$$

This action is not cocompact but one can try to attach to $X \times X$ a set Ω of ideal boundary points such that the action on $X \times X \cup \Omega$ is cocompact.

The space $(X \times X, d_{X\times X})$, where $d_{X\times X}$ denotes the standard product metric, is a proper metric space and hence, it can be compactified by means of horofunctions; see, for instance, Ballmann–Gromov–Schroeder [2, Sec. 3]. Since it is a CAT(0) space, the horofunctions are, in fact, Busemann functions and the horofunction boundary coincides with the boundary by rays; see Ballmann [1, Prop. 2.5]. The action of Γ extends continuously to an action by homeomorphisms on the boundary of the compactification, but it is not clear if there is a subset of the boundary that can be a good candidate for Ω. For this reason we introduce the maximum metric in $X \times X$, defined by

$$d_{\max}((x, y), (x', y')) = \max\{d_X(x, x'), d_X(y, y')\},$$

for any (x, y), (x', y') in $X \times X$. The space $(X \times X, d_{\max})$ is also a proper metric space and therefore, it can also be compactified by horofunctions. However, it is not a CAT(0) space, since the geodesic segment joining two points is not unique. The group Γ acts on $(X \times X, d_{\max})$ by isometries and the action can be extended to an action by homeomorphisms on the ideal boundary, which we denote by $\partial_\infty^{\max}(X \times X)$. The compactification with respect to the metric $d_{\max}$ turns out to be more adapted to our problem. The main result of this work is the fact that we can find a subset of $\partial_\infty^{\max}(X \times X)$ where the action of Γ is properly discontinuous, and which compactifies the action of Γ on $X \times X$:

Theorem 1.1. *Let X be a proper, non-compact CAT(-1) space and Γ a group of isometries of X acting in a properly discontinuous and cocompact way on X. There exist an open set $\Omega \subset \partial_\infty^{\max}(X \times X)$ such that the diagonal action of Γ on $X \times X \cup \Omega$ is properly discontinuous and cocompact.*

The ideal boundary $\partial_\infty^{\max}(X \times X)$ can be interpreted in terms of the ideal boundary of X, which we denote by $\partial_\infty X$. We will see that $\partial_\infty^{\max}(X \times X)$ splits in a singular part, which is naturally homeomorphic to $\partial_\infty X \sqcup \partial_\infty X$, and a regular part, which is homeomorphic, also in a natural way, to $\partial_\infty X \times \partial_\infty X \times \mathbb{R}$. If we denote by D the diagonal of $\partial_\infty X \times \partial_\infty X$, the set Ω is just the subset of the regular part of the boundary which corresponds, under the homeomorphism, to the set $((\partial_\infty X \times \partial_\infty X) \setminus D) \times \mathbb{R}$.

The set Ω is naturally homeomorphic to the set G of parametrized geodesics in X, equipped with the topology of uniform convergence on compact sets. The identification gives more geometrical insight to the solution of the problem. We consider the diagonal Δ of $X \times X$, and the nearest point retraction that sends each point of $X \times X$ to its nearest point in Δ. This map can be extended in a continuous way to G, by sending each geodesic g to the point $(g(0), g(0))$ in Δ. We use the continuous extension of the nearest point retraction to show that the action of Γ on Ω is properly discontinuous and cocompact.

2. Definitions and notations

In this section we review very briefly the main concepts that appear through the paper.

For us, a parametrized geodesic (or simply a geodesic) is an isometric embedding $g\colon \mathbb{R} \to X$. We call the image of a geodesic a *geodesic line*. A *ray* is an isometric embedding $r\colon [0,\infty) \to X$. And similarly, a *geodesic segment* joining two points x and y is an isometric embedding $xy\colon [a,b] \to X$ such that $xy(a) = x$ and $xy(b) = y$. We will make no distinction between a ray or a geodesic segment and their images.

A *geodesic space* is a metric space such that any two points can be joined by a geodesic segment. A metric space is *proper* if its closed balls are compact. A geodesic metric space is proper if and only if it is complete and locally compact; see Gromov [7, Thm. 1.10].

Let X be a metric space and Δ a geodesic triangle in X. A comparison triangle $\bar{\Delta}_{\mathbb{H}^2}$ in the hyperbolic plane $\mathbb{H}^2$ (or $\bar{\Delta}_{\mathbb{E}^2}$ in the Euclidean plane $\mathbb{E}^2$) is a geodesic triangle in $\mathbb{H}^2$ (or in $\mathbb{E}^2$) with sides of the same length than those of Δ. The space X is $CAT(-1)$ (respectively, $CAT(0)$) if for any triangle Δ, any x, y in Δ and their comparison points $\bar{x}, \bar{y}$ in $\bar{\Delta}_{\mathbb{H}^2}$ (respectively, $\bar{\Delta}_{\mathbb{E}^2}$) satisfy $d(x,y) \leq d(\bar{x},\bar{y})$. A CAT$(-1)$ space is, in particular, CAT(0); see Bridson–Haefliger [4, Part II, Thm. 1.12].

The action of Γ on X is *properly discontinuous* if, for every compact $K \subset X$, the set $K \cap \gamma K$ is non empty for finitely many $\gamma \in \Gamma$. And it is *cocompact* if there exists a compact $K \subset X$ such that $X = \Gamma K$.

Given a metric space (X,d), the *Gromov product* of two points x and y in X with respect to a third point z in X is defined as

$$(x|y)_z = \frac{1}{2}\big(d(z,y) + d(z,x) - d(x,y)\big).$$

In a CAT(-1) space the product can be extended to points at infinity. Observe that a geodesic g determines two rays, one with the same orientation than the geodesic itself and the other with the reversed orientation. The equivalence classes of these rays are the ideal endpoints of the geodesic, we denote them by $g(+\infty)$ and $g(-\infty)$ respectively. The Gromov product of the points $g(+\infty)$ and $g(-\infty)$ with respect to a base point o, is given by

$$(g(+\infty)|g(-\infty))_o = \frac{1}{2}\big(\beta^{g(0)}_{g(+\infty)}(o) + \beta^{g(0)}_{g(-\infty)}(o)\big),$$

where the functions $\beta^{g(0)}_{g(+\infty)}$, $\beta^{g(0)}_{g(-\infty)}$ will be defined in Section 3. The definition of the Gromov product is independent from the parametrization of the geodesic line L with ideal endpoints $g(+\infty)$ and $g(-\infty)$. Observe that if $o \in L$ then it is 0. See, for instance, Ghys–de-la-Harpe [6, Ch. 2] and Bourdon [3, Sect. 2], for more about Gromov products.

3. Horofunction boundary of $(X \times X, d_{\max})$

As we have stated in the introduction, let (X, d_X) be a proper, non-compact CAT(-1) space and consider the product space $X \times X$ together with the maximum metric $d_{\max}$,

$$d_{\max}((x,y),(x',y')) = \max\{d_X(x,x'), d_X(y,y')\}.$$

The compactification via horofunctions of a proper metric space is explained in detail in Ballmann [1, Ch.2] and Bridson–Haefliger [4, Part II, Sect.8]. The idea, which is in fact valid for any complete locally compact metric space, is due to Gromov [2, Sect.3]. It consists on embedding the space, in our case $(X \times X, d_{\max})$, into the space C_* of its continuous functions (with the topology of uniform convergence on compact sets) modulo additive constant, via the map

$$\begin{array}{rcl} \iota\colon X \times X & \longrightarrow & C_* \\ (x,y) & \mapsto & [d_{\max}((x,y),\cdot)], \end{array}$$

assigning to each point in the space the class in C_* of the distance function with respect to this point. The closure of the space, denoted by $\overline{X \times X}^{\max}$, is the closure of $\iota(X \times X)$ in C_*, and the ideal boundary, denoted by $\partial_\infty^{\max}(X \times X)$, is the set $(\overline{X \times X}^{\max}) \setminus \iota(X \times X)$. Both $\overline{X \times X}^{\max}$ and $\partial_\infty^{\max}(X \times X)$ are compact since $X \times X$ is locally compact.

A horofunction is a continuous function such that its class belongs to $\partial_\infty^{\max}(X \times X)$. One can think of horofunctions as limits of normalized distance functions. The level sets of a horofunction are known as horospheres and the sublevel sets as horoballs. Observe that two horofunctions in the same equivalence class differ by a constant and share the same set of horospheres and horoballs.

We can compactify in the same way the original space (X, d_X), since it is also a proper metric space. In this case, since (X, d_X) is CAT(-1), the compactification obtained is homeomorphic to the compactification by rays, where the points at infinity are equivalence classes of rays, and two rays c, c' are in the same class if $d_X(c(t), c'(t)) \leq C$ for all t. The equivalence class of a ray c is denoted by $c(\infty)$, the ideal boundary of X by $\partial_\infty X$, and an arbitrary point in $\partial_\infty X$ by ξ. The horofunctions of a CAT(0) space are in fact Busemann functions and can be written as $\beta_\xi^p(\cdot) = \lim_{t\to\infty} d_X(\cdot, c(t)) - t$, where c is a ray such that $c(0) = p$ and $c(\infty) = \xi \in \partial_\infty X$. Observe that β_ξ^p is the representative of $\xi \in \partial_\infty X$ that satisfies $\beta_\xi^p(p) = 0$. The following two properties about Busemann functions will be used:

(i) If a sequence $\{x_n\}_n$ converges to a point ξ in $\partial_\infty X$, then $\beta_\xi^o(y) = \lim_{n\to\infty} d(x_n, y) - d(x_n, o)$; see Ballmann [1, Prop. 2.5].

(ii) $\beta_\xi^o(y) = -\beta_\xi^y(o)$; see Bourdon [3, Sect.2.1].

Example 3.1. Let X be a complete simply connected n-dimensional Riemannian manifold X of sectional curvature ≤ -1, then $\partial_\infty X \cong S^{n-1}$. In particular, $\partial_\infty \mathbb{H}^n \cong S^{n-1} \cong (T_{x_0}\mathbb{H}^n)^1$ for any $x_0 \in X$.

Now we list all the possible boundary points obtained when compactifying $(X \times X, d_{\max})$. For the calculations, we have chosen a base point $O = (o, o)$ with $o \in X$ and then, as a representative of a class of distance functions, the function $d_{\max}^O((x,y),\cdot) = d_{\max}((x,y),\cdot) - d_{\max}((x,y),(o,o))$. Then, $\{[d_{\max}(P_n,\cdot)]\}_n \to \xi$ if and only if $\{d_{\max}^O(P_n,\cdot)\}_n \to h_\xi^O$, where h_ξ^O is the horofunction in ξ satisfying $h_\xi^O(O) = 0$.

Proposition 3.2. *Every divergent sequence $\{(x_n, y_n)\}_n \subset (X \times X, d_{\max})$ has a subsequence that satisfies, except for permutations of x_n and y_n, one of the following possibilities:*

(i) $d_X(x_n, o)$ is bounded for all n, and $\{y_n\}_n \to \xi' \in \partial_\infty X$;

(ii) $\{x_n\}_n \to \xi \in \partial_\infty X$, $\{y_n\}_n \to \xi' \in \partial_\infty X$, and $d_X(x_n, o) - d_X(y_n, o) \to C$;

(iii) $\{x_n\}_n \to \xi \in \partial_\infty X$, $\{y_n\}_n \to \xi' \in \partial_\infty X$, and $d_X(x_n, o) - d_X(y_n, o) \to +\infty$.

Furthermore, for every $(z,z') \in X \times X$*, the limit* $\lim_{n\to\infty} d^O_{\max}((x_n,y_n),(z,z'))$ *equals* $\beta^o_{\xi'}(z')$ *for the possibility (i),* $\max\{\beta^o_\xi(z), \beta^o_{\xi'}(z')+C\}$ *for (ii), and* $\beta^o_\xi(z)$ *for (iii).*

Proof. We do the proof for case (II), the other two cases can be obtained in a similar fashion. Denote $C_n = d(x_n,o) - d(y_n,o)$ and assume $C_n \geq 0$, to simplify. Then,

$$d_{\max}((x_n,y_n),(x,y)) - d_{\max}((x_n,y_n),(o,o)) = \max\{d(x_n,x) - d(x_n,o), d(y_n,y) - d(y_n,o) - C_n\}.$$

Given $\epsilon > 0$ and $r > 0$, we want to see that there is an N such that, for all $n > N$ and all $(x,y) \in B^{\max}(O,r)$,

$$|\max\{d(x_n,x) - d(x_n,o), d(y_n,y) - d(y_n,o) - C_n\} - \max\{\beta^o_\xi(x), \beta^o_{\xi'}(y) - C\}| < \epsilon,$$

where $B^{\max}(O,r)$ is the ball of center $O = (o,o)$ and radius r in $(X \times X, d_{\max})$. There are four cases to check. We do the case for which the first maximum is $d(x_n,x) - d(x_n,o)$ and the second maximum is $\beta^o_{\xi'}(y) - C$, the other are similar. Because of the definition of Busemann function, there is an N such that, for all $(x,y) \in B^{\max}(O,r)$ and all $n > N$,

$$d(x_n,x) - d(x_n,o) - \beta^o_\xi(x) < \epsilon,$$

$$d(y_n,y) - d(y_n,o) - C_n - (\beta^o_{\xi'}(y') - C) > -\epsilon.$$

On the one hand, we have

$$d(x_n,x) - d(x_n,o) - (\beta^o_{\xi'}(y) - C) = d(x_n,x) - d(x_n,o) - \beta^o_\xi(x) + \beta^o_\xi(x) - (\beta^o_{\xi'}(y) - C) < \epsilon$$

and

$$\begin{aligned} d(x_n,x) - d(x_n,o) - (\beta^o_{\xi'}(y) - C) =& d(x_n,x) - d(x_n,o) - (d(y_n,y) - d(y_n,o) - C_n) + \\ & d(y_n,y) - d(y_n,o) - C_n - (\beta^o_{\xi'}(y') - C) > -\epsilon. \end{aligned}$$ □

The set of boundary points with a representative of the form $\beta^o_\xi(z)$ or $\beta^o_{\xi'}(z')$ is called the *singular part* of the boundary, denoted $\partial^{\max}_\infty(X \times X)_{\text{sing}}$. The rest of the points, i.e., those with a representative of the form $\max\{\beta^o_\xi(z), \beta^o_{\xi'}(z') + C\}$, form the *regular part* of the boundary, denoted by $\partial^{\max}_\infty(X \times X)_{\text{reg}}$.

Observe that, for any constant C', the function $\max\{\beta^o_\xi(z), \beta^o_{\xi'}(z') + C\} + C' = \max\{\beta^o_\xi(z) + C', \beta^o_{\xi'}(z') + C + C'\}$ is in the same class as $\max\{\beta^o_\xi(z), \beta^o_{\xi'}(z') + C\}$. Since two Busemann functions of X associated to the same point $\xi \in \partial_\infty X$ differ by a constant, for each C' we can find points p and p' in X such that $\beta^p_\xi(z) = \beta^o_\xi(z) + C'$ and $\beta^{p'}_{\xi'}(z') = \beta^o_{\xi'}(z') + C + C'$. So, the regular points are in fact the classes modulo constant of the functions $\max\{\beta^p_\xi(z), \beta^{p'}_{\xi'}(z')\}$ for all $p, p' \in X$ and $\xi, \xi' \in \partial_\infty X$.

Proposition 3.3. *There is a natural homeomorphism* $\varphi_{sing} : \partial^{\max}_\infty(X \times X)_{sing} \longrightarrow \partial_\infty X \sqcup \partial_\infty X$ *that consists in associating to a Busemann function that takes values only in the first (second) factor of* $X \times X$*, the same Busemann function viewed as a point of the first (second) factor in* $\partial_\infty X \sqcup \partial_\infty X$*.*

Proof. It follows from the fact that the set of Busemann functions in one factor is naturally identified to the boundary of X. □

The regular part of the boundary can also be identified with a more easy to handle object.

Proposition 3.4. *For each choice of base point $(o, o') \in X \times X$, there is a natural homeomorphism*

$$\begin{aligned} \varphi_{reg} : \partial_\infty^{\max}(X \times X)_{reg} &\longrightarrow \partial_\infty X \times \partial_\infty X \times \mathbb{R} \\ \left[\max\{\beta_\xi^p(z), \beta_{\xi'}^{p'}(z')\}\right] &\mapsto (\xi, \xi', \beta_{\xi'}^{p'}(o) - \beta_\xi^p(o')). \end{aligned} \tag{1}$$

Remark 3.5. Under our choice of base point, the homeomorphism (1) takes the form

$$\begin{aligned} \varphi_{reg} : \partial_\infty^{\max}(X \times X)_{\mathrm{reg}} &\longrightarrow \partial_\infty X \times \partial_\infty X \times \mathbb{R} \\ \max\{\beta_\xi^o(z), \beta_{\xi'}^o(z') + C\} &\mapsto (\xi, \xi', C). \end{aligned} \tag{2}$$

Proof of Proposition 3.4. The map (1) is well defined since two horofunctions in the same class differ by a constant, and $\max\{\beta_\xi^p, \beta_{\xi'}^{p'}\} \neq \max\{\beta_\eta^q, \beta_{\eta'}^{q'}\}$ for $(\xi, \xi') \neq (\eta, \eta')$. To see this, normalize the Busemann functions with respect to the same point; i.e., $\max\{\beta_\xi^p, \beta_{\xi'}^{p'}\} = \max\{\beta_\xi^p, \beta_{\xi'}^p + A\}$ and $\max\{\beta_\eta^q, \beta_{\eta'}^{q'}\} = \max\{\beta_\eta^p + B, \beta_{\eta'}^p + C\}$ for some constants A, B, C. Choose a sequence $z_n \to \xi$ along the geodesic ray joining p and ξ. Then $\beta_\xi^p(z_n) \to -\infty$ and $\beta_\nu^p(z_n) \to +\infty$ for all $\nu \in \partial_\infty X$ such that $\nu \neq \xi$. Using this property one can see that $\xi = \eta$ and $\xi' = \eta'$ if $\max\{\beta_\xi^p, \beta_{\xi'}^{p'}\} = \max\{\beta_\eta^q, \beta_{\eta'}^{q'}\}$.

Now, for each class we choose the representative of the form $\max\{\beta_\xi^o(z), \beta_{\xi'}^o(z') + C\}$ and prove that the map (2) is a homeomorphism. Injectivity is clear. For the exhaustivity, given (ξ, ξ', C), one can see that the sequence $(g(n), g(-n))$ where g is a parameterization of the geodesic line joining ξ' and ξ with $\beta_\xi^o(g(0)) - \beta_{\xi'}^o(g(0)) = C$, has limit $\max\{\beta_\xi^o, \beta_{\xi'}^o + C\}$.

For the continuity, take a sequence $\max\{\beta_{\xi_n}^o, \beta_{\xi'_n}^o + C_n\}$ converging to a point $\max\{\beta_\xi^o, \beta_{\xi'}^o + C\}$. The C_n must be bounded, otherwise the sequence would converge to a Busemann function in one factor. This and the compactness of the set of Busemann fuctions of X, implies that (ξ_n, ξ'_n, C_n) converges to some point (η, η', C'). Since the maximum function is continuous, $\max\{\beta_{\xi_n}^o, \beta_{\xi'_n}^o + C_n\}$ should also converge to $\max\{\beta_\eta^o, \beta_{\eta'}^o + C\}$. Therefore, $\max\{\beta_\xi^o, \beta_{\xi'}^o + C\} = \max\{\beta_\eta^o, \beta_{\eta'}^o + C\}$ and $(\xi, \xi', C) = (\eta, \eta', C')$.

The continuity of the maximum function also assures that the inverse of(2) is continuous. □

Example 3.6. For X a complete simply connected n-dimensional Riemannian manifold of sectional curvature ≤ -1, $\partial_\infty^{\max}(X \times X)_{reg} \cong S^{n-1} \times S^{n-1} \times \mathbb{R}$ and $\partial_\infty^{\max}(X \times X)_{sing} \cong S^{n-1} \sqcup S^{n-1}$. It can be shown that the boundary of $X \times X$ is homeomorphic to a $(2n-1)$-sphere:

$$\partial_\infty^{\max}(X \times X) \cong \mathrm{Join}(S^{n-1}, S^{n-1}) \cong S^{2n-1}.$$

4. An ideal domain for the action of Γ

From now on, Γ will be a discrete and cocompact subgroup of the isometries of X. Recall that we were interested in the diagonal action of Γ on $X \times X$. In this section we look for an open subset of $\partial_\infty^{\max}(X \times X)$, where the action of Γ is good enough. Let D denote the diagonal in $\partial_\infty X \times \partial_\infty X$,

$$D = \{(\xi, \xi) \mid \xi \in \partial_\infty X\} \cong \partial_\infty X,$$

and let $\Lambda_{(x,y)}$ be the limit set of the orbit of the point (x, y) in $X \times X$,

$$\Lambda_{(x,y)} = \overline{\Gamma(x, y)} \cap \partial_\infty^{\max}(X \times X).$$

Notice that $\Lambda_{(x,y)}$ depends on $(x, y) \in X \times X$ since this space is not CAT(-1). Define the limit set of Γ as

$$\Lambda_\Gamma = \bigcup_{(x,y)\in X\times X} \Lambda_{(x,y)}.$$

Proposition 4.1. *The diagonal action of Γ on $X \times X$ satisfies: (i) $\Lambda_\Gamma \subset \partial_\infty^{\max}(X \times X)_{reg}$; and (ii) $\varphi_{reg}(\Lambda_\Gamma) = D \times \mathbb{R}$.*

Proof. First observe that the limit of any sequence $(\gamma_n x, \gamma_n y)$ is in $D\times\mathbb{R}$. Indeed, by the triangle inequality, $|d(\gamma_n x, o)-d(\gamma_n y, o)| \le d(x, y)$ so, the limit is a regular point and, since $d(\gamma_n x, \gamma_n y) = d(x, y)$, if $\gamma_n x \to \xi$ then $\gamma_n y \to \xi$ because X is CAT(-1).

Let us see next that any point (ξ, ξ, C) is in the limit set. Let γ_n be a sequence in Γ such that $\gamma_n \to \xi$. Observe that such a sequence exists since Γ is cocompact and hence its limit set in $\overline{X}$ is the whole $\partial_\infty X$. Let $\xi' \in \partial_\infty X$ be the limit of the sequence γ_n^{-1} and take any point (x, y) satisfying $\beta^o_{\xi'}(x) - \beta^o_{\xi'}(y) = C$. For instance, one can take a point $(g(t), g(t'))$, where g is the ray joining o and ξ', and $t - t' = C$. Then

$$\begin{aligned}\lim_{n\to\infty} d(\gamma_n x, o) - d(\gamma_n y, o) &= \lim_{n\to\infty} d(x, \gamma_n^{-1}o) - d(\gamma_n^{-1}o, o) - (d(y, \gamma_n^{-1}o) - d(\gamma_n^{-1}o, o)) \\ &= \beta^o_{\xi'}(x) - \beta^o_{\xi'}(y) = C.\end{aligned}$$

Hence, the limit of the sequence $(\gamma_n x, \gamma_n y)$ is the point (ξ, ξ, C). □

As a candidate for the domain at infinity, we choose $\Omega \subset \partial_\infty^{\max}(X \times X)_{\rm reg}$ such that $\Omega \cong (\partial_\infty X \times \partial_\infty X \setminus D) \times \mathbb{R}$ under the homeomorphism (1). Observe that we have excluded the whole region $D \times \mathbb{R}$.

Now consider the set G of parameterized geodesics in X. This set is the same as the set of oriented geodesic lines with a distinguished base point and, as we show next, it is in correspondence with the points of Ω. Observe that there is a natural action of Γ on G: an element $\gamma \in \Gamma$ sends a geodesic g to γg.

Lemma 4.2. *The following map is a bijection*

$$\begin{aligned} f\colon G &\longrightarrow \Omega \\ g &\mapsto \lim_{n\to\infty}(g(n), g(-n)). \end{aligned}$$

Proof. First we check that, given a geodesic g in G, the limit of the sequence $\{(g(n), g(-n))\}_n$ belongs to Ω. In order for this limit to be a regular point of the boundary, the sequence $\{(g(n), g(-n))\}_n$ needs to belong to case (ii) of Proposition 3.2 so, the limit of the difference $d_X(g(n), o) - d_X(g(-n), o)$ has to be a real constant. This follows from the next calculation:

$$\begin{aligned}\lim_{n\to\infty}\Big(d_X(g(n), o) - d_X(g(-n), o)\Big) &= \lim_{n\to\infty}\Big(d_X(g(n), o) - d_X(g(n), g(0)) \\ &\quad + d_X(g(-n), g(0)) - d_X(g(-n), o)\Big) \\ &= \beta^{g(0)}_{g(+\infty)}(o) + \beta^{g(0)}_{g(-\infty)}(o) \\ &= \beta^o_{g(-\infty)}(g(0)) - \beta^o_{g(+\infty)}(g(0)) \in \mathbb{R}.\end{aligned}$$

Here, we have used the fact that $d_X(g(n), g(0)) = d_X(g(-n), g(0))$ and the definition of Busemann function. Henceforth, the limit of the sequence $\{(g(n), g(-n))\}_n$ is the point

$$(g(+\infty), g(-\infty), \beta^o_{g(-\infty)}(g(0)) - \beta^o_{g(+\infty)}(g(0)))$$

in $\partial_\infty X \times \partial_\infty X \times \mathbb{R}$ under the homeomorphism (2). Since $g(+\infty) \neq g(-\infty)$, this limit belongs to Ω.

Let us check the injectivity of the map f. Suppose we have two different geodesics g and g' such that $f(g) = f(g')$. Then, $g(+\infty) = g'(+\infty)$ and $g(-\infty) = g'(-\infty)$ and, since given two ideal points in a CAT(-1) space there is a unique geodesic line having them as ideal endpoints (see Bridson–Haefliger [4, Thm. 9.33]), both geodesics must be different parametrizations of the same geodesic line L. Now, observe that the difference $\beta^o_{g(-\infty)}(g(0)) - \beta^o_{g(+\infty)}(g(0))$ can be rewritten using the Gromov product $(g(+\infty)|g(-\infty))_o$ as

$$\begin{aligned}\beta^o_{g(-\infty)}(g(0)) - \beta^o_{g(+\infty)}(g(0)) &= \beta^{g(0)}_{g(+\infty)}(o) - \beta^{g(0)}_{g(-\infty)}(o)) \\ &= 2(\beta^{g(0)}_{g(+\infty)}(o) - (g(+\infty)|g(-\infty))_o) \\ &= -2(\beta^o_{g(+\infty)}(g(0)) + (g(+\infty)|g(-\infty))_o).\end{aligned}$$

Therefore, the two parametrizations satisfy

$$-2(\beta^o_{g(+\infty)}(g(0)) + (g(+\infty)|g(-\infty))_o) = -2(\beta^o_{g'(+\infty)}(g'(0)) + (g'(+\infty)|g'(-\infty))_o)$$

and hence, $\beta^o_\xi(g(0)) = \beta^o_\xi(g'(0))$. Since both $g(0)$ and $g'(0)$ belong to L, we must have $g(0) = g'(0)$ so, both parametrizations of L do coincide.

To finish, let us check the exhaustivity of f. Let (ξ_+, ξ_-, r) be a point in Ω (seen through the homeomorphism (1)). We are looking for a geodesic g such that $\lim_{n\to\infty}\{(g(n), g(-n))\}_n = (\xi_+, \xi_-, r)$. We have already calculated the limit of such a sequence at the beginning of the proof, and we know it is the point $\left(g(+\infty), g(-\infty), \beta^o_{g(-\infty)}(g(0)) - \beta^o_{g(+\infty)}(g(0))\right)$. So, we look for a geodesic such that $g(+\infty) = \xi_+$, $g(-\infty) = \xi_-$, and $\beta^o_{g(-\infty)}(g(0)) - \beta^o_{g(+\infty)}(g(0)) = r$. Let L be the geodesic line with ideal endpoints ξ_+ and ξ_-. Consider a parametrization $g(t)$ of L. Since the functions $\beta^o_{\xi_+}(g(t))$ and $\beta^o_{\xi_-}(g(t))$ are lineal with slope ± 1, respectively, there is a unique point p in L satisfying $\beta^o_{\xi_-}(p) - \beta^o_{\xi_+}(p) = r$. The parametrization $g'(t)$ of L such that $g'(0) = p$ is the one we are looking for, since it satisfies $g'(+\infty) = \xi_+$, $g'(-\infty) = \xi_-$, and $\beta^o_{g'(-\infty)}(g'(0)) - \beta^o_{g'(+\infty)}(g'(0))) = r$. □

Remark 4.3. Observe that what we have checked in the proof of Lemma 4.2 is, in fact, the bijectivity of the map $\varphi_{reg} \circ f$.

We consider in G the topology of uniform convergence on compact sets.

Theorem 4.4. *The following map is an equivariant homeomorphism*

$$\begin{aligned} f\colon G &\longrightarrow \Omega \\ g &\mapsto \lim_{n\to\infty}(g(n), g(-n)).\end{aligned}$$

Proof. To see that f is a homeomorphism, consider the map

$$\begin{aligned} f'\colon G &\longrightarrow ((\partial_\infty X \times \partial_\infty X)\setminus D)\times\mathbb{R} \\ g &\mapsto \left(g(+\infty),\, g(-\infty),\, \beta^{g(0)}_{g(+\infty)}(o) - \beta^{g(0)}_{g(-\infty)}(o)\right),\end{aligned}$$

which is just $\varphi_{reg} \circ f$ and which, as we have seen along the proof of Lemma 4.2, it is a bijection. Let ϕ be the Hopf parametrization

$$\begin{aligned} \phi\colon G &\longrightarrow ((\partial_\infty X \times \partial_\infty X)\setminus D)\times\mathbb{R}) \\ g &\mapsto \left(g(+\infty), g(-\infty), \beta^{g(0)}_{g(+\infty)}(o)\right),\end{aligned}$$

which is a homeomorphism (see Bourdon [3, Sect. 2.9]), and let h be the map

$$\begin{array}{rcl} h\colon ((\partial_\infty X \times \partial_\infty X) \setminus D) \times \mathbb{R}) & \longrightarrow & ((\partial_\infty X \times \partial_\infty X) \setminus D) \times \mathbb{R}) \\ (\xi_+, \xi_-, r) & \mapsto & (\xi_+, \xi_-, 2(r - (\xi_+|\xi_-)_O)), \end{array}$$

which is also a homeomorphism. Now, the following diagram commutes

$$\begin{array}{ccc} G & \xrightarrow{f'} & ((\partial_\infty X \times \partial_\infty X) \setminus D) \times \mathbb{R}). \\ \downarrow \phi & \nearrow h & \\ ((\partial_\infty X \times \partial_\infty X) \setminus D) \times \mathbb{R}) & & \end{array}$$

Indeed, for any $g \in G$,

$$\begin{aligned} h \circ \phi(g) &= h((g(+\infty), g(-\infty), \beta^{g(0)}_{g(+\infty)}(o))) \\ &= (g(+\infty), g(-\infty), 2(\beta^{g(0)}_{g(+\infty)}(o) - (g(+\infty)|g(-\infty))_o) \\ &= (g(+\infty), g(-\infty), \beta^{g(0)}_{g(+\infty)}(o) - \beta^{g(0)}_{g(-\infty)}(o)) \\ &= f'(g). \end{aligned}$$

Since both ϕ and g are homeomorphisms, f' is a homeomorphism. And since $f' = \varphi_{reg} \circ f$ and ϕ_{reg} is a homeomorphism, our map f is a homeomorphism too.

To finish, observe that the map f is equivariant since, for every $\gamma \in \Gamma$ and $g \in G$,

$$f(\gamma g) = \lim_{n\to\infty} (\gamma g(n), \gamma g(-n)) = \gamma \lim_{n\to\infty} (g(n), g(-n)) = \gamma f(g). \qquad \square$$

The next proposition is a consequence of Theorem 4.4 and the properties of G. It will also be a consequence of Theorem 5.9.

Proposition 4.5. *The action of Γ on Ω is cocompact and properly discontinuous.* □

5. Compactness of $(X \times X \cup \Omega)/\Gamma$

We define a topology in $X \times X \cup G$ in the following way. We keep the same topology in $X \times X$ and G, and we say a sequence $\{(x_n, y_n)\}_n$ in $X \times X$ converges to a point g in G if and only if the following three conditions hold: (i) $\{x_n\}_n \to g(+\infty)$, (ii) $\{y_n\}_n \to g(-\infty)$; and (iii) $d_X(x_n, o) - d_X(y_n, o) \to \beta^o_{g(-\infty)}(g(0)) - \beta^o_{g(+\infty)}(g(0))$. With this topology, $X \times X \cup G$ and $X \times X \cup \Omega$ are homeomorphic.

Now, consider the diagonal in $X \times X$, $\Delta = \{(x, x) \mid x \in X\}$, and let $\rho\colon X \times X \to \Delta$ be the map that sends each point in $X \times X$ to its nearest point in Δ with respect to the metric $d_{\max}$.

Lemma 5.1. *The fibre $\rho^{-1}(a, a)$ is the set of point $(x, y) \in X \times X$ such that a is the midpoint of the segment xy.* □

Definition 5.2. We extend the projection ρ to a map $\tilde{\rho}\colon X \times X \cup G \to \Delta$ as follows: if $p = (x, y) \in X \times X$, then $\tilde{\rho}(p) = \rho(p)$; if $g \in G$, then $\tilde{\rho}(g) = (g(0), g(0))$.

Next, we prove that this extension is continuous. Before, we need a couple of auxiliary lemmas.

Lemma 5.3. *In a proper CAT(-1) space, $\lim_{i,j}(x_i|y_j)_o = (\xi|\xi')_o$ for any sequences $x_i \to \xi$, $y_j \to \xi'$; see Buyalo–Schroeder [5, Prop 3.4.2]. Moreover, $(\xi|\xi')_o = +\infty$ if and only if $\xi = \xi'$; see Bridson–Haefliger [4, Ch.III.H. Rmk. 3.17].* □

Lemma 5.4. *Given a sequence of geodesic segments $g_n\colon [a_n, b_n] \to X$ such that $g_n(a_n) \to \xi$, $g_n(b_n) \to \xi'$, and $g_n(0) \to m \in X$, there exists a convergent subsequence to a geodesic g satisfying $g(+\infty) = \xi$, $g(-\infty) = \xi'$, and $g(0) = m$.*

Proof. The fact that the sequence $\{g_n\}_n$ converges to a geodesic g such that $g(0) = m$ is a consequence of Arzela–Ascoli theorem for proper metric spaces; see Papadopoulos [8, Thm. 1.4.9]. Now, for any $t \in \mathbb{R}$, observe that $(g_n(a_n)|g_n(b_n))_{g_n(t)} = 0$ since, for all n, $g_n(t)$ is a point of the segment g_n, and $\lim_n(g_n(a_n)|g_n(b_n))_{g_n(t)} = (\xi|\xi')_{g(t)}$ by the continuity of the Gromov product. Therefore, for all t, $(\xi|\xi')_{g(t)} = 0$ and $g(t)$ belongs to the line joining ξ and ξ'. Since this line is unique, $g(+\infty) = \xi$ and $g(-\infty) = \xi'$. □

Proposition 5.5. *The map $\tilde{\rho}\colon X \times X \cup G \to \Delta$ is continuous.*

Proof. The restrictions of $\tilde{\rho}$ to $X \times X$ and to G are continuous. Let $\{(x_n, y_n)\}_n$ be a sequence in $X \times X$ converging to a geodesic g in G so, $x_n \to g(+\infty)$, $y_n \to g(-\infty)$, and $d(x_n, o) - d(y_n, o) \to C$. First, we will prove that the geodesic segments $x_n y_n$ converge to a parameterization of the geodesic line L with ideal endpoints $\xi = g(+\infty)$ and $\xi' = g(-\infty)$; then we will see that this parameterization is precisely g.

For each pair (x_n, y_n) let m_n be the middle point of the segment $x_n y_n$. The points m_n lie in a compact set. Indeed, suppose $d(o, m_n) \to +\infty$ so, $m_n \to \eta \in \partial_\infty X$. Using the definition of Gromov product,

$$(x_n|m_n)_o = \frac{1}{2}\big((x_n|y_n)_o + d(m_n, o) + \frac{1}{2}(d(x_n, o) - d(y_n, o))\big).$$

Since $\xi \neq \xi'$, by Lemma 5.3 $\lim_n(x_n|y_n)_o$ is bounded. By hypothesis, $d(x_n, o) - d(y_n, o)$ is also bounded and $d(m_n, o) \to +\infty$. Therefore, $\lim_n(x_n|m_n)_o = +\infty$ and, by Lemma 5.3 again, $\eta = \xi$. Similarly, one could find that $\eta = \xi'$, so $\xi = \xi'$ and arrive to a contradiction. Therefore, $m_n \to m$ for some $m \in X$ and, by Lemma 5.4, m must be a point in L.

Now we have, on one hand,

$$d_X(m_n, y_n) - d_X(o, y_n) - (d_X(m_n, x_n) - d_X(o, x_n)) = d_X(o, x_n) - d_X(o, y_n) \to C.$$

On the other hand, since $d_X(\cdot, y_n) - d_X(o, y_n) \to \beta^o_{\xi'}(\cdot)$ and $d_X(\cdot, x_n) - d_X(o, x_n) \to \beta^o_{\xi}(\cdot)$ uniformly on compact sets, we have

$$d_X(m_n, y_n) - d_X(o, y_n) - (d_X(m_n, x_n) - d_X(o, x_n)) \to \beta^o_{\xi'}(m) - \beta^o_{\xi}(m)$$

so, $\beta^o_{\xi'}(m) - \beta^o_{\xi}(m) = C$. But the only point in L satisfying this equation is precisely $g(0)$. Hence $m = g(0)$.

Therefore, $\tilde{\rho}(x_n, y_n) = (m_n, m_n) \to (m, m) = (g(0), g(0)) = \tilde{\rho}(g)$ and the map $\tilde{\rho}$ is everywhere continuous. □

Corollary 5.6. *The fibre $\tilde{\rho}^{-1}(x, x)$ restricted to G is the set $G_x = \{\sigma \in G \mid \sigma(0) = x\} \subset G$.* □

Example 5.7. In a Riemannian manifold X of dimension n and sectional curvature ≤ -1, G_x is identified with the unitary tangent at x, $(T_xX)^1 \cong S^{n-1}$.

Observe that, since $X \times X \cup G$ and $X \times X \cup \Omega$ are homeomorphic, we also have a continuous projection from $X \times X \cup \Omega$ to Δ, which we also call $\tilde{\rho}$.

Now, let $K \subset X$ be a compact in X such that $K/\Gamma \cong X/\Gamma$ and consider $K_\Delta = \{(x,x) \in \Delta \mid x \in K\}$.

Lemma 5.8. *The set $\tilde{\rho}^{-1}(K_\Delta)$ is compact.*

Proof. Consider a sequence $\{(x_n, y_n)\}_n$ in $\tilde{\rho}^{-1}(K_\Delta)$. Then, since K_Δ is compact, $\{\tilde{\rho}((x_n, y_n))\}_n$ has a convergent subsequence $\{(m_n, m_n)\}_n$ in K_Δ. Take a sequence of points $\{(x_n, y_n)\}_n$ that project to this convergent subsequence. They have a subsequence $\{(x'_n, y'_n)\}_n$ converging in $\overline{X \times X}^{\max}$. If the limit point of this subsequence is in $X \times X \cup \Omega$ we are done. If not, either the limit is in $X \times X \cup \Delta$ or $|\{d(x'_n, o) - d(y'_n, o)\}_n|$ is unbounded for all n. For the first case, observe that, for every n, $(x'_n|y'_n)_o \leq d_X(m_n, o) < C$, but $\{x'_n\}_n$ and $\{y'_n\}_n$ have the same limit if and only if $(x'_n|y'_n)_o \to \infty$ by Lemma 5.3, hence, this case is not possible. The second case is not possible either since $|\{d(x'_n, o) - d(y'_n, o)\}_n|$ unbounded implies $\{d(m_n, o)\}_n$ unbounded.

Now, consider a sequence of the form $\{g_n\}_n$ in $\tilde{\rho}^{-1}(K_\Delta)$. The geodesics in the sequence satisfy $g_n(0) \in K$ for all n. Since the set of geodesics going through a compact set is compact, $\{g_n\}$ has a convergent subsequence. □

Theorem 5.9. *The action of Γ on $X \times X \cup \Omega$ is properly discontinuous and cocompact.*

Proof. In order to see that Γ acts properly discontinuously, take $K \subset (X \times X \cup \Omega)$ any compact subset, and let $\gamma \in \Gamma$ be such that $\gamma K \cap K \neq \emptyset$. Then, $\tilde{\rho}(K) \cap \tilde{\rho}(\gamma K) = \tilde{\rho}(K) \cap \gamma\tilde{\rho}(K) \neq \emptyset$. Since $\tilde{\rho}$ is continuous, $\tilde{\rho}(K)$ is compact and, since the action of Γ on Δ is properly discontinuous, $\tilde{\rho}(K) \cap \gamma\tilde{\rho}(K) \neq \emptyset$ only for a finite number of elements $\gamma \in \Gamma$. Therefore $\gamma K \cap K \neq \emptyset$ only for a finite number of $\gamma \in \Gamma$.

For the cocompactness observe that, by Lemma 5.1 and Corollary 5.6, $(X \times X \cup \Omega)/\Gamma \cong (X \times X \cup G)/\Gamma = \tilde{\rho}^{-1}(K_\Delta)/\Gamma$ which is compact by Lemma 5.8. □

References

[1] W. Ballmann, "Lectures on spaces of nonpositive curvature", *DMV Seminar* **25**, Birkhäuser Verlag, Basel, 1995. (With an appendix by M. Brin).

[2] W. Ballmann, M. Gromov, and V. Schroeder, "Manifolds of nonpositive curvature", *Progress in Mathematics* **61**, Birkhäuser, Boston, MA, 1985.

[3] M. Bourdon, "Structure conforme au bord et flot géodésique d'un CAT(−1)-espace", *Enseign. Math.* **41**(1-2) (1995), 63–102.

[4] M. Bridson and A. Haefliger, "Metric spaces of non-positive curvature", *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **319**, Springer-Verlag, Berlin, 1999.

[5] S. Buyalo and V. Schroeder, "Elements of asymptotic geometry", *EMS Monographs in Mathematics*, European Mathematical Society (EMS), Zürich, 2007.

[6] É. Ghys and P. de la Harpe, eds., "Sur les groupes hyperboliques d'après Mikhael Gromov", *Progress in Mathematics* **83**, Birkhäuser Boston, MA, 1990.

[7] M. Gromov, "Metric structures for Riemannian and non-Riemannian spaces", Modern Birkhäuser Classics, Birkhäuser, Boston, MA, 2007.

[8] A. Papadopoulos, "Metric spaces, convexity and non-positive curvature", *IRMA Lectures in Mathematics and Theoretical Physics* **6**, European Mathematical Society (EMS), Zürich, 2014.

Reports@SCM
Volume 3, number 1, 2017

Table of Contents

THE WAVE EQUATION FOR STIFF STRINGS AND PIANO TUNING
Xavier Gracia and Tomas Sanz-Perela 1

EXPLICIT BOUNDS FOR GROWTH OF SETS IN NON-ABELIAN GROUPS
Alberto Espuny Díaz 17

COMPACTIFICATION OF A DIAGONAL ACTION ON THE PRODUCT OF CAT (-1) SPACES
Teresa García 27