New Trends in Analysis and Geometry

Edited by **Ali Hussain Alkhaldi**, **Mohammed Kbiri Alaoui** and **Mohamed Amine Khamsi**

New Trends in Analysis and Geometry

New Trends in Analysis and Geometry

Edited by

Ali Hussain Alkhaldi,
Mohammed Kbiri Alaoui
and Mohamed Amine Khamsi

Cambridge
Scholars
Publishing

New Trends in Analysis and Geometry

Edited by Ali Hussain Alkhaldi, Mohammed Kbiri Alaoui
and Mohamed Amine Khamsi

This book first published 2020

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-5275-4563-6
ISBN (13): 978-1-5275-4563-2

Contents

Preface

The present edited collection is an attempt to report recent progress on a number of interconnected mathematical results in analysis, partial differential equations, geometry and applications. This book is composed of eleven chapters on various theoretical topics and its applications, among others, partial differential equations, modular function spaces, Musielak-Orlicz spaces, Sobolev embeddidngs, fixed point theory on graphs, modular convexity, hyperbolic geometry, hyperbolic kinematics and variational inequalities on manifolds.

Chapters are written by different authors who present their latest contributions in a style aimed at a wide audience, ranging from beginning students to specialists. Each Chapter has also been thought of as a source of examples, references, open questions, and occasional new approaches to traditional topics, aiming at opening new directions of research and shading new light on long-standing problems. Graduate students will find in this monograph a wide variety of topics among which to select a mathematical field to focus their interest. It is the hope of the authors that the monograph will be a useful tool to mathematicians interested in the general aspects of the expounded topics, as well as to specialists seeking to explore the deeper aspects of the presented themes.

We collectively thank our many friends and colleagues who, through their encouragement and help, influenced the development of this book. In particular, we are especially grateful to the Rector of King Khalid University, Prof. Faleh Al-Solamy, for his support in the organization of the first International Conference of Mathematics and its Applications, that took place in March 2018, as well as to the Chairman of the Department of Mathematics of said Institution, Dr. Ibrahim Almanjahie. Most of the authors who contributed in this book attended this conference and readily agreed to be part of this project. Finally, we would like to extend our special thanks to Helen Edwards (Commissioning Editor, Cambridge Scholars Publishing) for her interest in publishing this book.

Abha, Saudi Arabia	Ali Hussain Alkhaldi
Abha, Saudi Arabia	Mohammed Kbiri Alaoui
El Paso, USA	Mohamed Amine Khamsi
December 2019	

Chapter 1

A voting system with a fixed point

C. de Kerchove
Department of Mathematics, Namur University, Namur, Belgium,
E-mail: cristobald.dekerchove@unamur.be

P. Van Dooren
Department of Mathematical Engineering, Université catholique de Louvain, Louvain-la-Neuve, Belgium,
E-mail: paul.vandooren@uclouvain.be

A voting system is presented that is based on an iterative procedure that converges to a unique fixed point. The votes are casted by m raters regarding the reputation of n items, are organized as a $m \times n$ voting matrix X, which is possibly sparse when each rater does not evaluate all items. From this matrix X, a unique rating of the considered items is finally obtained via an iterative procedure which updates the reputations of the n items as well as that of the m raters. For any rating matrix the proposed method converges linearly to the unique vector of reputations. Some applications of this voting system will be presented.

1.1 Introduction

One of the most influential changes in our generation is undoubtedly the internet and its use for the communication of information. It is via the internet that most individuals are looking for information or are provided with information when signing up for some *information channel*. Typical tools for the finding of information are search engines, such as Google and Yahoo, whereas Facebook, CNN, various newspapers and several entertainment channels are examples of information channels. Many of these channels use databases of *opinions* gathered from a pool of arbitrary users and are based on votes that are ultimately used to rate the objects that users are interested in. These objects could be books, hotels, restaurants, movies or touristic locations: one may e.g. refer to Amazon to find out about popular books, Booking or Trivago to inquire about hotels, Movielens to look into movies, or Tripadvisor to investigate about various touristic hotspots. The list of such interactive sites is continuously growing. These actions can all be interpreted as a form of voting; but the honesty or reliability of the *raters* cannot always be verified.

A rater on the Movielens database may give random ratings to movies he/he has not even seen, or a dishonest voter on Tripadvisor may post biased opinions just to favor his or her ”friends”.

Clearly these websites can only benefit from their rating system being as trustworthy as possible. At first sight, this looks like an impossible task since one cannot verify the honesty of all raters.

However, the coherence of the ratings provided by isolated raters can at least be checked against that of the *average* opinion. This is the approach that we propose here. We will actually try to achieve two simultaneous goals. The first such objective is to assign a *reputation* to each of the evaluated items, and the second one is to assign a grading of *reliability* or *trust* to each of the raters who evaluated the items.

We establish a clear difference between the *reputation* of an item, that is, between what is generally said or believed about the quality of characteristics of an item, and the *reliability* of a rater, which is our expectation that the rater gives a fair or relevant evaluation to the item in question.

We illustrate the need of such a voting system by recalling a voting scandal at the 2002- Winter Olympics. To this effect, we quote the following account from Wikipedia [8] (see also [1, 7]):

> At the 2002 Winter Olympics held in Salt Lake City, allegations arose that the pairs' figure skating competition had been fixed. The controversy led to two pairs teams receiving gold medals: the original winners Elena Berezhnaya and Anton Sikharulidze of Russia and original silver-medalists Jamie Salé and David Pelletier of Canada. The scandal was one of the causes for the revamp of scoring in figure skating to the new ISU Judging System. [...]. The ISU Judging System replaced the previous 6.0 system in 2004. This new system was created in response to the 2002 Winter Olympics figure skating scandal, in an attempt to make the scoring system more objective and less vulnerable to abuse.

During this event, two of the judges favored the Russian team with scores that were quite different from the averages scores of the remaining voters. The press and the public immediately, openly criticized the results of the voting, and the reaction was so strong that the president of the International Olympic Committee announced the decision that, though Russia would be allowed to keep the Gold Medal, the Canadian team, to whom the silver medal had been awarded as a result of the competition, would also get a gold medal.

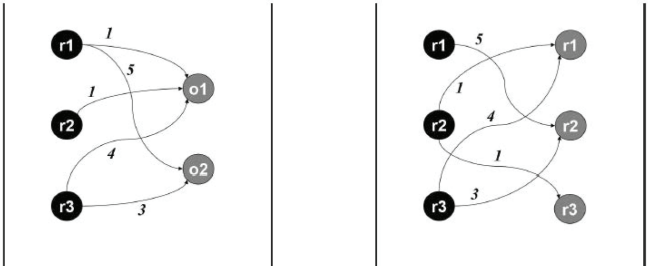
The new ISU system is actually quite complex and tailored to the the specific case of figure skating events. We present a different voting system, based on a fixed point iteration and that turns out to be the solution of an optimization problem. We delve into the advantages of this system and illustrate its use in several applications.

1.2 Reputation and trust

Various measures of reputation have been proposed in recent years under the names of reputation, voting, ranking or trust systems, among others and they deal with a number of contexts ranging from the classification of football teams to the reliability of each individual in peer to peer systems. Surprisingly enough, the most used method for measuring reputation on the Web, amounts simply to averaging the votes. In that case, the reputation is, for instance, the average of scores represented by stars in YouTube, or the percentage of positive transactions in eBay. Such a method, then, implicitly trusts evenly each rater of the system. Besides this method, many other algorithms exploit the structure of networks generated by the votes: raters and evaluated items are nodes connected by votes, as illustrated in Figure 1.1.

There are many different ways of defining trust or reputation and each of them has advantages and shortcomings. We refer here to [2] for a short survey on these ideas and the principles they are based on. Obviously the choice of a specific reputation system depends on subjective properties that are just accepted. For example, in the averaging method mentioned above, it is tacitly agreed that every rater is taken into account in the same manner, whereas the PageRank algorithm is based on the principle that a random walk over the network is a good model for the navigation of a web surfer. The fundamental assumption underlying the method we present here is the following:

Raters diverging often from other raters' opinions ought to be taken less



$$\begin{bmatrix} 1 & 1 & 4 \\ 5 & . & 3 \end{bmatrix} \qquad \begin{bmatrix} . & 1 & 4 \\ 5 & . & 1 \\ . & 1 & . \end{bmatrix}$$

FIGURE 1.1: Network and matrix of votes from raters to objects and from raters to other raters

into account than the remaining raters.

This principle is the basis of our filtering process and implies that all votes are taken into account, but with a continuous validation scale, in contrast for instance, to the direct deletion of outliers. Moreover, the weight of each rater depends on the distance between his/her votes and the reputation of the objects he/she evaluates: typically weights of random raters and outliers decrease during the iterative filtering. The main criticism to be raised against this method is that it discriminates marginal evaluators.

Votes, raters and objects can appear, disappear or change, making the system dynamical. This is for example the case when we consider a stream of news like in [19]: news sources and articles are ranked according to their publications over time. Nowadays, most sites driven by raters involve dynamical opinions. For instance, the blogs, the site Digg and the site Flickr are good places to exchange and discuss ideas, remarks and votes about various topics ranging from political election to photos and videos. We will see that our proposed system allows for the consideration of evolving voting matrices and provides time varying reputations.

1.3 Weighted averages of votes

A natural way of tackling the problem of unreliable or unfair raters in reputation systems is to assign a weight to the evaluations of the raters.

Hence the range of weights corresponds to a continuous scale of validation of the votes. These weights change via an iterative refinement that is guaranteed to converge to a reputation score for every evaluated item and to a reliability score for every rater. At each step the reliability of a rater is calculated according to some distance between his/her given evaluations and the reputations of the items he/she evaluates. This distance is interpreted as the *belief-divergence*. Typically, a rater diverging too much from the group should be distrusted to some extent. The same definition of distance appears in [5, 6, 9] and is used for the same purpose. The strength of the reputation system we will describe here is that it can be applied to any static network of raters and items and that it converges to a unique fixed point. Moreover, our reputation system can also be extended to dynamical systems with time-dependent votes.

We describe our approach for a static system with m raters and n objects to be rated. The entry X_{ij} represents the vote of rater $j \in \{1,\ldots,m\}$ for item $i \in \{1,\ldots,n\}$, the matrix $X \in [a,b]^{n\times m}$ is the *voting matrix*. Each vote is in the positive real interval $[a,b]$, and the vector $\vec{x}_j$, the j-th column of X, is the vector of votes of rater j.

The graph of votes and raters can be represented by an adjacency matrix $A \in \{0,1\}^{n\times m}$ where $A_{ij} = 1$ if object i is evaluated by rater j, and is equal to 0 otherwise. For the sake of simplicity, it is first assumed that every rater evaluates all items. Then the item's reputation vector $\vec{r}$ is defined as the weighted sum of the votes

$$\vec{r} := X\frac{\vec{w}}{\vec{1}_m^T\vec{w}}, \tag{1.1}$$

where $\vec{1}_m$ is the vector of all ones. Since this is a convex combination of the vectors $\{\vec{x}_1,\ldots,\vec{x}_m\}$, it follows that $\vec{r} \in [a,b]^n$. The rater's weight vector $\vec{w}$ depends on the discrepancy with the other votes, interpreted as belief-divergence, which we define as

$$\vec{w} := G(\vec{r}) := \vec{1}_m - k\vec{d}, \quad \text{where} \quad \vec{d} := \frac{1}{n}\begin{bmatrix} \|\vec{x}_1 - \vec{r}\|_2^2 \\ \vdots \\ \|\vec{x}_m - \vec{r}\|_2^2 \end{bmatrix}, \tag{1.2}$$

and k is a positive parameter. Clearly, as k tends to zero, $\vec{w}$ tends to $\vec{1}_m$, and $\vec{r}$ tends to the average of the votes. Increasing k corresponds to more stringent discrimination toward outliers.

We proved in [2] that for $k < 1/b$ the vector $\vec{w}$ is always positive, and that this implies then that there exists a unique pair of vectors $\vec{r}$, $\vec{w}(\vec{r})$ satisfying both (nonlinear) equalities (1.1) and (1.2). Moreover the nonlinear iteration given by

$$\vec{w}^0 := \vec{1}_m, \quad \vec{r}^{t+1} := X\vec{w}^t/(\vec{1}_m^T\vec{w}^t),\ \vec{w}^{t+1} := G(\vec{r}^{t+1}), \quad \text{for} \quad t = 0,1,2,\ldots$$

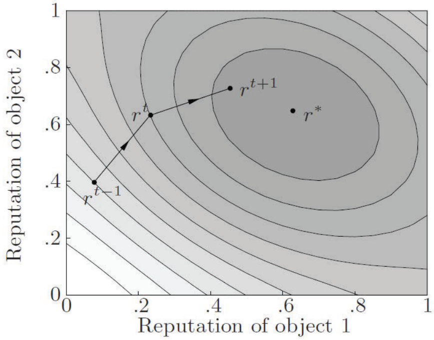


FIGURE 1.2: Representation of steps $\vec{r}^t$ of the nonlinear iteration in the unit box $[0,1]\times[0,1]$. The sequence $E(\vec{r}^t)$ decreases with t and converges to $E(\vec{r}^*)$.

converges to a unique fixed point $[\vec{w}^*, \vec{r}^*]$ that satisfies equations (1.1) and (1.2).

The uniqueness of the solution is established via the definition of the cost function $E(\vec{x}) = -\frac{1}{2k}\vec{w}(\vec{x})^T\vec{w}(\vec{x})$ that is minimized for $\vec{x}$ equal to the reputation vector $\vec{r}^*$. Moreover each step given by the nonlinear iteration resulting from formulas (1.1) and (1.2) corresponds to the steepest descent direction

$$\nabla_{\vec{r}}E(\vec{r}^t) = -\frac{1}{\alpha^t}(\vec{r}^{t+1} - \vec{r}^t),$$

of the cost function with step size $\alpha^t := \frac{n}{2\vec{1}_m^T\vec{w}^t} \geq \frac{n}{2m}$. It is shown in [2] that the corresponding steepest descent iteration

$$\vec{r}^{t+1} := \vec{r}^t - \alpha^t\nabla_{\vec{r}}E(\vec{r}^t)$$

converges to the fixed point $\vec{r}^*$ of the iteration and that $\vec{r}^*$ is the unique minimum of the cost function $E(\vec{x})$ for $\vec{x}$ in the hypercube $[a,b]^m$. This is illustrated in Figure 1.2 when only two objects are considered, that is, in the particular case when $m = 2$, and where the voting interval is $[0,1]$.

Therefore the solution should not only be viewed as the fixed point of a nonlinear iteration, but it can also be interpreted as the minimizer of $-\vec{w}^T\vec{w}$ (and hence, as the maximizer of the 2-norm of $\vec{w}$). It therefore remains to be shown that there is a *unique minimizer* $\vec{r}^*$ in the imposed constrained set $\vec{r} \in \mathscr{H} := [a,b]^n$. This was again analyzed in [2], where it is shown that the function $E(\vec{r})$ is convex in $\mathscr{H}$, *provided* that the positive parameter

k is smaller than the value $\frac{1}{b}$. This indeed guarantees that the weighting vector $\vec{w}$ is strictly positive, from which it then also follows that the function $E(\vec{r})$ is strictly convex in $\mathscr{H}$. In Figure 1.3 we show (for a 2-dimensional vector $\vec{r}$ and the set $\mathscr{H} = [0,1]^2$), a plot of four possible configurations of the function $E(\vec{r})$ depending on the value of k. In (a) we chose $k < 1$ which implies that the energy function is convex, in (b), (c) and (d), we gradually increase k which introduces saddle points and local minima and maxima, and eventually makes the function concave.

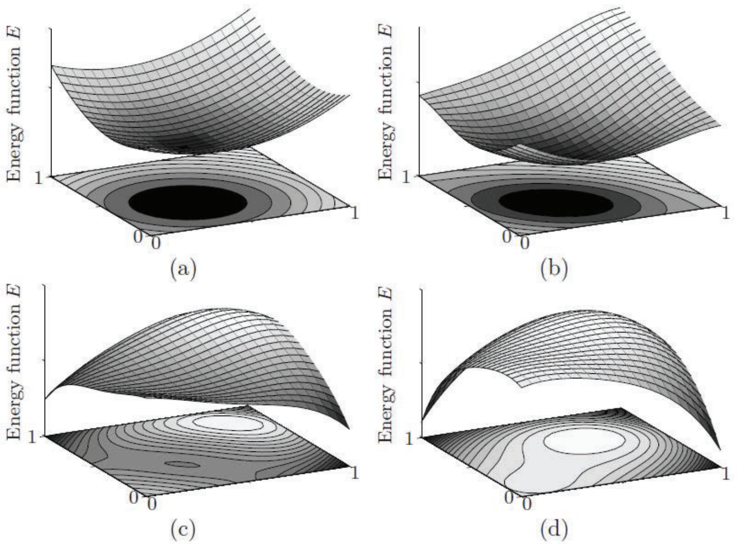


FIGURE 1.3: The function $E(\vec{r})$ for increasing values of k. In (a), $k < 1$ and there is a unique minimum. In (b), (c) and (d), k is increased and saddle points and local maxima and minima are observed.

1.4 Extensions

Here we briefly mention three different extensions that have been analyzed in [2].

Sparse votes

In reputation systems like Amazon, Tripadvisor and the Movielens

Database, the votes are clearly sparse since most raters do not give their opinion about all objects. If we assume that the entry X_{ij} of the voting matrix X is set to 0 when the entry A_{ij} of the adjacency matrix is 0, then $A \circ X = X$, where the symbol "$\circ$" is used to denote the elementwise product of two matrices of the same dimensions (also called the *Hadamard* product). This property turns out to be crucial for the derivation of the sparse voting scheme explained in [2] and is based on a fixed point idea. The formulas (1.1) and (1.2) are now replaced with the following expressions :

$$\vec{r} = \frac{[X\vec{w}]}{[A\vec{w}]} \tag{1.3}$$

$$\vec{w} = G(\vec{r}) := \vec{1}_m - k\vec{d}, \quad \text{where} \quad \vec{d} = \begin{bmatrix} \frac{1}{n_1}\|\vec{x}_1 - \vec{a}_1 \circ \vec{r}\|_2^2 \\ \vdots \\ \frac{1}{n_m}\|\vec{x}_m - \vec{a}_m \circ \vec{r}\|_2^2 \end{bmatrix}, \tag{1.4}$$

where $\vec{a}_j$ is the j-th column of the adjacency matrix A, and n_j is the j-th element of the vector $\vec{n}$ containing the numbers of votes for each item, i.e., $\vec{n} = A^T\vec{1}_n$, whereas the scalar n denotes the total number of items. We point out that $\frac{[\cdot]}{[\cdot]}$ is the componentwise division of two vectors of the same dimension, which implies that every item is evaluated by at least one rater with nonzero weight. It is easy to verify that when the matrix A is the matrix of all ones, one retrieves the formulas for the dense voting matrix. Moreover, the nonlinear iteration

$$\vec{w}^0 := \vec{1}_m, \quad \vec{r}^{t+1} := \frac{[X\vec{w}^t]}{[A\vec{w}^t]}, \; \vec{w}^{t+1} := G(\vec{r}^{t+1}), \quad \text{for} \quad t = 0, 1, 2, \dots$$

converges to a unique fixed point $[\vec{w}^*, \vec{r}^*]$ that satisfies equations (1.3) and (1.4). We refer again to [2] for the proofs of these assertions.

Time-varying votes

This extension makes it possible to also consider dynamical votes where the rating matrix changes over time. Clearly, votes and web users are constantly evolving on the Web, therefore it appears necessary to develop also dynamical reputation systems. In this scenario, we consider discrete sequences of votes and adjacency matrices such as

$$\{X^s,\ s = 1, 2, 3, \dots\} \quad \text{and} \quad \{A^s,\ s = 1, 2, 3, \dots\},$$

that evolve on a discrete time axis, and we again assume that $A^s \circ X^s = X^s$ for every s. The iteration then becomes

$$\vec{w}^0 := \vec{1}_m, \quad \vec{r}^{t+1} := \frac{[X^{s+1}\vec{w}^t]}{[A^{s+1}\vec{w}^t]}, \; \vec{w}^{t+1} := G_{s+1}(\vec{r}^{t+1}), \quad \text{for} \quad t = 0, 1, 2, \dots,$$

where

$$G_{s+1}(\vec{r}^{t+1}) := \vec{1}_m - k \begin{bmatrix} \frac{1}{n_1^{s+1}} \|\vec{x}_1^{s+1} - \vec{a}_1^{s+1} \circ \vec{r}^{t+1}\|_2^2 \\ \vdots \\ \frac{1}{n_m^{s+1}} \|\vec{x}_m^{s+1} - \vec{a}_m^{s+1} \circ \vec{r}^{t+1}\|_2^2 \end{bmatrix}.$$

The convergence issue is clearly more delicate here but in the case of periodic votes it is e.g. shown in [2] that under mild conditions, the iteration converges also to a “fixed” periodic limit cycle.

Other discriminant functions

The scalar function

$$g(d) = 1 - kd$$

links the belief-divergence $\vec{d}$ to the weights $\vec{w}$ by an affine function. A similar idea was already present in [6], [9] and [5], but using different scalar functions, namely

$$g(d) = \frac{1}{d}, \quad g(d) = \frac{1}{\sqrt{d}}, \quad \text{and} \quad g(d) = e^{-kd}.$$

The motivation for using these more complex functions is that the corresponding minimization problem has a statistical interpretation, but the use of these functions also makes the problem of characterizing the fixed points harder. We refer to [2] for a further discussion on this issue.

1.5 A worked example

We illustrate the sparse extension of the method described in Section 1.3 with an experiment involving a data set (supplied by the GroupLens Research Project) of 100,000 ratings given by 943 users on 1682 movies. The votes ranged from 1 to 5 and the movies were selected such that each rater voted on at least 20 of the 1682 movies. This corresponds to a very sparse voting matrix but since every voter has a sufficient overlap with other raters, the computation of the divergence between raters remains sufficiently relevant. In order to test the robustness of our reputation system, we added 237 spammers giving always a vote of 1 except for their preferred movie, which they rated with a vote of 5.

Let $\vec{r}^*$ and $\tilde{\vec{r}}^*$ be respectively the reputation vector before and after the addition of these spammers. We expect that such behavior will be penalized by decreasing the spammer’s weights. Figure 1.4 illustrates the effect of adding spammers for the two different methods we want to compare: firstly for our method using weighted averages, where the total perturbation can be measured by the distance $\|\vec{r}^* - \tilde{\vec{r}}^*\|_1 = 267$, and secondly by taking the unweighted average, where it is clearly seen that all reputations tend to be diminished. The distance is then given by $\|\vec{r}^* - \tilde{\vec{r}}^*\|_1 = 638$ and, as expected, this is greater than in the previous case, since spammers receive as much weight as the other raters.

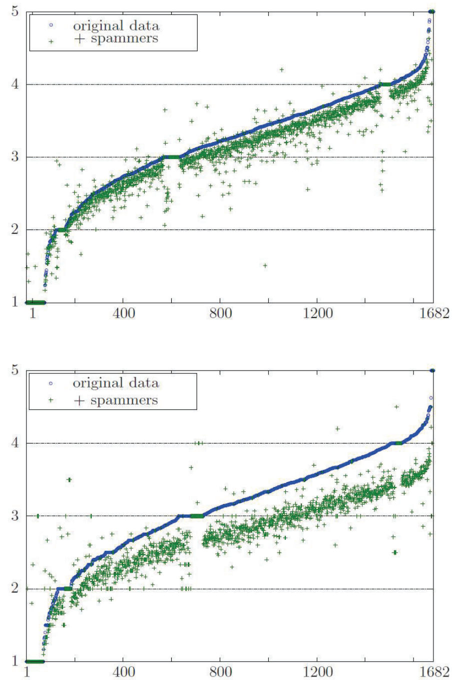


FIGURE 1.4: X-Axis: the sorted movies according to their reputations before the addition of spammers. Y-Axis: their reputations according to our algorithm (Top) and to the average (Bottom).

Let us now look at the evolution of the weights during the iterations. The distribution of weights is shown in Figure 1.5 after one step, two steps and after convergence. Clearly the spammers receive eventually a much smaller weight than the original voters, and the method converges in very few steps.

This shows that our method could also be used to characterize outliers. Spammers could be detected by setting a threshold on the rater's weights w_i. In our example it follows from the last plot in Figure 1.5 that most raters with a weight below 0.6 are spammers and that most of those with a weight larger than 0.6 are not. Dismissing completely the raters with a low converged weight can thus be viewed as a method to eliminate spammers. However, we prefer to take them into account, though with a reduced weight.

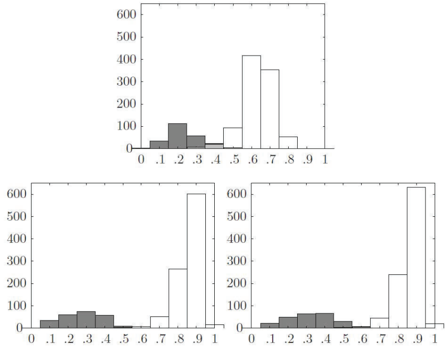


FIGURE 1.5: X-Axis: the weights of the raters. Y-Axis: the density after one iteration (Top), after two iterations (Left), and after convergence (Right). In dark grey: the spammers. In white: the original raters. In light grey: overlap of both raters.

1.6 Concluding remarks

Several other examples can be found in [2] and [3]. It is shown there that the voting system in competitions like the Eurosong contest suffers from so-called cultural voting (see [4]) and that this deficiency can be corrected by employing our filtering techniques. We point out that our technique could also be applied to websites such as Booking.com, Amazon.com, Tripadvisor and so on, where rankings are being offered based on anonymous votes. Bringing order to votes on the Web is certainly a promising topic that requires further investigations. Refining votes and hence reputations is one way to achieve that aim.

In conclusion, the main issues in reputation systems are, in our view, the relevance of the measure, the robustness against different sort of attackers, the application of the method to any sorts of data and the easiness for users to understand the measure.

It is surely tricky to determine how relevant a measure is in the context of voting, In our case we accept the idea of *belief-divergence* as a basis for calculating the rater's weight, even if it implies the disqualification of marginal users. Nevertheless the parameter d allows us to quantify the degree of discrimination. Moreover the exponent in equation (1.2) can be chosen to be higher than 2, but then the uniqueness of the fixed point is no longer guaranteed. Several fixed points, however, could be interpreted as different opinion groups. In that way a marginal group would maintain its reputation if the number of its members is large enough. It makes sense to allow several opinions for a same movie, but providing one intrinsic value for each item should be more relevant in most contexts (such as the reputation of sellers and buyers on E-bay).

A durable reputation system must be robust. Smart cheaters who understand the system well enough to take advantage of it are certainly the most effective spammers. The way to proceed according to our approach is simple: they need to establish their reliability by correctly evaluating a group of items and then with that trust, they can rate some target items. In order to significantly change the reputation of these target items, they must have a number of coordinated evaluations that is larger than the one of honest raters. Such cheaters can thus be easily disqualified by looking at their coordinated ratings of one or several items. Unfortunately, this sort of spammer requires an extra process, similar to the procedure used by Google to detect spam farms that create thousands of links to boost their page ranks

The constantly increasing size of web data sets requires algorithms that are not too time consuming. Typically, a linear complexity in the number of votes is ideal. This is especially true for dynamical reputation systems, where the frequency of updates is high. In addition to this efficiency requirement, the method must be applicable to any ”sparse” data. In general, the network resulting from votes between raters and objects is not complete, i.e., each object is not evaluated by all raters. These two points - complexity and sparse data - are a must if one wants a widespread use of a reputation system.

It is also desirable that reputation systems be able to cope with time-dependent data. More recent opinions may be considered more valuable than older ones, especially in the case of timely items such as news, arts, fashion, etc. The approach we described above can be easily extended to incorporate dynamical systems with time-dependent votes (see [2]) but the convergence to a fixed point is then replaced by the tracking of a time-varying trend.

Last but not least, the method must be understandable by those to whom it is directed, namely, by the users. Indeed, users cannot be confident in a voting if the measure of reputation looks like a black box. Although our method is more complicated that a simple E-bay-like system in which all ratings have the same weight, it remains relatively simple. In addition, users like the voting system to be transparent. For example, a record of the voting history and comments from the raters helps users to develop their own opinions.

Note that all figures in this chapter are taken from Chapter 7 of the first author's Ph.D. thesis, Cristobald de Kerchove [3].

Bibliography

[1] CNN article, *Canadian skaters get gold; judge suspended,* https://www.cnn.com/2002/US/02/15/oly.skate.row/

[2] C. de Kerchove, P. Van Dooren, *Iterative Filtering for a Dynamical Reputation System*, Siam J. Matrix Anal. Appl. (2008).

[3] C. de Kerchove, *Ranking Large Networks: Leadership, Optimization, and Distrust* Ph.D. thesis, Universtié catholique de Louvain, Louvain-la-Neuve, Belgium, 2009.

[4] V. Ginsburgh, A. Noury, *Cultural Voting: The EuroVision Song Contest*, mimeo, 2004; available online from http://ideas.repec.org/p/cor/louvco/2005006.html#download

[5] E. Kotsovinos, P. Zerfos, N. M. Piratla, N. Cameron, S. Agarwal, *Jiminy: A Scalable Incentive-Based Architecture for Improving Rating Quality*, iTrust06, LNCS 3986 (2006), 221-236.

[6] P. Laureti, L. Moret, Y.-C. Zhang, Y.-K. Yu, *Information Filtering via Iterative Refinement*, EuroPhysic Letter 75 (2006), 1006-1012.

[7] The New York Times article, *FIGURE SKATING; 2 French Officials Suspended 3 Years In Skating Scandal,* https://www.nytimes.com/2002/05/01/sports/figure-skating-2-french-officials-suspended-3-years-in-skating-scandal.html

[8] Wikipedia 2002 Winter Olpympics Figure Skating Scandal, https://en.wikipedia.org/wiki/2002_Winter_Olympics_figure_skating_scandal#The_competition

[9] S. Zhang, Y. Ouyang, J. Ford, F. Make, *Analysis of a Low-dimensional Linear Model under Recommendation Attacks*, Proceedings of the 29th annual International ACM SIGIR conference on Research and development in information retrieval, (2006), 517-524.

Chapter 2

A priori regularity of parabolic partial differential equations

F. Berkemeier
King Abdullah University of Science and Technology (KAUST), CEMSE Division, Thuwal 23955-6900, Saudi Arabia,
E-mail: francisco.berkemeier.1@kaust.edu.sa

D. A. Gomes
King Abdullah University of Science and Technology (KAUST), CEMSE Division, Thuwal 23955-6900, Saudi Arabia,
E-mail: diogo.gomes@kaust.edu.sa

This Chapter is devoted to the analysis of parabolic partial differential equations and to the development of methods that provide *a priori* estimates for solutions with singular initial data. These estimates are obtained by understanding the time decay of the norms of the solutions. First, regularity results are derived for the Fokker-Planck equation by estimating the decay of Lebesgue norms. These estimates depend on integral bounds for the advection and diffusion. Then, we apply similar methods to the heat equation. Finally, we conclude by extending our techniques to the porous media equation. The sharpness of our results is confirmed by examining known solutions of these equations. Our main contribution is the use of functional inequalities to establish the decay of norms by means of nonlinear differential inequalities. These are then combined with ODE methods to deduce estimates for the norms of the solutions and their derivatives.

2.1 Introduction

Parabolic partial differential equations are often used to describe the diffusion of mass, momentum or heat through a material. A classical parabolic

PDE is the heat equation:

$$u_t(x,t) = \Delta u(x,t), \tag{2.1}$$

where $u : \mathbb{R}^d \times [0,T] \to \mathbb{R}$. It is well known that the solution to (2.1) with singular initial data $u(x,0) = \delta_{x_0}$ is the fundamental solution

$$\Phi(x,t) = \frac{1}{(4\pi t)^{d/2}} e^{-\frac{|x|^2}{4t}}.$$

Although when $t \to 0$, Φ becomes singular, for $t > 0$, Φ is smooth in x and in any L^p space. More precisely, the L^1-norm of this solution is conserved and the L^p-norms decay in time as follows

$$\|\Phi\|_{L^p(\mathbb{R}^d)} = C_p t^{-\frac{1}{2p}d(p-1)}$$

for some constant $C_p > 0$. The preceding identity can be checked by direct computation. Here, we will prove similar bounds for solutions of parabolic equations without relying on explicit formulas for the solutions.

We begin by investigating the Fokker-Planck equation

$$u_t(x,t) = \operatorname{div}(b(x,t)u(x,t)) + \operatorname{div}(a(x,t)\nabla u(x,t)),$$

where a is a positive scalar diffusion coefficient and b is a smooth advection vector field . This second-order equation, also known as the Kolmogorov forward equation, models the behavior of a particle under the effect of drag (corresponding to the advection term, b) and random forces (corresponding to the diffusion coefficient, a) and has applications in physics, polymer fluids, plasma, surface physics, and finance, to name just a few. Here, for initial data u_0 and a domain Ω, we obtain estimates of the form

$$\|D^k u\|_{L^p(\Omega)} \le C\|u_0\|_{L^1(\Omega)}^{f(p,d,k)} t^{-g(p,d,k)},$$

where $k \in \mathbb{N}_0$, $f, g \geq 0$ are functions of dimension d, k and p, and C is a non-negative constant depending on the space and the problem parameters. Moreover, these estimates depend only on the L^1-norm of the initial data and do not depend on the particular solution.

Our main results on the Fokker-Planck equation are as follows. First, under assumptions on the divergence of the advection, we obtain the following theorem:

Theorem 2.1. *Let u be a solution of* (2.9) *with $u \in C^\infty(\mathbb{R}^d \times [0,\infty))$. Let $a > 0$. Moreover, assume $a \in L^{\frac{1}{1-q}}(\mathbb{R}^d)$ for some $1 < q < 2$. Then, for $d \geq 2$, the following holds:*

1. If $d=2$ and $1<q<2$, or $q<(d+2)/d$ and $d\geq 3$, $divb=0$, and $p>1$, then, for all $t>0$,

$$\|u\|_{L^p(\mathbb{R}^d)} \leq C\|u_0\|_{L^1(\mathbb{R}^d)} t^{-\frac{d(p-1)}{p(2-d(q-1))}}. \tag{2.2}$$

2. If $divb \in L^r(\mathbb{R}^d)$ and p,q are such that

$$2 \leq d < 2r \text{ and } 1 < q < \frac{2r+dr-d}{dr}, \tag{2.3}$$

then, there exists $T>0$ such that

$$\|u\|_{L^p(\mathbb{R}^d)} \leq C\|u_0\|_{L^1(\mathbb{R}^d)} t^{-\frac{d(p-1)}{p(2-d(q-1))}} \tag{2.4}$$

for all $t<T$. For $t>T$, $\|u\|_{L^p(\mathbb{R}^d)} \leq C\|u_0\|_{L^1(\mathbb{R}^d)} T^{-\frac{d(p-1)}{p(2-d(q-1))}}$.

Remark 2.1. The exponent on the right-hand side of (2.2) is negative if $d=2$ and $1<q<2$, or $q<(d+2)/d$ and $d\geq 3$.

Under further integrability assumptions on the advection, we have the following result.

Theorem 2.2. *Let u solve (2.9) with $u\in C^\infty(\mathbb{R}^d\times[0,\infty))$. Moreover, assume that $a^{-1}\in L^r(\mathbb{R}^d)$ and that $|b|\in L^{\frac{2rq}{r-1}}(\mathbb{R}^d)$ for some $q>1$, $r>2$. Then, for any $p>1$ and $d\geq 2$, the following holds:*

1. Let q be such that

$$q > \frac{d(1-r)}{d-2r} \text{ for } 2<d<2r. \tag{2.5}$$

If a is bounded by above and below, there exists $T>0$ such that

$$\|u\|_{L^p(\mathbb{R}^d)} \leq C\|u_0\|_{L^1(\mathbb{R}^d)} t^{-\frac{qr(p-1)}{p(r+q-1)}} \tag{2.6}$$

for all $t<T$. For $t>T$, $\|u\|_{L^p(\mathbb{R}^d)} \leq C\|u_0\|_{L^1(\mathbb{R}^d)} T^{-\frac{qr(p-1)}{p(r+q-1)}}$.

2. Let q be such that

$$q > \frac{d(1-r)}{dr(s-1)+d-2r} \text{ for } \frac{2}{s} < d < \frac{2r}{1+r(s-1)}. \tag{2.7}$$

Moreover, if $a\in L^{\frac{1}{1-s}}(\mathbb{R}^d)$, there also exists $T>0$ such that

$$\|u\|_{L^p(\mathbb{R}^d)} \leq C\|u_0\|_{L^1(\mathbb{R}^d)} t^{-\frac{qr(p-1)}{p(r+q-1)}} \tag{2.8}$$

for $1<s<2$ and $t<T$. For $t>T$, $\|u\|_{L^p(\mathbb{R}^d)} \leq C\|u_0\|_{L^1(\mathbb{R}^d)} T^{-\frac{qr(p-1)}{p(r+q-1)}}$.

The proofs of the prior theorems are presented in Section 2.2. There, we also discuss an application to L^∞ bounds for the solutions of (2.9) with singular initial data. Then, in Section 2.3, we study a particular case, the heat equation. There, we compare our methods with the entropy method [8] and hypercontractivity [1, 5, 7, 11, 12].

Finally, in Section 2.4, we extend our results to the porous media equation,

$$u_t(x,t) = \Delta(u(x,t)^m),$$

where $m \geq 1$. This equation models diffusion processes and fluid flow through porous media (such as a sponge or wood, for example) and has applications to mathematical biology, lubrication and boundary-layer theory.

Our main contribution is the use of functional inequalities and of a differential argument to derive a method to prove estimates for the norms of solutions of linear and nonlinear equations. This method systematizes techniques to infer estimates for solutions of parabolic PDE.

Similar techniques were studied in [2, 14, 15] and used to establish smoothing effects and the time decay of solutions of the heat equation and of the porous media equation. A method comparable to ours was studied in [9, 10]. There, the regularizing effect and the long- and short-time decay were studied for the parabolic Cauchy-Dirichlet problem and the viscous Hamilton-Jacobi equation with a superlinear Hamiltonian.

There are three key techniques used to prove our results. First, we expand the time derivative of the L^p-norms and use integration by parts to establish the decay of these norms. Then, we combine Gagliardo-Nirenberg and Sobolev inequalities with the conservation of L^1-norms to obtain a nonlinear dissipation estimate. Finally, we apply a nonlinear Grönwall-type estimate to get decay in time.

2.2 Fokker-Planck equations

Consider the Fokker-Planck equation with initial data in L^1:

$$\begin{cases} u_t(x,t) = \operatorname{div}(b(x)u(x,t)) + \operatorname{div}(a(x)\nabla u(x,t)) & \text{in } \mathbb{R}^d \times (0,\infty) \\ u(x,0) = u_0(x) & \text{in } \mathbb{R}^d, \end{cases} \tag{2.9}$$

where a is a positive scalar diffusion coefficient and b is a smooth advection vector field. In this section, we derive integrability conditions on a and b that imply decay estimates for the Lebesgue norms. To simplify the discussion, we assume that a and b are time-independent. We are interested in two scenarios. In the first one, we assume integrability on the divergence of b. In the second such scenario, we assume integrability on b. When $b = 0$ and $a = 1$, (2.9) becomes the heat equation for which we deduce further regularity in the following section.

2.2.1 Integrability conditions on the divergence of the advection

Here, we prove Theorem 2.1 and obtain the two estimates for the solutions of (2.9) depending on the properties of div$b$.

Proof of Theorem 2.1. First, observe that

$$\frac{d}{dt}\int_{\mathbb{R}^d} u^p\,dx = p\int_{\mathbb{R}^d} u^{p-1}\mathrm{div}(bu)\,dx + p\int_{\mathbb{R}^d} u^{p-1}\mathrm{div}(a\nabla u)\,dx. \tag{2.10}$$

The reverse Hölder inequality, for functions f and g, states that

$$\|fg\|_{L^1(\mathbb{R}^d)} \geq \|f\|_{L^{\frac{1}{q}}(\mathbb{R}^d)}\|g\|_{L^{\frac{1}{1-q}}(\mathbb{R}^d)},$$

whenever $q > 1$. Then, since $a \in L^{\frac{1}{1-q}}(\mathbb{R}^d)$, we have

$$\int_{\mathbb{R}^d} u^{p-1}\mathrm{div}(a\nabla u)\,dx = -C\int_{\mathbb{R}^d} au^{p-2}|\nabla u|^2\,dx \leq -C\left(\int_{\mathbb{R}^d}(u^{p-2}|\nabla u|^2)^{\frac{1}{q}}\,dx\right)^q.$$

Fix $\gamma = p/2$. Then, by the Gagliardo-Nirenberg-Sobolev inequality for $q < 2$, it follows that

$$\left(\int_{\mathbb{R}^d}(u^{p-2}|\nabla u|^2)^{\frac{1}{q}}\,dx\right)^q = \left(\int_{\mathbb{R}^d}|\nabla(u^\gamma)|^{\frac{2}{q}}\,dx\right)^q \geq C\left(\int_{\mathbb{R}^d} u^{\gamma q^*}\,dx\right)^{\frac{2}{q^*}}, \tag{2.11}$$

where q^* is the Sobolev conjugate exponent to $\frac{2}{q}$, given by $q^* = \frac{2d}{dq-2}$. Using the interpolation inequality, L^1-norm conservation, and $0 < \lambda < 1$ with

$$\frac{1}{p} = 1 - \lambda + \frac{\lambda}{\gamma q^*},$$

we have that

$$\left(\int_{\mathbb{R}^d} u^{\gamma q^*}\,dx\right)^{\frac{\lambda}{q^*}} = \|u\|^{\gamma\lambda}_{L^{\gamma q^*}(\mathbb{R}^d)} = \|u\|^{\gamma\lambda}_{L^{\gamma q^*}(\mathbb{R}^d)} \geq \|u\|^{\gamma}_{L^p(\mathbb{R}^d)}\|u_0\|^{\gamma(\lambda-1)}_{L^1(\mathbb{R}^d)},$$

where $\lambda = \frac{d(p-1)}{2+d(p-q)}$. Combining the previous estimates it is clear that

$$\int_{\mathbb{R}^d} u^{p-1}\operatorname{div}(a\nabla u)\,dx \le -C\left(\int_{\mathbb{R}^d} u^p\,dx\right)^{\beta} \|u_0\|_{L^1(\mathbb{R}^d)}^{\frac{2\gamma(\lambda-1)}{\lambda}},$$

where $\beta = \frac{1}{\lambda} = \frac{2+d(p-q)}{d(p-1)}$. For the other term in (2.10), it follows that

$$\begin{aligned}\int_{\mathbb{R}^d} u^{p-1}\operatorname{div}(bu)\,dx &= -\int_{\mathbb{R}^d} u^{p-1}\nabla u\cdot b\,dx\\ &= -C\int_{\mathbb{R}^d} \nabla(u^p)\cdot b\,dx = C\int_{\mathbb{R}^d} u^p \operatorname{div} b\,dx.\end{aligned}$$

Therefore, if $\operatorname{div} b = 0$, with $z(t) = \int_{\mathbb{R}^d} u^p\,dx$, one has the inequality

$$\dot z \le -C\|u_0\|_{L^1(\mathbb{R}^d)}^{\frac{2\gamma(\lambda-1)}{\lambda}} z^{\beta}.$$

Thus, by Lemma 2.1,

$$z(t) \le C\|u_0\|_{L^1(\mathbb{R}^d)}^{\frac{2\gamma(\lambda-1)}{\lambda(1-\beta)}} t^{\frac{1}{1-\beta}} = C\|u_0\|_{L^1(\mathbb{R}^d)}^{p} t^{-\frac{d(p-1)}{2-d(q-1)}},$$

which yields the estimate in (2.2).

2. Now, assume that $\operatorname{div} b \in L^r(\mathbb{R}^d)$. Hence, Hölder's inequality leads to

$$\int_{\mathbb{R}^d} u^p \operatorname{div} b\,dx \le \left(\int_{\mathbb{R}^d} u^{pr'}\,dx\right)^{\frac{1}{r'}}\left(\int_{\mathbb{R}^d} (\operatorname{div} b)^r\,dx\right)^{\frac{1}{r}},$$

where $1/r' + 1/r = 1$. From (2.11), we have

$$\int_{\mathbb{R}^d} u^{p-1}\operatorname{div}(a\nabla u)\,dx \le -C\left(\int_{\mathbb{R}^d} u^{\gamma q^*}\,dx\right)^{\frac{2}{q^*}},$$

where $\gamma = p/2$ and $q^* = 2d/(dq-2)$. Then,

$$\frac{d}{dt}\int_{\mathbb{R}^d} u^p\,dx \le C\left(\int_{\mathbb{R}^d} u^{pr'}\,dx\right)^{\frac{1}{r'}} - C\left(\int_{\mathbb{R}^d} u^{\gamma q^*}\,dx\right)^{\frac{2}{q^*}}. \tag{2.12}$$

Note that it follows by interpolation that

$$\left(\int_{\mathbb{R}^d} u^{pr'}\,dx\right)^{\frac{1}{r'}} \le \left(\int_{\mathbb{R}^d} u^{\gamma q^*}\,dx\right)^{\frac{p\theta}{\gamma q^*}} \|u_0\|_{L^1(\mathbb{R}^d)}^{p(1-\theta)},$$

where θ is such that $\frac{1}{pr'} = \frac{\theta}{\gamma q^*} + 1 - \theta$. Observe that the previous inequality only holds if

$$pr' < \gamma q^*;$$

that is, if

$$\frac{pr}{r-1} < \frac{pd}{dq-2},$$

which is true if (2.3) holds. Therefore, with $y(t) = \int_{\mathbb{R}^d} u^{\gamma q^*}\,dx$, the right-hand side of (2.12) is bounded by

$$C_1\|u_0\|_{L^1(\mathbb{R}^d)}^{p(1-\theta)} y^{\frac{p\theta}{\gamma q^*}} - C_2 y^{\frac{2}{q^*}} = C_1\|u_0\|_{L^1(\mathbb{R}^d)}^{p(1-\theta)} y^{\frac{2\theta}{q^*}} - C_2 y^{\frac{2}{q^*}}.$$

Then, since $\theta < 1$, with $z(t) = \int_{\mathbb{R}^d} u^p\,dx$, using Lemma 2.2 and interpolation again, one concludes that there exists $T > 0$ such that, for all $t < T$,

$$\begin{aligned}
\dot{z} \le -Cy^{\frac{2}{q^*}} &= -C\left(\int_{\mathbb{R}^d} u^{\gamma q^*}\,dx\right)^{\frac{2}{q^*}} \le -C\left(\int_{\mathbb{R}^d} u^p\,dx\right)^{\frac{1}{\lambda}} \|u_0\|_{L^1(\mathbb{R}^d)}^{\frac{2\gamma(\lambda-1)}{\lambda}} \\
&= -C\|u_0\|_{L^1(\mathbb{R}^d)}^{\frac{2\gamma(\lambda-1)}{\lambda}} z^{\frac{1}{\lambda}},
\end{aligned}$$

where $\lambda = \frac{d(p-1)}{2+dp-dq}$. Then

$$z(t) \le C\|u_0\|_{L^1(\mathbb{R}^d)}^{\frac{2\gamma(\lambda-1)}{\lambda(1-1/\lambda)}} t^{\frac{1}{1-1/\lambda}} = C\|u_0\|_{L^1(\mathbb{R}^d)}^{p} t^{\frac{d(p-1)}{d(q-1)-2}},$$

and thus (2.4) follows, for all $t < T$. For $t > T$,

$$\|u\|_{L^p(\mathbb{R}^d)} \le C\|u_0\|_{L^1(\mathbb{R}^d)} T^{-\frac{d(p-1)}{p(2-d(q-1))}}.$$

□

2.2.2 Integrability conditions on the advection

Recall now Theorem 2.2, where the integrability of the advection was considered.

Proof of Theorem 2.2. We have

$$\begin{aligned}
\frac{d}{dt}\int_{\mathbb{R}^d} u^p\,dx &= C\int_{\mathbb{R}^d} u^{p-1}\operatorname{div}(bu)\,dx + C\int_{\mathbb{R}^d} u^{p-1}\operatorname{div}(a\nabla u)\,dx \\
&= -C\int_{\mathbb{R}^d} u^{p-1}\nabla u\cdot b\,dx - C\int_{\mathbb{R}^d} au^{p-2}|\nabla u|^2\,dx \\
&= -C\int_{\mathbb{R}^d} a^{\frac{1}{2}}u^{\frac{p}{2}-1}\nabla u\cdot bu^{\frac{p}{2}}a^{-\frac{1}{2}}\,dx - C\int_{\mathbb{R}^d} au^{p-2}|\nabla u|^2\,dx.
\end{aligned}$$

Then, reorganizing the previous inequality and using Cauchy's inequality with ε, it is clear that

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{R}^d} u^p\,dx + C\int_{\mathbb{R}^d} au^{p-2}|\nabla u|^2\,dx &= -C\int_{\mathbb{R}^d} a^{\frac{1}{2}}u^{\frac{p}{2}-1}\nabla u\cdot bu^{\frac{p}{2}}a^{-\frac{1}{2}}\,dx\\ &\leq \left|C\int_{\mathbb{R}^d} a^{\frac{1}{2}}u^{\frac{p}{2}-1}\nabla u\cdot bu^{\frac{p}{2}}a^{-\frac{1}{2}}\,dx\right| \qquad (2.13)\\ &\leq \varepsilon C\int_{\mathbb{R}^d}|a|u^{p-2}|\nabla u|^2\,dx + C_\varepsilon\int_{\mathbb{R}^d}|b|^2u^p|a|^{-1}\,dx.\end{aligned}$$

Hence, for ε small, (2.13) can be rewritten as

$$\frac{d}{dt}\int_{\mathbb{R}^d} u^p\,dx + C\int_{\mathbb{R}^d} au^{p-2}|\nabla u|^2\,dx \leq C_\varepsilon\int_{\mathbb{R}^d}|b|^2u^p|a|^{-1}\,dx. \qquad (2.14)$$

Now, applying Hölder's inequality twice to the last term in the previous inequality it follows:

$$\begin{aligned}\int_{\mathbb{R}^d}|b|^2u^p|a|^{-1}\,dx &\leq \left(\int_{\mathbb{R}^d}|b|^{2r'}u^{pr'}\right)^{\frac{1}{r'}}\left(\int_{\mathbb{R}^d}|a|^{-r}\right)^{\frac{1}{r}}\\ &\leq C\left(\int_{\mathbb{R}^d}u^{pr'q'}\right)^{\frac{1}{r'q'}}\left(\int_{\mathbb{R}^d}|b|^{2r'q}\right)^{\frac{1}{r'q}} \leq C\left(\int_{\mathbb{R}^d}u^{pr'q'}\right)^{\frac{1}{r'q'}},\end{aligned}$$

where $\frac{1}{r}+\frac{1}{r'}=1=\frac{1}{q}+\frac{1}{q'}$ and $r'q=\frac{rq}{r-1}$. Accordingly, defining $\gamma = pr'q' = \frac{pqr}{(q-1)(r-1)}$, one has from (2.14),

$$\frac{d}{dt}\int_{\mathbb{R}^d} u^p\,dx \leq C_1\left(\int_{\mathbb{R}^d}u^\gamma\,dx\right)^{\frac{p}{\gamma}} - C_2\int_{\mathbb{R}^d}a|\nabla(u^{\frac{p}{2}})|^2\,dx,$$

where $C_1, C_2 > 0$ are constants depending on η and ε. Now the two cases are considered separately.

1. If a is bounded above and below, then, by virtue of Sobolev's inequality, one has

$$\int_{\mathbb{R}^d}a|\nabla(u^{\frac{p}{2}})|^2\,dx \geq C\left(\int_{\mathbb{R}^d}u^{\frac{2^*p}{2}}\,dx\right)^{\frac{2}{2^*}}.$$

Then, using interpolation,

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{R}^d} u^p\,dx &\leq C_1\left(\int_{\mathbb{R}^d}u^\gamma\,dx\right)^{\frac{p}{\gamma}} - C_2\left(\int_{\mathbb{R}^d}u^{\frac{2^*p}{2}}\,dx\right)^{\frac{2}{2^*}}\\ &\leq C_1\left(\int_{\mathbb{R}^d}u^{\frac{2^*p}{2}}\,dx\right)^{\frac{2\theta}{2^*}}\|u_0\|_{L^1(\mathbb{R}^d)}^{p(1-\theta)} - C_2\left(\int_{\mathbb{R}^d}u^{\frac{2^*p}{2}}\,dx\right)^{\frac{2}{2^*}},\end{aligned}$$

where θ is such that $\frac{1}{\gamma} = \frac{2\theta}{2^* p} + 1 - \theta$. The previous inequality only holds if $\gamma \leq 2^* p/2$. This is true for q such that (2.5) holds. Hence, since $\theta < 1$, using Lemma 2.2 and interpolation again, it is concluded that there exists $T > 0$ such that, for all $t < T$,

$$\frac{d}{dt}\int_{\mathbb{R}^d} u^p\,dx \leq -C\left(\int_{\mathbb{R}^d} u^p\,dx\right)^{\frac{1}{\lambda}} \|u_0\|_{L^1(\mathbb{R}^d)}^{\frac{p(\lambda-1)}{\lambda}}$$

for some $\lambda > 0$ such that $\frac{1}{p} = \frac{\lambda}{\gamma} + 1 - \lambda \Leftrightarrow \lambda = \frac{\gamma(p-1)}{p(\gamma-1)}$, which yields

$$\lambda = \frac{qr(p-1)}{qr(p-1)+q+r-1}.$$

Hence, setting $z(t) = \int_{\mathbb{R}^d} u^p\,dx$, there follows an inequality of the type $\dot{z} \leq -C\|u_0\|_{L^1(\mathbb{R}^d)}^{\frac{p(\lambda-1)}{\lambda}} z^{\frac{1}{\lambda}}$. Thus,

$$z(t) \leq C\|u_0\|_{L^1(\mathbb{R}^d)}^{p} t^{\frac{1}{1-1/\lambda}} = C\|u_0\|_{L^1(\mathbb{R}^d)}^{p} t^{\frac{qr(1-p)}{r+q-1}},$$

which combined with (2.5), yields (2.6), for $t < T$. For $t > T$, $\|u\|_{L^p(\mathbb{R}^d)} \leq C\|u_0\|_{L^1(\mathbb{R}^d)} T^{-\frac{qr(p-1)}{p(r+q-1)}}$.

2. If $a \in L^{\frac{1}{1-s}}(\mathbb{R}^d)$, by virtue of the reverse Hölder's inequality, it follows that

$$\begin{aligned}\int_{\mathbb{R}^d} a|\nabla(u^{\frac{p}{2}})|^2\,dx &\geq \left(\int_{\mathbb{R}^d} a^{\frac{1}{1-s}}\,dx\right)^{1-s}\left(\int_{\mathbb{R}^d} |\nabla(u^{\frac{p}{2}})|^{\frac{2}{s}}\,dx\right)^{s}\\ &\geq C\left(\int_{\mathbb{R}^d} |\nabla(u^{\frac{p}{2}})|^{\frac{2}{s}}\,dx\right)^{s}.\end{aligned}$$

Then, for $s < 2$, the Gagliardo-Nirenberg-Sobolev inequality yields

$$\left(\int_{\mathbb{R}^d} |\nabla(u^{\frac{p}{2}})|^{\frac{2}{s}}\,dx\right)^{s} \geq C\left(\int_{\mathbb{R}^d} u^{\frac{mp}{2}}\,dx\right)^{\frac{2}{m}}$$

with $m = \frac{2d}{ds-2}$. Furthermore, interpolation and L^1-norm conservation yield

$$\left(\int_{\mathbb{R}^d} u^{\gamma}\,dx\right)^{\frac{p}{\gamma}} \leq \left(\int_{\mathbb{R}^d} u^{\frac{mp}{2}}\,dx\right)^{\frac{2\theta}{m}} \|u_0\|_{L^1(\mathbb{R}^d)}^{p(1-\theta)}, \tag{2.15}$$

where θ is such that $\frac{1}{\gamma} = \frac{2\theta}{mp} + 1 - \theta$. Note that (2.15) holds if $\gamma < mp/2$.

This is true for q such that (2.7) holds. Following the same steps as before, since $\theta < 1$, one concludes that there exists $T > 0$ such that, for all $t < T$,

$$z(t) \leq C\|u_0\|^p_{L^1(\mathbb{R}^d)} t^{\frac{1}{1-1/\lambda}} = C\|u_0\|^p_{L^1(\mathbb{R}^d)} t^{\frac{qr(1-p)}{r+q-1}},$$

which combined with (2.7), yields (2.8), for $1 < s < 2$ and $t < T$. For $t > T$, one has $\|u\|_{L^p(\mathbb{R}^d)} \leq C\|u_0\|_{L^1(\mathbb{R}^d)} T^{-\frac{qr(p-1)}{p(r+q-1)}}$. □

2.2.3 The adjoint method

As an application of our estimates we present bounds of the form

$$\|v(\cdot,0)\|_{L^\infty(\Omega)} \leq C\|f\|_{L^b([0,T],L^q(\Omega))} \tag{2.16}$$

for solutions of

$$\begin{cases} v_t + b\cdot\nabla v = \operatorname{div}(a\nabla v) + f & \text{in } \Omega\times(0,T] \\ v(x,T) = v_T(x) & \text{in } \Omega, \end{cases} \tag{2.17}$$

where $\Omega = \mathbb{R}^d$ or $\Omega = \mathbb{T}^d$ and $v_T \in W^{1,\infty}(\Omega)$. Such bounds can be proved by means of the adjoint method. Estimates such as (2.16) arise in the theory of mean-field games, for example. As in [3, 4, 6], the adjoint problem to (2.17) is

$$\begin{cases} u_t = \operatorname{div}(ub) + \operatorname{div}(a\nabla u) & \text{in } \Omega\times(0,T] \\ u(x,0) = \delta_{x_0} & \text{in } \Omega. \end{cases} \tag{2.18}$$

The central idea of the adjoint method is to derive a representation formula for solutions of (2.17) in terms of solutions of (2.18). Arguing as in [6], it follows that

$$v(\cdot,0) = \int_0^T\int_\Omega f(x,t)u(x,t)\,dxdt + \int_\Omega v_T(x)u(x,T)\,dx.$$

Thus

$$|v(\cdot,0)| \leq \int_0^T\int_\Omega |f(x,t)u(x,t)|\,dxdt + \int_\Omega |v_T(x)u(x,T)|\,dx. \tag{2.19}$$

Therefore, to estimate the left-hand side, it is enough to bound each of the two terms on the right-hand side of the preceding inequality. For the second term on the right-hand side, we have that, by Hölder's inequality,

$$\int_\Omega |v_T(x)u(x,T)|\,dx \leq \|v_T\|_{L^\infty(\Omega)}\|u(x,T)\|_{L^1(\Omega)} = \|v_T\|_{L^\infty(\Omega)} \leq C,$$

since $v_T \in W^{1,\infty}(\Omega)$. For the first term in (2.19), we apply Hölder's inequality twice to conclude that

$$\int_0^T \int_\Omega |fu|\,dxdt \leq \int_0^T \|f\|_{L^q(\Omega)} \|u\|_{L^p(\Omega)}\,dt \leq \|f\|_{L^b([0,T],L^q(\Omega))} \|u\|_{L^c([0,T],L^p(\Omega))}, \tag{2.20}$$

where $\frac{1}{b} + \frac{1}{c} = 1 = \frac{1}{p} + \frac{1}{q}$. It is thus clear that bounds for u can be converted into bounds for v. Therefore, the estimates from Theorems 2.1 and 2.2, which still hold for the Fokker-Planck equation with singular initial data, yield estimates for $\|v(\cdot,0)\|_{L^\infty(\Omega)}$. We have the following result.

Theorem 2.3. *Let v,u solve* (2.17) *and* (2.18)*, respectively, in* $\mathbb{R}^d$. *Let* $\frac{1}{b} + \frac{1}{c} = 1 = \frac{1}{p} + \frac{1}{q}$. *Then,*

1. *Under the assumptions of Theorem 2.1, if*

$$c > \frac{p(2-d(q-1))}{d(p-1)}, \tag{2.21}$$

then $\|v(\cdot,0)\|_{L^\infty(\mathbb{R}^d)} \leq C\|f\|_{L^b([0,T],L^q(\mathbb{R}^d))}$.

2. *Under the assumptions of Theorem 2.2, if*

$$c > \frac{p(r+q-1)}{qr(p-1)},$$

then $\|v(\cdot,0)\|_{L^\infty(\mathbb{R}^d)} \leq C\|f\|_{L^b([0,T],L^q(\mathbb{R}^d))}$.

Proof. 1. By (2.20), one concludes

$$\|v(\cdot,0)\|_{L^\infty(\mathbb{R}^d)} \leq \|f\|_{L^b([0,T],L^q(\mathbb{R}^d))} \|u\|_{L^c([0,T],L^p(\mathbb{R}^d))}.$$

Then, by Theorem 2.1,

$$\|u\|^c_{L^c([0,T],L^p(\mathbb{R}^d))} = \int_0^T \|u\|^c_{L^p(\mathbb{R}^d)}\,dt \leq C\int_0^T t^{-\frac{cd(p-1)}{p(2-d(q-1))}}\,dt,$$

which is finite if and only if (2.21) holds. Hence, the estimate follows.

2. The proof is analogous, using Theorem 2.2. □

2.3 The heat equation

Here the methods from the previous Section are applied to the homogeneous heat equation, which corresponds to (2.9) with $b=0$ and $a=1$:

$$\begin{cases} u_t(x,t) = \Delta u(x,t) & \text{in } \Omega \times (0,\infty) \\ u(x,0) = u_0(x) & \text{in } \Omega. \end{cases} \tag{2.22}$$

We consider the cases where $\Omega = \mathbb{R}^d$ and $\Omega = \mathbb{T}^d$.

2.3.1 Main estimate

We now give an estimate for the L^p-norm of a derivative of any order of the solution of the heat equation.

Theorem 2.4. *Let u solve (2.22) with $u \in C^\infty(\Omega \times [0,\infty))$. Then, there exists $T > 0$ such that, for any $k \in \mathbb{N}_0$, $p > 1$, the following estimate holds*

$$||D^k u||_{L^p(\Omega)} \leq C||u_0||_{L^1(\Omega)} t^{-\frac{dp+kp-d}{2p}} \tag{2.23}$$

for all $t > 0$ with $\Omega = \mathbb{R}^d$ and for $t \in [0,T)$ with $\Omega = \mathbb{T}^d$. For $t > T$, the norm is bounded.

Proof. Fix $\gamma = p/2$. Then,

$$\frac{d}{dt}\int_\Omega |D^k u|^p \, dx = C\int_\Omega |D^k u|^{p-2} D^k u D^k \Delta u \, dx = -C\int_\Omega |\nabla(|D^k u|^\gamma)|^2 \, dx.$$

For $\Omega = \mathbb{R}^d$, on account of the Sobolev and Gagliardo-Nirenberg inequalities, it is clear that

$$C\left(\int_{\mathbb{R}^d} |\nabla(|D^k u|^\gamma)|^2 \, dx\right)^{\frac{\lambda}{2}} \geq ||D^k u||^{\gamma\lambda}_{L^{2^*\gamma}(\mathbb{R}^d)} \geq ||D^k u||^{\gamma}_{L^p(\mathbb{R}^d)} ||u_0||^{\gamma(\lambda-1)}_{L^1(\mathbb{R}^d)},$$

where $\lambda = \frac{d(p-1)+kp}{2+d(p-1)+kp}$ satisfies

$$\frac{1}{p} = 1 - \lambda + \frac{k}{d} + \lambda\left(\frac{1}{2^*\gamma} - \frac{k}{d}\right).$$

Then, with $z(t) = \int_{\mathbb{R}^d} |D^k u|^p \, dx$, it follows the inequality

$$\dot{z} \leq -C||u_0||^{\frac{2\gamma(\lambda-1)}{\lambda}}_{L^1(\mathbb{R}^d)} z^{\frac{1}{\lambda}}.$$

Thus, by Lemma 2.1, $z(t) \leq C\|u_0\|^p_{L^1(\mathbb{R}^d)} t^{\frac{1}{1-1/\lambda}}$; (2.23) follows from

$$\frac{1}{1/\lambda - 1} = \frac{1}{2}(d(p-1)+kp).$$

For $\Omega = \mathbb{T}t^d$, the Gagliardo-Nirenberg inequality for bounded domains yields

$$\left(\int_{\mathbb{T}^d} |D^k u|^{2^*\gamma} dx\right)^{\frac{1}{2^*}} \leq C\left(\int_{\mathbb{T}^d} |D^k u|^{2\gamma} dx + \int_{\mathbb{T}^d} |\nabla(|D^k u|^{\gamma})|^2 dx\right)^{\frac{\alpha}{2}}.$$

Next, observe that by virtue of the Gagliardo-Nirenberg inequality,

$$\left(\int_{\mathbb{T}^d} |D^k u|^p dx\right)^{\frac{1}{p}} \leq C\left(\int_{\mathbb{T}^d} |D^k u|^{2^*\gamma} dx\right)^{\frac{\lambda}{2^*}} \|u_0\|_{L^1(\mathbb{T}t^d)},$$

where

$$\frac{1}{p} - \frac{k}{d} = 1 - \lambda + \lambda\left(\frac{1}{2^*\gamma} - \frac{k}{d}\right).$$

The preceding identity yields

$$\lambda = \frac{d(p-1)+kp}{2+d(p-1)+kp}.$$

Setting $z(t) = \int_{\mathbb{T}^d} |D^k u|^p dx$, the following differential inequality follows:

$$\dot{z} \leq C_1 z - C_2 \|u_0\|^{\gamma\frac{\lambda-1}{\lambda}}_{L^1(\mathbb{T}^d)} z^{\frac{\gamma}{\lambda p}}.$$

Hence, by Lemma 2.2, there exists $T > 0$ such that z satisfies

$$z(t) \leq C\|u_0\|^p_{L^1(\mathbb{T}^d)} t^{\frac{1}{1-1/\lambda}} = C\|u_0\|^p_{L^1(\mathbb{T}^d)} t^{-\frac{1}{2}(d(p-1)+kp)}$$

for $t \in [0,T)$. Thus, we get a similar estimate for $\mathbb{T}^d$. Also, by the same lemma, the norm is bounded for $t > T$. □

Remark 2.2. Comparing again with the fundamental solution, one has

$$\int_{\mathbb{R}^d} |D^k\Phi|^p dx \leq Ct^{-\frac{dp}{2}-\frac{kp}{2}} \int_{\mathbb{R}^d} e^{-C\frac{p|x|^2}{t}} dx = Ct^{-\frac{dp+kp-d}{2}},$$

which is the same estimate as (2.23). Hence, our estimates are sharp.

In the following two Sections, this method is compared with two alternative approaches: the entropy and hypercontractivity methods.

2.3.2 Entropy methods

We follow the discussion in [8] for the Fokker-Planck equation and present the entropy method applied to the heat equation. We define the entropy

$$H(t) = \int_{\mathbb{R}^d} \phi(u)\, dx,$$

where u solves (2.22) and ϕ is a convex function. Integration by parts yields

$$\dot{H}(t) = \frac{d}{dt}\int_{\mathbb{R}^d} \phi(u)\, dx = -\int_{\mathbb{R}^d} \phi''(u)|\nabla u|^2\, dx \leq 0.$$

Furthermore,

$$\ddot{H}(t) = -\int_{\mathbb{R}^d} \phi^{(3)}(u)u_t|\nabla u|^2 + 2\phi''(u)\nabla u \cdot \nabla(u_t)\, dx = I_1 + I_2,$$

where

$$I_1 = -\int_{\mathbb{R}^d} \phi^{(3)}(u)u_t|\nabla u|^2\, dx = \int_{\mathbb{R}^d} \phi^{(4)}(u)|\nabla u|^4 + 2\phi^{(3)}(u)\Delta u|\nabla u|^2\, dx$$

and

$$I_2 = -2\int_{\mathbb{R}^d} \phi''(u)\nabla u \cdot \nabla(u_t)\, dx = 2\int_{\mathbb{R}^d} \phi^{(3)}(u)\Delta u|\nabla u|^2 + \phi''(u)(\Delta u)^2\, dx.$$

Hence,

$$\ddot{H}(t) = \int_{\mathbb{R}^d} \phi^{(4)}(u)|\nabla u|^4 + 4\phi^{(3)}(u)\Delta u|\nabla u|^2 + 2\phi''(u)(\Delta u)^2\, dx.$$

We now set $\phi(u) = u^2$. Accordingly,

$$\ddot{H}(t) = 4\int_{\mathbb{R}^d} (\Delta u)^2\, dx.$$

Hence, for some constant, $C > 0$, the Gagliardo-Nirenberg inequality yields

$$\ddot{H}(t) = 4\int_{\mathbb{R}^d} (\Delta u)^2\, dx \geq C\left(\int_{\mathbb{R}^d} |\nabla u|^2\, dx\right)^{\alpha} = C(-\dot{H}(t))^{\alpha},$$

where α satisfies $\frac{1}{2} = \frac{2}{d} + \left(\frac{1}{2} - \frac{1}{d}\right)\alpha + 1 - \alpha$. Hence,

$$z(t) = -\dot{H}(t)$$

satisfies the following differential inequality

$$\dot{z} \leq -Cz^{\alpha}.$$

Hence, as before, $\dot{H}$ satisfies

$$|\dot{H}(t)| \leq Ct^{\frac{1}{1-\alpha}}$$

and thus, for some C depending on α,

$$\int_{\mathbb{R}^d} u^2\,dx = H(t) \leq Ct^{1+\frac{1}{1-\alpha}} = Ct^{-\frac{d}{2}},$$

which is the same estimate as the one obtained in Theorem 2.4 with $k = 0$ and $p = 2$. We have thus shown that our technique yields results similar to those obtained by means of entropy methods..

2.3.3 On logarithmic Sobolev inequalities and hypercontractivity

The gain of regularity in time can also be understood using the results in [2, 7, 11] on logarithmic Sobolev inequalities and hypercontractivity. Contractivity principles, which appear in quantum field theory, are often used to describe operators such as contractions between Lebesgue spaces, the case from L^p to L^q when $p \leq q$ being of particular interest.

Next, we state a result from [5] that yields a generalization of the logarithmic Sobolev inequality presented in [7]. First, we recall that the Fenchel-Legendre transform of a convex function φ is the function $\varphi^* : \mathbb{R}^d \to \mathbb{R}$ given by

$$\varphi^*(\mu) = \sup_{x \in \mathbb{R}^d} \{\mu \cdot x - \varphi(x)\}.$$

Proposition 2.1 (Gentil-Gross). *Let φ be a C^1 strictly convex function on $\mathbb{R}^d$ such that*

$$\lim_{|x| \to +\infty} \frac{\varphi(x)}{||x||} = +\infty.$$

Then, for all $\lambda > 0$ and for any smooth function g on $\mathbb{R}^d$, we have the following Euclidean logarithmic Sobolev inequality

$$\int_{\mathbb{R}^d} e^g \log\left(\frac{e^g}{\int_{\mathbb{R}^d} e^g\,dx}\right) dx \leq -d\log(\lambda e) \int_{\mathbb{R}^d} e^g\,dx + \int_{\mathbb{R}^d} \varphi^*(-\lambda \nabla g) e^g\,dx. \tag{2.24}$$

We begin by considering a time-dependent Lebesgue norm. More specifically, we are interested in bounding

$$\|u\|_{L^{s(t)}(\mathbb{R}^d)} = \left(\int_{\mathbb{R}^d} u^{s(t)}\,dx\right)^{\frac{1}{s(t)}},$$

where $1 \leq s(t) < \infty$.

Proposition 2.2. *Let u be a solution to the d-dimensional heat equation (2.22). Assume that $1 \leq s(t) < \infty$ is a nondecreasing function, with $s(0) = p \geq 1$ and such that*

$$s(t) = 1 + (p-1)e^{\frac{2t}{\lambda^2}}, \tag{2.25}$$

where $\lambda = e^{-1}$. Then, the following estimate holds for all $t : t > 0$:

$$\|u\|_{L^{s(t)}(\mathbb{R}^d)} \leq \|u_0\|_{L^p(\mathbb{R}^d)}. \tag{2.26}$$

Proof. Let $s \equiv s(t)$. As before, we have that

$$\frac{d}{dt}\int_{\mathbb{R}^d} u^s\,dx = s\int_{\mathbb{R}^d} u^{s-1}\Delta u\,dx + \dot{s}\int_{\mathbb{R}^d} u^s \log u\,dx. \tag{2.27}$$

Thus,

$$\begin{aligned} s\int_{\mathbb{R}^d} u^{s-1}\Delta u\,dx &= -s(s-1)\int_{\mathbb{R}^d} u^{s-2}|\nabla u|^2\,dx \\ &= -\frac{4(s-1)}{s}\int_{\mathbb{R}^d} |\nabla(u^{\frac{s}{2}})|^2\,dx \leq 0. \end{aligned} \tag{2.28}$$

Fix $g = \log(u^s)$ in (2.24) to get

$$\int_{\mathbb{R}^d} u^s \log\left(\frac{u^s}{\int_{\mathbb{R}^d} u^s\,dx}\right)dx \leq -d\log(\lambda e)\int_{\mathbb{R}^d} u^s\,dx + \int_{\mathbb{R}^d} \varphi^*(-\lambda\nabla\log(u^s))u^s\,dx.$$

Taking $\lambda = e^{-1}$, we estimate the second term on the right-hand side of (2.27) as

$$\dot{s}\int_{\mathbb{R}^d} u^s\log u\,dx \leq \frac{\dot{s}}{s}\left[\int_{\mathbb{R}^d} \varphi^*(-\lambda\nabla\log(u^s))u^s\,dx + \log\left(\int_{\mathbb{R}^d} u^s\,dx\right)\int_{\mathbb{R}^d} u^s\,dx\right]. \tag{2.29}$$

Fix $\varphi(x) = \frac{|x|^2}{2}$. Then, $\varphi^*(\mu) = \frac{|\mu|^2}{2}$. Thus

$$\int_{\mathbb{R}^d} \varphi^*(-\lambda\nabla\log(u^s))u^s\,dx = \frac{1}{2}(\lambda s)^2\int_{\mathbb{R}^d} |\nabla u|^2 u^{s-2}\,dx = 2\lambda^2\int_{\mathbb{R}^d} |\nabla(u^{\frac{s}{2}})|^2\,dx. \tag{2.30}$$

Combining (2.27), (2.28), (2.29) and (2.30), we obtain

$$\frac{d}{dt}\int_{\mathbb{R}^d} u^s\,dx \le (g(t)-f(t))\int_{\mathbb{R}^d} |\nabla(u^{\frac{s}{2}})|^2\,dx + \frac{\dot{s}}{s}\log\left(\int_{\mathbb{R}^d} u^s\,dx\right)\int_{\mathbb{R}^d} u^s\,dx,$$

where

$$f(t) = \frac{4(s(t)-1)}{s(t)} \text{ and } g(t) = \frac{2\lambda^2 \dot{s}(t)}{s(t)}.$$

Now, we select $\dot{s} \ge 0$ such that

$$g - f = 0;$$

that is,

$$\dot{s} = \frac{2s-2}{\lambda^2}, \tag{2.31}$$

whose solution is (2.25). Hence, for s such that (2.31) holds, we have the following differential inequality

$$\frac{d}{dt}\int_{\mathbb{R}^d} u^s\,dx \le \frac{\dot{s}}{s}\log\left(\int_{\mathbb{R}^d} u^s\,dx\right)\int_{\mathbb{R}^d} u^s\,dx.$$

Fix $z(t) = \int_{\mathbb{R}^d} u^s\,dx$ and $h(t) = \frac{\dot{s}}{s} = \frac{d}{dt}\log(s(t))$. Thus, the previous inequality simplifies to

$$\dot{z}(t) \le h(t)\log(z(t))z(t).$$

The preceding inequality can be rewritten as

$$\frac{\dot{z}(t)}{\log(z(t))z(t)} \le h(t)$$

and thus

$$\frac{d}{dt}(\log(\log(z(t)))) \le \frac{d}{dt}\log(s(t)).$$

Finally, with $s(0) = p$, the integration of the above expression leads to

$$\log(\log(z(t))) \le \log(s(t)) + \log(\log(z(0))) - \log p$$

and thus

$$z(t) \le \exp\{\exp\{\log(s(t)) + \log(\log(z(0))) - \log p\}\} = z(0)^{\frac{s(t)}{p}} = \|u_0\|_{L^p(\mathbb{R}^d)}^{s(t)}.$$

Hence, (2.26) follows. $\square$

Remark 2.3. 1. If $s(0) = 1$ in the previous proposition, (2.25) forces $s(t) = 1$ for all $t \geq 0$, which makes (2.26) trivial.

2. For t such that $s(t) > 2$, interpolation yields, for some $\lambda(t)$,

$$\|\Phi\|_{L^2(\mathbb{R}^d)} \leq \|\Phi\|_{L^1(\mathbb{R}^d)}^{\lambda(t)} \|\Phi\|_{L^{s(t)}(\mathbb{R}^d)}^{1-\lambda(t)} \leq C.$$

By the estimate in Theorem 2.4, we have $\|\Phi\|_{L^2(\mathbb{R}^d)} \leq Ct^{-\frac{d}{4}}$. Hence, our estimate still yields a sharper result.

3. For the fundamental solution Φ of the heat equation, the above hypercontractivity result yields $\|\Phi\|_{L^{s(t)}(\mathbb{R}^d)} \leq C$, where C is a fixed constant. On the other hand, a direct estimate yields

$$\|\Phi\|_{L^{s(t)}(\mathbb{R}^d)} = (4\pi t)^{-\frac{d}{2}} \left(\int_{\mathbb{R}^d} e^{-\frac{s(t)|x|^2}{4t}} \, dx \right)^{\frac{1}{s(t)}} = s(t)^{-\frac{d}{2s(t)}} (4\pi t)^{-\frac{d(s(t)-1)}{2s(t)}}.$$

Since $s(t)^{-\frac{d}{2s(t)}} \to 1$ and $\frac{s(t)-1}{s(t)} \to 1$ as $t \to \infty$, we have that the hypercontractivity estimate does not provide information about the decay of the Lebesgue norms.

2.3.3.1 Estimate curves

We are now interested in finding a norm function, $s(t)$, for a specific estimate. We start by analyzing estimates for the fundamental solution. By Remark 2.3, for a fixed $a > 0$, the curve $s(t)$ such that $\|\Phi\|_{L^{s(t)}(\mathbb{R}^d)} = a$ is given implicitly by

$$s(t) = a^{-\frac{2s(t)}{d}} (4\pi t)^{1-s(t)}. \tag{2.32}$$

Figure 2.1, generated by using a numerical solver in *Mathematica*, with $d = 3$, shows the curve $s(t)$ for different time intervals and values of a.

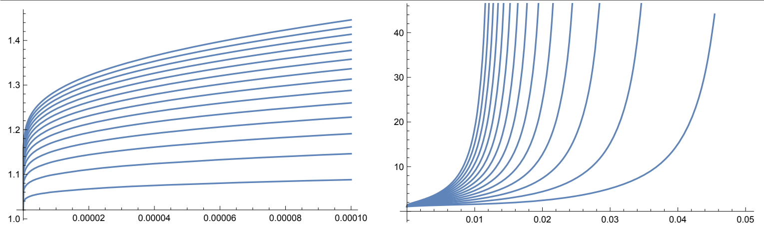


FIGURE 2.1: $s(t)$ paths for $2 \leq a \leq 15$ up to $t = 0.0001$ and $t = 0.05$

Here, we are considering solutions of (2.32) such that $s(t) \geq 1$. Such solutions only occur up to a certain time T_a, depending on a, which defines a vertical asymptote of $s(t)$. Using *Mathematica* again, we conclude that, for any dimension, T_a is given explicitly by $T_a = 1/(4\pi a^{\frac{2}{d}})$. Next, we deduce a similar estimate for the curves $\tilde{s}(t)$ regarding the result from Theorem 2.4 for general solutions of (2.22). Fixing $s(t) = p$ and $\gamma = p/2$, we have that, by Sobolev's inequality and interpolation,

$$\begin{aligned}\frac{d}{dt}\|u\|^p_{L^p(\mathbb{R}^d)} &= -\frac{4(p-1)}{p}\int_{\mathbb{R}^d}|D(u^\gamma)|^2\,dx \\ &\leq -\frac{4(p-1)}{pC_d^2}\left(\int_{\mathbb{R}^d}u^{2^*\gamma}\,dx\right)^{\frac{2}{2^*}} \\ &\leq -\frac{4(p-1)}{pC_d^2}\|u\|^{\frac{2\gamma(\lambda-1)}{\lambda}}_{L^1(\mathbb{R}^d)}\left(\int_{\mathbb{R}^d}u^p\,dx\right)^{\frac{1}{\lambda}},\end{aligned}$$

where $\lambda = \frac{d(p-1)}{2+d(p-1)}$ and C_d is the constant in the Sobolev's inequality, which only depends on the dimension. By [13], the sharp Sobolev's constant is given explicitly by

$$C_d = (\pi d(d-2))^{-\frac{1}{2}}\left(\frac{\Gamma(d)}{\Gamma\left(\frac{d}{2}\right)}\right)^{\frac{1}{d}}.$$

Then, as in the proof of Theorem 2.4, we have that

$$\|u\|_{L^p(\mathbb{R}^d)} \leq \left(\frac{4(p-1)(1/\lambda-1)}{pC_d^2}\right)^{\frac{1}{p(1-1/\lambda)}}\|u\|_{L^1(\mathbb{R}^d)}t^{\frac{1}{p(1-1/\lambda)}}.$$

Now, for a fixed a, the curve $\tilde{s}(t)$ such that $\|u\|_{L^{\tilde{s}(t)}(\mathbb{R}^d)} \leq a$ is given implicitly by

$$2^{-\frac{3d(\tilde{s}(t)-1)}{2\tilde{s}(t)}}\left(\frac{1}{\tilde{s}(t)}(d-2)\pi^{1+\frac{1}{d}}\left(2^{d-1}\Gamma\left(\frac{d+1}{2}\right)\right)^{-\frac{2}{d}}\right)^{-\frac{d(\tilde{s}(t)-1)}{2\tilde{s}(t)}}\|u\|_{L^1(\mathbb{R}^d)}t^{-\frac{d(\tilde{s}(t)-1)}{2\tilde{s}(t)}} = a.$$

With $d = 3$ and $\|u\|_{L^1(\mathbb{R}^d)} = 1$, Figure 2.2 shows the curve $\tilde{s}(t)$ for different time intervals and values of a.

We now compare both norm curves. Fix a such that $\|\Phi\|_{L^{s(t)}(\mathbb{R}^d)} = a$. Figure 2.3 illustrates $\|\Phi\|_{L^{\tilde{s}(t)}(\mathbb{R}^d)}$, for different values of a.

Hence, for all $t > 0$, $\|\Phi\|_{L^{\tilde{s}(t)}} \leq a$ and norm decay is still verified. Furthermore, the nature of both norms near $t = 0$ is compared by studying the limit of $\frac{s(t)-1}{\tilde{s}(t)-1}$ as $t \to 0$. Figure 2.4 suggests that $\lim\limits_{t\to 0}\frac{s(t)-1}{\tilde{s}(t)-1} < \infty$, also indicating that $s(t)$ and $\tilde{s}(t)$ might have similar behavior near $t = 0$.

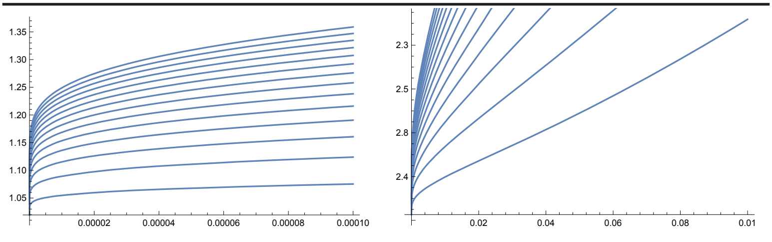


FIGURE 2.2: $\tilde{s}(t)$ paths for $2 \leq a \leq 15$ up to $t = 0.0001$ and $t = 0.05$

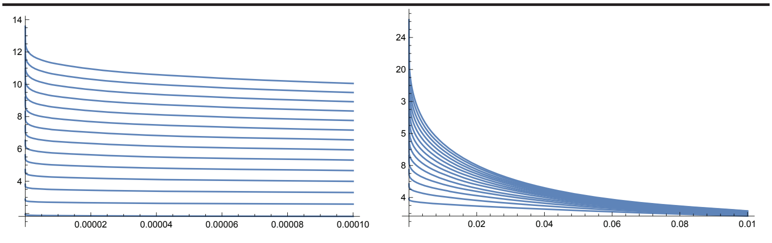


FIGURE 2.3: $\|\Phi\|_{L^{\tilde{s}(t)}(\mathbb{R}^d)}$ for $2 \leq a \leq 15$ up to $t = 0.0001$ and $t = 0.05$

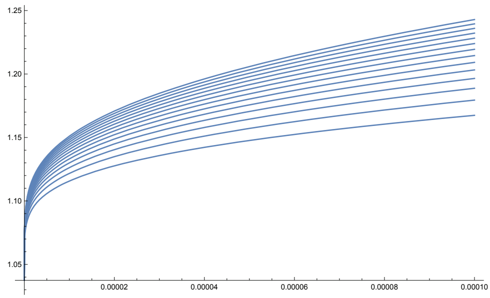


FIGURE 2.4: $\frac{s(t)-1}{\tilde{s}(t)-1}$

2.4 The porous media equation

The porous media equation (PME) is the following PDE

$$\begin{cases} u_t(x,t) = \Delta(u(x,t)^m) & \text{in } \mathbb{R}^d \times (0,T) \\ u(x,0) = u_0(x) & \text{in } \mathbb{R}^d, \end{cases} \tag{2.33}$$

for some $m \in [1,\infty)$, $u \geq 0$. Note that $m = 1$ corresponds to the heat equation. Here, we extend the ideas from the previous sections to obtain integrability estimates for the solution of the PME. Next, we examine the Barenblatt solutions to show that our bounds are sharp. We conclude this section by comparing our method with the results in [15].

2.4.1 Estimate methods revisited

We begin by applying our method to (2.33).

Theorem 2.5. *Let u solve* (2.33) *with $u \in C^\infty(\mathbb{R}^d \times [0,\infty))$. Then, for $p \geq 1$, the following estimate holds:*

$$\|u\|_{L^p(\mathbb{R}^d)} \leq C\|u_0\|_{L^1(\mathbb{R}^d)}^{\frac{d(m-1)+2p}{p(d(m-1)+2)}} t^{-\frac{d(p-1)}{p(d(m-1)+2)}} \tag{2.34}$$

for all $t > 0$.

Proof. Start by noticing that

$$\frac{d}{dt}\int_{\mathbb{R}^d} u^p\,dx = p\int_{\mathbb{R}^d} u^{p-1}\Delta(u^m)\,dx = -mp(p-1)\int_{\mathbb{R}^d} u^{m+p-3}|\nabla u|^2\,dx \leq 0. \tag{2.35}$$

Fix $\gamma = (m+p-1)/2$. Then, (2.35) yields

$$\frac{d}{dt}\int_{\mathbb{R}^d} u^p\,dx = -C\int_{\mathbb{R}^d} u^{2\gamma-2}|\nabla u|^2\,dx = -C\int_{\mathbb{R}^d} |\nabla(u^\gamma)|^2\,dx. \tag{2.36}$$

On account of the Sobolev inequality, it follows that

$$\left(\int_{\mathbb{R}^d} u^{2^*\gamma}\,dx\right)^{\frac{1}{2^*}} \leq C\left(\int_{\mathbb{R}^d} |\nabla(u^\gamma)|^2\,dx\right)^{\frac{1}{2}}. \tag{2.37}$$

Using the interpolation inequality it is clear that

$$\begin{aligned}\left(\int_{\mathbb{R}^d} u^{2^*\gamma}\,dx\right)^{\frac{2\lambda}{2^*}} &= \|u\|_{L^{2^*\gamma}(\mathbb{R}^d)}^{2\gamma\lambda} = \|u\|_{L^{2^*\gamma}(\mathbb{R}^d)}^{2\gamma\lambda} \\ &\geq \|u\|_{L^p(\mathbb{R}^d)}^{2\gamma}\|u_0\|_{L^1(\mathbb{R}^d)}^{2\gamma(\lambda-1)},\end{aligned} \tag{2.38}$$

where $\lambda = \frac{d(p-1)(m+p-1)}{p(2+d(m+p-2))}$. Hence, (2.36), (2.37) and (2.38) lead to

$$\frac{d}{dt}\int_{\mathbb{R}^d} u^p\,dx \leq -C\|u_0\|_{L^1(\mathbb{R}^d)}^{\frac{2\gamma(\lambda-1)}{\lambda}}\left(\int_{\mathbb{R}^d} u^p\,dx\right)^{\frac{2\gamma}{\lambda p}}.$$

Let $z(t) = \int_{\mathbb{R}^d} u^p\,dx$. Then, the previous inequality can be written as $\dot{z} \leq -C\|u_0\|_{L^1(\mathbb{R}^d)}^{\frac{2\gamma(\lambda-1)}{\lambda}} z^{\beta}$, where $\beta = 2\gamma/(\lambda p) > 1$. As before, we get the following time estimate

$$z(t) \leq C\|u_0\|_{L^1(\mathbb{R}^d)}^{\frac{2\gamma(\lambda-1)}{\lambda(1-\beta)}} t^{\frac{1}{1-\beta}} = C\|u_0\|_{L^1(\mathbb{R}^d)}^{\frac{d(m-1)+2p}{p(d(m-1)+2)}} t^{\frac{d(1-p)}{d(m-1)+2}}.$$

Thus, (2.34) follows. □

Next, we consider an estimate for a known solution to (2.33) and compare it to the prior estimate.

2.4.2 Barenblatt solutions

The Barenblatt solution of the PME has the following explicit formula, for an arbitrary constant $C > 0$:

$$\mathscr{U}(x,t) = t^{-\alpha}(C - k|x|^2 t^{-2\sigma})_+^{\frac{1}{m-1}},$$

where $(s)_+ = \max\{s, 0\}$ and

$$\alpha = \frac{d}{d(m-1)+2}, \ \sigma = \frac{\alpha}{d}, \ k = \frac{\alpha(m-1)}{2md}.$$

Denote the ball centered at the origin with radius $R = (Ct^{2\sigma}/k)^{\frac{1}{2}}$ by B_R. Then, with $u = \mathscr{U}$, we have

$$\begin{aligned}
\int_{\mathbb{R}^d} \mathscr{U}^p\,dx &= \int_{B_R} \mathscr{U}^p\,dx = t^{-p\alpha}\int_{B_R}(C - k|x|^2 t^{-2\sigma})^{\frac{p}{m-1}}\,dx \\
&= t^{-p\alpha}\int_{B_R}(C - k|y|^2)^{\frac{p}{m-1}} t^{\sigma d}\,dy \\
&= C_{m,p,k} t^{-p\alpha+\sigma d} = C_{m,p,k} t^{\alpha(1-p)} = C_{m,p,k} t^{-\frac{d(p-1)}{d(m-1)+2}},
\end{aligned}$$

where we considered the change of variables $y = x/t^{\sigma}$, with $dx = t^{\sigma d}\,dy$. Then, by comparison with (2.34), we conclude that our estimate is sharp.

2.4.3 Comparison with previous work

We now compare the results of our method with estimates in the literature. In [14], using phase-plane analysis, scaling techniques, and self-similarity, it was shown that

$$\|u\|_{L^p(\mathbb{R}^d)} \leq C\|u_0\|_{L^q(\mathbb{R}^d)}^{\sigma(p,q)} t^{-\alpha(p,q)}$$

with

$$\alpha(p,q)=\frac{d(p-q)}{p(d(m-1)+2q)},\ \sigma(p,q)=\frac{q(d(m-1)+2p)}{p(d(m-1)+2q)}.$$

In our case, we fix $q=1$ to get

$$\|u\|_{L^p(\mathbb{R}^d)}\leq C\|u_0\|_{L^1(\mathbb{R}^d)}^{\sigma(p,1)}t^{-\alpha(p,1)}=C\|u_0\|_{L^1(\mathbb{R}^d)}^{\frac{d(m-1)+2p}{p(d(m-1)+2)}}t^{-\frac{d(p-1)}{p(d(m-1)+2)}}$$

which yields the same estimate as in (2.34). Hence, our technique provides a different method to establish the results in [14] without relying on symmetry arguments.

2.4.4 Periodic solutions of the porous media equation

We now the deduce a similar estimate for the porous media equation on $\mathbb{T}^d$.

Proposition 2.3. *Let u solve* (2.33) *on the torus with $u\in C^\infty(\mathbb{T}^d\times[0,\infty))$. Then, there exists $T>0$ such that the following holds*

$$\|u\|_{L^p(\mathbb{T}^d)}\leq C\|u_0\|_{L^1(\mathbb{T}^d)}^{\frac{d(m-1)+2p}{p(d(m-1)+2)}}t^{-\frac{d(p-1)}{p(d(m-1)+2)}} \tag{2.39}$$

for all $t\in[0,T)$. For $t>T$, $\|u\|_{L^p(\mathbb{T}^d)}\leq C\|u_0\|_{L^1(\mathbb{T}^d)}^{\frac{d(m-1)+2p}{p(d(m-1)+2)}}T^{-\frac{d(p-1)}{p(d(m-1)+2)}}$.

Proof. Fix $\gamma=(m+p-1)/2$, thus $2\gamma>p$. From (2.36), we have that $\frac{d}{dt}\int_{\mathbb{R}^d}u^p\,dx=-C\int_{\mathbb{R}^d}|\nabla(u^\gamma)|^2\,dx$. Then,

$$\begin{aligned}\left(\int_{\mathbb{T}^d}u^p\,dx\right)^{\frac{\gamma}{p}}&\leq\|u_0\|_{L^1(\mathbb{T}^d)}^{\gamma(1-\lambda)}\left(\int_{\mathbb{T}^d}u^{2^*\gamma}\,dx\right)^{\frac{\lambda}{2^*}}\\&\leq\|u_0\|_{L^1(\mathbb{T}^d)}^{\gamma(1-\lambda)}\left(\|u_0\|_{L^1(\mathbb{T}^d)}+\int_{\mathbb{T}^d}|D(u^\gamma)|^2\,dx\right)^{\frac{\lambda}{2}}\\&\leq\|u_0\|_{L^1(\mathbb{T}^d)}^{\gamma(1-\lambda)}\left(\|u_0\|_{L^1(\mathbb{T}^d)}-C\frac{d}{dt}\int_{\mathbb{T}^d}u^p\,dx\right)^{\frac{\lambda}{2}},\end{aligned}$$

where $\lambda=\frac{d(p-1)(m+p-1)}{p(2+d(m+p-2))}$. Then, fixing $z(t)=\int_{\mathbb{T}^d}u^p\,dx$, we get the following differential inequality $\dot{z}\leq C_1\|u_0\|_{L^1(\mathbb{T}^d)}-C_2\|u_0\|_{L^1(\mathbb{T}^d)}^{\frac{2\gamma(\lambda-1)}{\lambda}}z^\beta$, where $\beta=\frac{2\gamma}{\lambda p}$. Hence, by Lemma 2.2, there exists $T>0$ such that z satisfies

$$z(t)\leq C\|u_0\|_{L^1(\mathbb{T}^d)}^{\frac{d(m-1)+2p}{p(d(m-1)+2)}}t^{\frac{d(1-p)}{d(m-1)+2}}$$

for $t \in [0,T)$, which yields (2.39). For $t > T$,

$$\|u\|_{L^p(\mathbb{T}^d)} \le C\|u_0\|_{L^1(\mathbb{T}^d)}^{\frac{d(m-1)+2p}{p(d(m-1)+2)}} T^{-\frac{d(p-1)}{p(d(m-1)+2)}}.$$

□

2.5 Differential inequalities

In this appendix, we present some of the estimates related to the differential inequalities used here.

Lemma 2.1. *Let $z : (0,\infty) \to (0,\infty)$ be a differentiable function satisfying the differential inequality*

$$z'(t) \le -Cz(t)^\beta \tag{2.40}$$

for some constant $C > 0$ and $\beta > 1$. Then, z satisfies

$$z(t) \le C_\beta t^{\frac{1}{1-\beta}}$$

for all $t > 0$.

Proof. Let $z \equiv z(t)$ and $\dot{z} \equiv z'(t)$. Since $\beta - 1 > 0$, multiplying both sides of (2.40) by $-(\beta-1)z^{-\beta}$ one gets $-(\beta-1)z^{-\beta}\dot{z} \ge (\beta-1)C$. Next, we observe that the left-hand side in the preceding equation is $\frac{d}{dt}(z(t)^{1-\beta})$. Hence, integrating in time, we get

$$z(t)^{1-\beta} \ge z(0)^{1-\beta}(1+z(0)^{\beta-1}(\beta-1)Ct).$$

Therefore,

$$\begin{aligned} z(t) &\le \frac{z(0)}{(1+z(0)^{\beta-1}(\beta-1)Ct)^{\frac{1}{\beta-1}}} \le \frac{1}{(z(0)^{1-\beta}+(\beta-1)Ct)^{\frac{1}{\beta-1}}} \\ &\le \frac{1}{((\beta-1)Ct)^{\frac{1}{\beta-1}}}. \end{aligned}$$

Hence, since $0 < z(0) < \infty$, z satisfies $z(t) \le C_\beta t^{\frac{1}{1-\beta}}$ for some constant $C_\beta > 0$, depending on β. □

Lemma 2.2. *Let $z : (0,\infty) \to (0,\infty)$ be a differentiable function satisfying the differential inequality*

$$\dot{z} \leq C_1 z^\theta - C_2 z^\beta \tag{2.41}$$

for constants $C_1, C_2 > 0$, and $1 \leq \theta < \beta$. Then, there exists $T > 0$ such that

$$z(t) \leq C_\beta t^{\frac{1}{1-\beta}}$$

for $t \in (0,T)$. Moreover, for $t > T$, $z(t) \leq C_\beta T^{\frac{1}{1-\beta}}$.

Proof. The function

$$z \mapsto C_1 z^\theta - C_2 z^\beta$$

has a single positive zero $\bar{z}$. Fix $z_0 > \bar{z}$ such that

$$C_1 (z)^\theta - C_2 (z)^\beta < -\tilde{C} z^\beta$$

for $z > \tilde{z}$. Consider the solution $z_*(t)$ of

$$\dot{z}_* = -\tilde{C} z_*^\beta$$

defined on $(0,+\infty)$ with $\lim\limits_{t\to 0} z_*(t) = +\infty$. Define T by

$$z_*(T) = \tilde{z}.$$

Then, if z satisfies (2.41), we have $z(t) \leq z_*(t)$ for $t \leq T$ and $z_*(t) \leq \tilde{z}$ for $t \geq T$. Thus, by computing z_* and then $\tilde{z}$ as a function of T, we conclude that $z(t) \leq C t^{\frac{1}{1-\beta}}$ for all $t \in (0,T)$ and $z(t) \leq C T^{\frac{1}{1-\beta}}$ for $t > T$. □

Bibliography

[1] A. Arnold, J. Dolbeault, *Refined Convex Sobolev Inequalities*. Angewandte Mathematik und Informatik. Univ., 2005.

[2] E. B. Davies, *Heat kernels and spectral theory*, volume 92. Cambridge University press, 1990.

[3] L. C. Evans, *A survey of partial differential equations methods in weak KAM theory*. Communications on pure and applied mathematics, 57(4):445–480, 2004.

[4] L. C. Evans, *Adjoint and compensated compactness methods for Hamilton–Jacobi pde*. Archive for rational mechanics and analysis, 197(3):1053–1088, 2010.

[5] I. Gentil, *Logarithmic Sobolev inequality for log-concave measure from Prekopa-Leindler inequality*. arXiv preprint math/0503476, 2005.

[6] D. A. Gomes, E. A. Pimentel, V. Voskanyan, *Regularity theory for mean-field game systems*. Springer, 2016.

[7] L. Gross, *Logarithmic Sobolev inequalities*. American Journal of Mathematics, 97(4):1061–1083, 1975.

[8] A. Jüngel, *Entropy methods for diffusive partial differential equations*. Springer, 2016.

[9] M. Magliocca, *Regularizing effect and decay results for a parabolic problem with repulsive superlinear first order terms*. arXiv preprint arXiv:1712.09246, 2017.

[10] M. Magliocca, A. Porretta, *Local and global time decay for parabolic equations with super linear first order terms*. arXiv preprint arXiv:1707.01761, 2017.

[11] E. Nelson, *The free Markoff field*. Journal of Functional Analysis, 12(2):211–227, 1973.

[12] B. Simon, R. Høegh-Krohn, *Hypercontractive semigroups and two dimensional self-coupled bose fields*. Journal of Functional Analysis, 9(2):121–180, 1972.

[13] G. Talenti, *Best constant in Sobolev inequality*. Annali di Matematica pura ed Applicata, 110(1):353–372, 1976.

[14] J. L. Vázquez, *Smoothing and decay estimates for nonlinear diffusion equations: equations of porous medium type*, volume 33. Oxford University Press, 2006.

[15] J. L. Vázquez, *The porous medium equation: mathematical theory*. Oxford University Press, 2007.

Chapter 3

Degenerate elliptic equations and sum operators

G. Di Fazio
Dipartimento di Matematica e Informatica Università di Catania Viale A. Doria 6, 95125, Catania, Italy,
E-mail: giuseppedifazio@unict.it

M. S. Fanciullo
Dipartimento di Matematica e Informatica Università di Catania Viale A. Doria 6, 95125, Catania, Italy,
E-mail: fanciullo@dmi.unict.it

P. Zamboni
Dipartimento di Matematica e Informatica Università di Catania Viale A. Doria 6, 95125, Catania, Italy,
E-mail: zamboni@dmi.unict.it

This Chapter is devoted to a new kind of degenerate elliptic operator. It is shown that it is possible to derive a regularity theory for this class. Despite the strong degeneracy of the operator, the smoothness of the generalized solutions can be proved.

3.1 Introduction

The regularity of generalized solutions of elliptic PDEs is a very important issue that has received a lot of attention in the past decades (see [2], [4], [5],[24], [27], [28] and [31]) and elliptic operators that are degenerate deserve a deeper study. This is mainly due to the fact that many operators that appear in applications are not uniformly elliptic. Some of our previous contributions are contained in [7] [8] [9] [10] [13] and [37], where Harnack inequalities and Hölder continuity for solutions of degenerate equations have been proved. There, one of the main tools is given by suitable sub representation formulas. In [21] it has been proved that these formulas are not always available and in particular this is due to the given geometrical setting and to the validity of Poincaré's inequality.

In this note some regularity results are presented in those cases where representation formulas are not available. Indeed such kind of formulas imply embedding results that we use to control the effect of lower order terms containing very strong degeneracy and singularity (see [10], and [11] [16], [18], [10], and [11]). Unfortunately, our setting is not suitable for the validity of a $(1-1)$-Poincaré inequality as the following example shows.

Let us consider the space $\mathbb{R}^2$ equipped with the Euclidean metric and the measure μ generated by the density $d\mu(x) = |x_2|^t dx$ ($t > 0$). In this case, the $(1-p)$-Poincaré inequality holds true if and only if $p > t+1$ (see [25]).

The lack of validity of the $(1-1)$-Poincaré inequality implies that we cannot have an explicit representation formula like that of the Euclidean case for smooth functions in terms of a quantity related to the system of vector fields.

We overcome this difficult task by using a special geometry introduced in [21] by Franchi Perez and Wheeden. Then we may assume a different instance of Poincaré inequality, namely we assume a $(1-p)$-Poincaré inequality. This allows us to use a different representation formula in terms of chains of balls related to a given one and then we can prove the embedding we need (see also [11], [12], [14] and [15]).

We now briefly describe the contents of the present note.

Let us consider a given system of m first order locally Lipschitz vector fields in $\mathbb{R}^n$, i.e

$$X = (X_1, X_2, \dots, X_m), \qquad m < n$$

and the degenerate elliptic equation

$$-X_j^*(a_{ij}X_i u + d_j u) + \frac{b_0}{\lambda} w|Xu|^2 + b_i X_i u + cu = f - X_i^* h_i\,, \tag{3.1}$$

where w is a 2-admissible weight (see Section 2 for definition) and $\{a_{ij}(x)\}$ is a symmetric matrix of measurable functions in Ω satisfying the following ellipticity condition

$$\exists\lambda > 0 : \lambda^{-1} w(x)|\xi|^2 \le a_{ij}(x)\xi_i\xi_j \le \lambda w(x)|\xi|^2 \text{ a.e. } x \in \Omega\ \forall\xi \in \mathbb{R}^m\,. \tag{3.2}$$

The lower order terms are taken in suitable weighted Stummel classes (see section 3.2 for definitions). The reason for the name is the optimality of the Stummel classes in those cases in which subrepresentation formulas are available.

Our main goal here is the Harnack inequality for positive weak solutions of the equation (3.1) and our results are achieved by adapting classical iteration procedures to the present setting. First we consider solutions ot equation (3.1) with $b_0 = 0$, and closely following the proof by Serrin ([32]) we prove local boundedness and the Harnack inequality. Later we study the case $b_0 \neq 0$ and, following the proof by Trudinger ([33]) we prove that the local, bounded solutions satisfy the Harnack inequality. In both cases, as a direct consequence, we will get continuity and Hölder continuity of the weak solutions.

In the last section of this note we study the following quasilinear equation

$$\sum_{i=1}^{m} X_i^* A_i(x,u,Xu) + B(x,u,Xu) = 0 \tag{3.3}$$

where A and B are measurable functions. We consider two kinds of structural assumptions satisfied by A and B, involving p-admissible weights and coefficients in Stummel classes. We refer to these assumptions as controlled or natural growth respectively. As in the linear case, we obtain Harnack inequality for positive solutions of (3.3) and, consequently, the continuity and Hölder continuity.

3.2 Sum operators and underlying geometry

Let $X = (X_1, X_2, \ldots, X_m)$ be a system of locally Lipschitz vector fields in $\mathbb{R}^n$ and d the associated Carnot-Carathéodory distance . We assume that d is finite for each pair of points $x, y \in \mathbb{R}^n$. Let us denote by $B = B_r = B(x,r)$ the Carnot-Carathéodory ball centered at $x \in \mathbb{R}^n$ with radius r.

Throughout the paper the following assumptions are made:

(A1) The distance d is continuous with respect to the Euclidean distance in $\mathbb{R}^n$.

(A2) Let w be a finite Borel measure, absolutely continuous with respect to the Lebesgue measure. We assume the following doubling condition, : there exists a positive constant C_D such that

$$w(B(x,2r)) \leq C_D w(B(x,r)) \qquad \forall x \in \mathbb{R}^n, r > 0,$$

where $w(B(x,r)) = \int_{B(x,r)} w\,dy$.

(A3) $(1-p)$-*Poincaré inequality*. If B_0 is a given ball in $\mathbb{R}^n$, $p > 1$, there exists a positive constant C_P such that

$$\frac{1}{w(B)}\int_B |u-u_B|wdy \le C_P\, r\left(\frac{1}{w(B)}\int_B |Xu|^p wdy\right)^{1/p}$$

for all $B\subset B_0$ and all $u\in C^\infty(\overline{B}_0)$. Here $u_B=\frac{1}{w(B)}\int_B uw\,dy$ and r is the radius of B.

The number $Q=\log_2 C_D$ will be called the homogeneous dimension of $\mathbb{R}^n$. In the sequel we will need the Sobolev and the Stummel-type spaces with respect to the measure $w\,dx$.

Definition 3.1. Let Ω be a bounded domain in $\mathbb{R}^n$ and $p>1$. We say that u belongs to $W^{1,p}(\Omega,w)$ if u, $X_iu\in L^p(\Omega,w)$ for $i=1,\dots m$. We denote by $W_0^{1,p}(\Omega,w)$ the closure of the smooth, compactly supported functions in $W^{1,p}(\Omega,w)$ and furnished this space with the norm

$$\|u\|_{W^{1,p}(\Omega,w)} = \|u\|_{L^p(\Omega,w)}+\sum_{i=1}^m \|X_iu\|_{L^p(\Omega,w)}\,.$$

A function u is said to belong to $W^{1,p}_{loc}(\Omega,w)$ if $u\in W^{1,p}(\Omega',w)$ for any $\Omega'\Subset\Omega$.

In order to clarify the weights that are going to be considered in this discussion, recall the concept of p-admissible weight given in [25] (see also [26] and [20]).

Definition 3.2 (p-admissible weights). Let w be a nonnegative locally integrable function and $1<p<\infty$. w is a p-admissible weight if the following conditions are satisfied:

1. There exists a positive constant C_D such that

$$w(B(x,2r))\le C_D w(B(x,r)) \qquad \forall x\in\mathbb{R}^n, r>0$$

2. If Ω is an open set in $\mathbb{R}^n$ and $\{\varphi_i\}\subset C^\infty(\Omega)$ is a sequence such that $\|\varphi_i\|_{L^p(\Omega,w)}\to 0$ and $\|X\varphi_i - v\|_{L^p(\Omega,w)}\to 0$ for some $v\in L^p(\Omega,w)$, then $v\equiv 0$.

3. There exist two constant $C>0$ and $k>1$ such that

$$\left(\frac{1}{w(B)}\int_B |\varphi|^{kp}wdy\right)^{\frac{1}{kp}} \le Cr\left(\frac{1}{w(B)}\int_B |X\varphi|^p wdy\right)^{\frac{1}{p}}, \qquad (3.4)$$

for any $\varphi\in C_0^\infty(B)$.

4. There exists a constant C such that

$$\left(\frac{1}{w(B)}\int_B |\varphi-\varphi_B|^p w dy\right)^{\frac{1}{p}} \le Cr\left(\frac{1}{w(B)}\int_B |X\varphi|^p w dy\right)^{\frac{1}{p}}, \quad (3.5)$$

for any $\varphi \in C^\infty(B)$.

Remark 3.1. Thanks to Theorem 13.1 in [25], all weights considered here are p - admissible weights. Examples of admissible weights are A_p weights and suitable powers of Jacobians of quasiconformal mappings (see Corollary 15.34 in [26]).

Next, the geometry and related function spaces to be used in the sequel are introduced.

Definition 3.3. Given $B_0 = B(x_0, r)$ and $x \in B_0$, let $\{B_i\} = \{B_i(x)\}_{i=1}^\infty$ be a chain of balls of radius $r(B_i)$, such that

(H1) $B_i \subset B_0$ for all $i \ge 0$

(H2) $r(B_i) \sim 2^{-i} r(B_0)$ for all $i \ge 0$

(H3) $\rho(B_i, x) \le Cr(B_i)$ for all $i \ge 0$

(H4) for all $i \ge 0$, $B_i \cap B_{i+1}$ contains a ball S_i with $r(S_i) \sim r(B_i)$.

Remark 3.2. It follows from Remark 2 in [21] that a chain such as the one described in Definition 3.3 actually exists in the present setting.

Next, the Stummel and Morrey classes adapted to the present setting are introduced.

Definition 3.4. Let $p > 1$, $B_0 = B(x_0, r)$ be a ball and $\{B_j(x)\}_{j=1}^\infty$ be a chain of balls as in Definition 3.3. We say that $V \in L^1_{loc}(\mathbb{R}^n, w)$ belongs to the class $\tilde{S}_p(\mathbb{R}^n, w)$ if

$$\eta(V;r) \equiv \sup_{x_0\in\mathbb{R}^n} \sup_{y\in B_0} \int_{B_0} \sum_{j=0}^\infty \frac{r^p(B_j(x))|V(x)|}{w(B_j(x))} \chi_{B_j(x)}(y) w(x) dx$$

is finite for all $r > 0$. $V \in \tilde{S}_p(\mathbb{R}^n, w)$ is said to belong to $S_p(\mathbb{R}^n, w)$ if in addition, it holds that $\lim_{r\to 0} \eta(V;r) = 0$. One sets $V \in S'_p(\mathbb{R}^n, w)$ if there exists $\delta > 0$ such that

$$\int_0^\delta \frac{\eta(V;t)}{t} dt < +\infty.$$

V is said to belong to the Morrey class $M_\sigma(\mathbb{R}^n, w)$ if there exist C and $\sigma > 0$ such that $\eta(V;r) \le Cr^\sigma$.

We denote the Stummel classes $\tilde{S}_2(\mathbb{R}^n,w)$, $S_2(\mathbb{R}^n,w)$, $S_2'(\mathbb{R}^n,w)$ by $\tilde{S}(\mathbb{R}^n,w)$, $S(\mathbb{R}^n,w)$ and $S'(\mathbb{R}^n,w)$, respectively.

The following embedding Theorem is crucial for the developement of the main results in this Chapter. The unweighted case and some corollaries have been proved in [11] (see also [3], [16], [18], [29] [34], [35], [36]).

A fundamental tool to prove the required embedding is the subrepresentation formula proved by Franchi, Perez and Wheeden, in a more general setting (see [21]).

Theorem 3.1. *Given a ball B_0 let $\{B_j(x)\}_{j=1}^{\infty}$ be a chain of balls as in Definition 3.3, let $p>1$ and let w be a p-admissible weight. Let $u \in W^{1,p}(B_0,w)$ be such that for any ball $B \subset B_0$*

$$\frac{1}{w(B)}\int_B |u-u_B|w dx \le Cs\left(\frac{1}{w(B)}\int_B |Xu|^p w dy\right)^{1/p}$$

where s is the radius of B. Then there exists $C'>0$ such that for a. e. $x \in B_0$

$$|u(x)-u_{B_0}| \le C' \sum_{j=0}^{\infty} r(B_j(x))\left(\frac{1}{w(B_j(x))}\int_{B_j(x)} |Xu|^p w(y)dy\right)^{1/p}$$

where C' is a geometric constant which also depends on C.

Theorem 3.2. *Let B_0 be a ball in $\mathbb{R}^n$, let $p>1$, w be a p-admissible weight and V a function in $\tilde{S}_p(\mathbb{R}^n,w)$. Then, there exists a positive constant C such that*

$$\int_{B_0} |V(x)|\,|u(x)-u_{B_0}|^p w\,dx \le C\,\eta(V;r)\int_{B_0} |Xu(x)|^p w\,dx$$

for any $u \in C^{\infty}(B_0)$.

Proof. Let u be a smooth function in B_0. Theorem 3.1 yields the following representation formula for u

$$|u(x)-u_{B_0}| \le C \sum_{j=0}^{\infty} r(B_j(x))\left(\frac{1}{w(B_j(x))}\int_{B_j(x)} |Xu|^p w(y)dy\right)^{1/p} \tag{3.6}$$

for a.e. $x \in B_0$. Now from (3.6) and Hölder's inequality it follows that

$$\int_{B_0} |V(x)||u(x)-u_{B_0}|^p w(x)dx \leq$$

$$\leq \int_{B_0} |V(x)||u(x)-u_{B_0}| \cdot$$

$$\left\{ \sum_{j=0}^{\infty} r(B_j(x)) \cdot \left[\frac{1}{w(B_j(x))} \int_{B_j(x)} |Xu(y)|^p w(y)dy \right]^{1/p} \right\}^{p-1} w(x)dx \leq$$

$$\leq \left[\int_{B_0} |V(x)||u(x)-u_{B_0}|^p w(x)dx \right]^{1/p} \cdot$$

$$\cdot \left[\int_{B_0} \sum_{j=0}^{\infty} |V(x)| \frac{r^p(B_j(x))}{w(B_j(x))} \int_{B_j(x)} |Xu(y)|^p w(y)dyw(x)dx \right]^{\frac{p-1}{p}} \leq$$

$$\leq \left[\int_{B_0} |V(x)||u(x)-u_{B_0}|^p w(x)dx \right]^{1/p} \cdot$$

$$\cdot \left[\int_{B_0} \sum_{j=0}^{\infty} |V(x)| \frac{r^p(B_j(x))}{w(B_j(x))} \int_{B_0} |Xu(y)|^p \chi_{B_j(x)}(y)w(y)dyw(x)dx \right]^{\frac{p-1}{p}} \leq$$

$$\leq \left[\int_{B_0} |V(x)||u(x)-u_{B_0}|^p w(x)dx \right]^{1/p} \cdot$$

$$\cdot \left[\int_{B_0} |Xu(y)|^p \int_{B_0} \sum_{j=0}^{\infty} |V(x)| \frac{r^p(B_j(x))}{w(B_j(x))} \chi_{B_j(x)}(y)w(x)dxw(y)dy \right]^{\frac{p-1}{p}} \leq$$

$$\leq \left[\int_{B_0} |V(x)||u(x)-u_{B_0}|^p w(x)dx \right]^{1/p} \eta^{\frac{p-1}{p}}(V;r) \cdot$$

$$\cdot \left[\int_{B_0} |Xu(y)|^p w(y)dy \right]^{\frac{p-1}{p}}$$

from which one readily obtains

$$\int_{B_0} |V(x)||u(x)-u_{B_0}|^p w(x)dx \leq C\eta(V;r) \int_{B_0} |Xu(x)|^p w(x)dx.$$

$\square$

From Theorem 3.2, one gets the following two Corollaries.

Corollary 3.1. *Let $p > 1$ and let V be a function in $\tilde{S}_p(\mathbb{R}^n, w)$. Then there exists a positive constant C such that*

$$\int_{\mathbb{R}^n} |V(x)|\,|u(x)|^p w\,dx \leq C\,\eta(V;r) \int_{\mathbb{R}^n} |Xu(x)|^p w\,dx,$$

for any smooth function u, compactly supported in $\mathbb{R}^n$.

Corollary 3.2. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain and $p > 1$. Let V be a function in $S_p(\Omega, w)$. Then, for any $\varepsilon > 0$ there exists a positive function $K(\varepsilon) \sim \frac{\varepsilon}{[\eta^{-1}(V;\varepsilon)]^{Q+p}}$, where η^{-1} is the inverse function of $\eta(V;r)$, such that*

$$\int_{\Omega} |V(x)|\,|u(x)|^p w\,dx \leq \varepsilon \int_{\Omega} |Xu(x)|^p\,dx + K(\varepsilon) \int_{\Omega} |u(x)|^p\,dx \qquad (3.7)$$

for any smooth function u compactly supported in Ω.

Proof. Let $\varepsilon > 0$. Let r be a positive number to be determined later. Let $\{\alpha_i\}$, $i = 1, 2, \ldots N(r)$, be a finite partition of unity of $\overline{\Omega}$, such that $\operatorname{supp}\alpha_i \subset B(x_i, r)$ with $x_i \in \overline{\Omega}$ (for the construction of cut off functions α_i see [22]).

It follows from Corollary 3.1 that

$$\int_{\Omega} |V(x)|\,|u(x)|^p w\,dx \leq \int_{\Omega} |V(x)|\,|u(x)|^p \sum_{i=1}^{N(r)} \alpha_i^p(x) w\,dx =$$

$$= \sum_{i=1}^{N(r)} \int_{\Omega} |V(x)|\,|u(x)|^p \alpha_i^p(x) w dx \leq$$

$$\leq C \sum_{i=1}^{N(r)} \eta(V;r) \left(\int_{\Omega} |Xu(x)|^p \alpha_i^p(x) w\,dx + \int_{\Omega} |X\alpha_i(x)|^p |u(x)|^p w\,dx \right) \leq$$

$$\leq C\eta(V;r) \left(\int_{\Omega} |Xu(x)|^p w\,dx + \frac{N(r)}{r^p} |u(x)|^p w\,dx \right).$$

At this point, choose r such that $C\eta(V;r) = \varepsilon$; since $N(r) \sim r^{-Q}$ (3.7) follows at once. □

We recall a lemma (see Lemma 3.4 in [30]) useful in the sequel.

Lemma 3.1. *Let $\mu(r)$ a continuous, positive, increasing function defined on $]0, +\infty[$ such that $\lim\limits_{r \to 0} \mu(r) = 0$. Let $0 < \theta < 1$. Then the series*

$$\sum_{i=0}^{+\infty} \theta^i log\, \mu^{-1} \left(\theta^{qi} \right),$$

where $q > 0$, is convergent if and only if there exists $\rho > 0$ such that

$$\int_0^{\rho} \frac{\mu^{\frac{1}{q}}(t)}{t}\,dt < +\infty.$$

Proof. We claim that

$$\int_0^\rho \frac{\mu^{\frac{1}{q}}(t)}{t}\,dt < +\infty$$

if and only if the series

$$\sum_{i=0}^{+\infty}(\theta a_i - a_{i+1})$$

is convergent, where

$$a_i = \theta^i \log \mu^{-1}\left(\mu(\rho)\theta^{qi}\right).$$

Indeed

$$\int_0^\rho \frac{\mu^{\frac{1}{q}}(t)}{t}\,dt = \int_0^{\mu(\rho)} \frac{s^{\frac{1}{q}}}{\mu^{-1}(s)}\frac{1}{\mu'(\mu^{-1}(s))}\,ds =$$

$$= \sum_{i=0}^{+\infty}\int_{\mu(\rho)\theta^{q(i+1)}}^{\mu(\rho)\theta^{qi}} \frac{s^{\frac{1}{q}}}{\mu^{-1}(s)}\frac{1}{\mu'(\mu^{-1}(s))}\,ds <$$

$$< \sum_{i=0}^{+\infty}\Big\{\mu^{\frac{1}{q}}(\rho)\theta^i \log\mu^{-1}\left(\mu(\rho)\theta^{qi}\right) -$$

$$-\frac{1}{\theta}\mu^{\frac{1}{q}}(\rho)\theta^{i+1}\log\mu^{-1}\left(\mu(\rho)\theta^{q(i+1)}\right)\Big\} = \frac{\mu^{\frac{1}{q}}(\rho)}{\theta}\sum_{i=0}^{+\infty}(\theta a_i - a_{i+1}).$$

Analogously, it is possible to show that

$$\int_0^\rho \frac{\mu^{\frac{1}{q}}(t)}{t}\,dt > \sum_{i=0}^{+\infty}(\theta a_i - a_{i+1}).$$

Thus, the claim is proved.

Since the series $\sum\limits_{i=0}^{+\infty}(\theta a_i - a_{i+1})$ and $\sum\limits_{i=0}^{+\infty} a_i$ share the same character, the conclusion is obtained. □

The following Lemma will be used in the sequel (see Lemma 8.23 in [23]).

Lemma 3.2. *Let φ and σ be non-decreasing functions in $(0, r_0]$ such that*

$$\varphi(\tau r) \le \gamma\varphi(r) + \sigma(r) \qquad 0 < r \le r_0$$

and $0 < \gamma, \tau < 1$, Then, for any $\mu \in (0,1)$, one has

$$\varphi(r) \le C\left[\left(\frac{r}{r_0}\right)^\alpha \varphi(r_0) + \sigma(r^\mu r_0^{1-\mu})\right],$$

where $C = C(\gamma,\tau)$ and $\alpha = \alpha(\gamma,\tau,\mu)$ are positive constants.

3.3 Harnack inequality for linear and quasilinear degenerate elliptic equations

In this Section, Harnack inequality for degenerate elliptic equations will be proved under two different kinds of growth conditions and it will be established that generalized solutions are locally bounded and that the positive solutions satisfy a Harnack-type inequality. Then the case of natural growth will be discussed, where - in general - solutions can fail to be locally bounded. Nevertheless, it is still possible to prove regularity for those solutions that are a-priori locally bounded.

3.3.1 Degenerate equations under controlled growth

Let Ω be a bounded domain in $\mathbb{R}^n$. Let $X = (X_1, X_2, \ldots, X_m)$ be a system of locally Lipschitz vector fields in $\mathbb{R}^n$. For $i = 1, 2, \ldots, m$, X_i^* will stand for the formal adjoint of the vector fields X_i. It will be assumed that $(A3)$ is valid, with $p = 2$. Let w be a 2-admissible weight and $\{a_{ij}(x)\}$ be a symmetric matrix of measurable functions in Ω satisfying the ellipticity condition (3.2).

Consider the following elliptic linear equation in divergence form

$$-X_j^*(a_{ij}X_iu + d_ju) + b_iX_iu + cu = f - X_i^*h_i\,, \tag{3.8}$$

where

$$\left(\frac{b_i}{w}\right)^2, \frac{c}{w}, \left(\frac{d_i}{w}\right)^2, \frac{f}{w}, \left(\frac{h_i}{w}\right)^2 \in S'(\Omega, w)\,. \tag{3.9}$$

Next, the notion of weak solution of equation (3.8) is defined.

Definition 3.5. $u \in W^{1,2}_{loc}(\Omega, w)$ is said to be a local, weak solution of (3.8) if

$$\int_\Omega [(a_{ij}X_ju + d_ju)X_j\phi + (b_iX_iu + cu)\phi]\,dx = \int_\Omega (f\phi + h_iX_i\phi)dx\,, \tag{3.10}$$

$\forall \phi \in W_0^{1,2}(\Omega, w)$.

Notice that the integrals appearing in (3.10) are all finite because of (3.9). The reasoning presented here follows along the lines of [32]. First, we will prove the local boundedness of solutions of equation (3.8).

Theorem 3.3. *Let u be a local, weak solution of the equation* (3.8) *in* Ω. *We assume that the conditions* (3.2) *and* (3.9) *hold true. Then, there exists a* *positive constant C, independent of u, such that for any B_r for which $B_{4r} \subset \Omega$, we have*

$$\|u\|_{L^\infty(B_r)} \le C \left\{ \left(\frac{1}{w(B_{2r})} \int_{B_{2r}} |u|^2 w\,dx \right)^{\frac{1}{2}} + \right.$$
$$\left. + \eta\left(\frac{f}{w};3r\right) + \left(\sum_{i=1}^m \eta\left(\left(\frac{h_i}{w}\right)^2 ;3r \right) \right)^{1/2} \right\}.$$

Proof. Let B_r be a ball such that $B_{4r} \subset \Omega$ and

$$h = h(r) = \eta\left(\frac{f}{w};r\right) + \left(\sum_{i=1}^m \eta\left(\left(\frac{h_i}{w}\right)^2 ;r \right) \right)^{1/2}$$

and $l > h$. For $q \ge 1$, we consider the function

$$G(u) = \operatorname{sign} u \left\{ F(v)F'(v) - qh^{2q-1} \right\}, \quad u \in]-\infty,+\infty[,$$

where

$$v = |u| + h,$$

and

$$F(v) = \begin{cases} v^q & \text{if} \quad h \le v \le l \\ ql^{q-1}v - (q-1)l^q & \text{if} \quad l \le v. \end{cases}$$

Let $\phi(x) = \varphi^2(x)G(u)$, where $\varphi \in C_0^\infty(\Omega)$ is such that $0 \le \varphi \le 1$ and $\operatorname{supp}\varphi \subseteq B_{2r}$. From the definition of solution and following the proof of Theorem 3.1 in [37], one has

$$\sum_{i,j=1}^m \int_\Omega a_{ij} X_i u [2\varphi X_j \eta G(u) + \eta^2 G'(u) X_j u]\,dx +$$
$$+ \sum_{j=1}^m \int_\Omega d_j u [2\eta X_j \eta G(u) + \eta^2 G'(u) X_j u]\,dx +$$
$$+ \sum_{i=1}^m \int_\Omega b_i X_i u \eta^2 G(u)\,dx + \int_\Omega c u \eta^2 G(u)\,dx =$$
$$= \int_\Omega f \eta^2 G(u)\,dx + \sum_{i=1}^m \int_\Omega h_i [2\eta X_i \eta G(u) + \eta^2 G'(u) X_i u]\,dx.$$

It follows from condition (3.2) that

$$\lambda^{-1}\int_\Omega |Xu|^2\eta^2 G'(u)w\,dx \le 2\lambda\int_\Omega |Xu|\,|X\eta|\,\eta\,|G(u)|w\,dx+$$
$$+\sum_{j=1}^m\int_\Omega |d_j|\,|u|\,\eta^2\,|G'(u)|\,|X_ju|\,dx+2\sum_{j=1}^m\int_\Omega |d_j|\,|u|\,\eta\,|X_j\eta|\,|G(u)|\,dx+$$
$$+\sum_{i=1}^m\int_\Omega |b_i|\,|X_iu|\,\eta^2\,|G(u)|\,dx+\int_\Omega |c|\,|u|\,\eta^2\,|G(u)|\,dx+$$
$$+\int_\Omega |f|\,\eta^2\,|G(u)|\,dx+2\sum_{i=1}^m\int_\Omega |h_i|\,\eta\,|X_i\eta|\,|G(u)|\,dx+$$
$$+\sum_{i=1}^m\int_\Omega |h_i|\,\eta^2\,|X_iu|\,|G'(u)|\,dx.$$

Since $v=|u|+h$ and

$$G'=\begin{cases}(2-1/q)(F')^2 & \text{if}\quad |u|<l-h\\ (F')^2 & \text{if}\quad |u|>l-h,\end{cases}$$

$$|G|\le F(F'),$$
$$vF'\le qF,$$

it is clear that

$$\int_\Omega |Xv|^2\varphi^2(F')^2w\,dx \le 2\lambda^2\int_\Omega |Xv|\,|X\varphi|\,\varphi\,F\,(F')w\,dx+$$
$$+\left(2-\frac{1}{q}\right)q\lambda\sum_{j=1}^m\int_\Omega |d_j|\,\varphi^2\,F\,(F')\,|X_jv|\,dx+$$
$$+2q\lambda\sum_{j=1}^m\int_\Omega |d_j|\,\varphi\,|X_j\varphi|\,F^2\,dx+\lambda\sum_{i=1}^m\int_\Omega |b_i|\,|X_iv|\,\varphi^2\,F\,(F')\,dx+$$
$$+\lambda q\int_\Omega |c|\,\varphi^2\,F^2\,dx+h^{-1}\lambda q\int_\Omega |f|\,\varphi^2\,F^2\,dx+$$
$$+2q\lambda h^{-1}\sum_{i=1}^m\int_\Omega |h_i|\,\varphi\,|X_i\varphi|\,F^2\,dx+$$
$$+\left(2-\frac{1}{q}\right)q\lambda h^{-1}\sum_{i=1}^m\int_\Omega |h_i|\,\varphi^2\,|X_iv|\,F\,(F')\,dx.$$

One gets from the inequality

$$ab\le\frac{\varepsilon}{2}a^2+\frac{1}{2\varepsilon}b^2,\quad \varepsilon>0$$

that

$$\int_\Omega |Xv|^2\varphi^2(F')^2 w\,dx \le C(\lambda)q^2\left\{\int_\Omega |X\varphi|^2F^2w\,dx + \int_\Omega V\varphi^2F^2\,dx\right\}, \tag{3.11}$$

where

$$V = \sum_{i=1}^m \frac{|b_i|^2}{w} + |c| + \sum_{j=1}^m \frac{|d_j|^2}{w} + h^{-1}|f| + h^{-2}\sum_{i=1}^m \frac{|h_i|^2}{w}.$$

Note that $\frac{V}{w} \in S'(\Omega, w)$ and

$$\eta\left(\frac{V}{w};r\right) \le C\Big\{\sum_{i=1}^m \eta\left(\left(\frac{b_i}{w}\right)^2;r\right) + \eta\left(\frac{c}{w};r\right) + \\ + \sum_{i=1}^m \eta\left(\left(\frac{d_i}{w}\right)^2;r\right) + h^{-1}\eta\left(\frac{f}{w};r\right) + h^{-2}\sum_{i=1}^m \eta\left(\left(\frac{h_i}{w}\right)^2;r\right)\Big\},$$

then

$$\eta\left(\frac{V}{w};r\right) \le C\Big\{\sum_{i=1}^m \eta\left(\left(\frac{b_i}{w}\right)^2;r\right) + \eta\left(\frac{c}{w};r\right) + \sum_{i=1}^m \eta\left(\left(\frac{d_i}{w}\right)^2;r\right) + 2\Big\},$$

Set $\mathscr{U} = F(v)$. From (5.1) it follows that

$$\int_\Omega \varphi^2|X\mathscr{U}|^2w\,dx \le Cq^2\left\{\int_\Omega |X\varphi|^2\mathscr{U}^2w\,dx + \int_\Omega V\varphi^2\mathscr{U}^2\,dx\right\}.$$

From Corollary 3.2, one has

$$\int_\Omega \varphi^2|X\mathscr{U}|^2w\,dx \le Cq^2\Big\{(1+\varepsilon)\int_\Omega |X\varphi|^2\mathscr{U}^2w\,dx + \\ + \varepsilon\int_\Omega \varphi^2|X\mathscr{U}|^2w\,dx + K(\varepsilon)\int_\Omega \varphi^2\mathscr{U}^2w\,dx\Big\} \quad \forall 0<\varepsilon<1.$$

Choosing now $\varepsilon = \frac{1}{2Cq^2}$ it follows that

$$\int_\Omega \varphi^2|X\mathscr{U}|^2w\,dx \le \\ \le C\left\{q^2\int_\Omega |X\varphi|^2\mathscr{U}^2w\,dx + q^2K\left(\frac{1}{2Cq^2}\right)\int_\Omega \varphi^2\mathscr{U}^2w\,dx\right\}. \tag{3.12}$$

From (3.4), with $p=2$, and (3.12),

$$\left(\int_{B_{2r}} |\varphi\mathscr{U}|^{2\tau}w\,dx\right)^{\frac{1}{\tau}} \le$$

$$\leq Cr^2 w(B_{2r})^{\frac{1}{\tau}-1} \left\{ q^2 \int_{B_{2r}} |X\varphi|^2 \mathscr{U}^2 w\,dx + q^2 K\left(\frac{1}{2Cq^2}\right) \int_{B_{2r}} |\varphi \mathscr{U}|^2 w\,dx \right\},$$

with $\tau > 1$.

Let r_1 and r_2 be such that $r \leq r_1 \leq r_2 \leq 2r$. Choose φ such that $\varphi(x) = 1$ in B_{r_1}, $0 \leq \varphi \leq 1$ in B_{r_2} and $|X\varphi| \leq \frac{2}{r_2 - r_1}$. Then

$$\left(\int_{B_{r_1}} \mathscr{U}^{2\tau} w\,dx\right)^{\frac{1}{\tau}} \leq Cr^2 w(B_{2r})^{\frac{1}{\tau}-1} \frac{1}{(r_2-r_1)^2} q^2 K\left(\frac{1}{2Cq^2}\right) \int_{B_{r_2}} \mathscr{U}^2 w\,dx.$$

Taking the $\frac{1}{2q}$-th root and letting $l \to +\infty$, it follows readily that

$$\left(\int_{B_{r_1}} v^{2q\tau} w\,dx\right)^{\frac{1}{2q\tau}} \leq C^{\frac{1}{2q}} r^{\frac{1}{q}} w(B_{2r})^{\frac{1}{2}\left(\frac{1}{\tau}-1\right)\frac{1}{q}} \left(\frac{1}{r_2-r_1}\right)^{\frac{1}{q}} q^{\frac{1}{q}} \cdot$$

$$\cdot \left(K\left(\frac{1}{2Cq^2}\right)\right)^{\frac{1}{2q}} \left(\int_{B_{r_2}} v^{2q} w\,dx\right)^{\frac{1}{2q}}.$$

For $\gamma = 2q$, one has

$$\|v\|_{L^{\tau\gamma}(B_{r_1},w)} \leq C^{\frac{1}{\gamma}} r^{\frac{2}{\gamma}} w(B_{2r})^{\left(\frac{1}{\tau}-1\right)\frac{1}{\gamma}} \left(\frac{1}{r_2-r_1}\right)^{\frac{2}{\gamma}} \cdot$$

$$\cdot \left[\frac{1}{\left(\eta^{-1}\left(\frac{V}{w}; \frac{1}{2C(\frac{\gamma}{2})^2}\right)\right)^{Q+2}}\right]^{\frac{1}{\gamma}} \|v\|_{L^{\gamma}(B_{r_2},w)}.$$

Setting $\gamma_i = 2\tau^i$, for $i = 1, 2, \ldots$, and $r_i = r + \frac{r}{2^i}$, the previous inequality becomes

$$\|v\|_{L^{\gamma_{i+1}}(B_{r_{i+1}},w)} \leq C^{\frac{1}{\gamma_i}} r^{\frac{2}{\gamma_i}} w(B_{2r})^{\left(\frac{1}{\tau}-1\right)\frac{1}{\gamma_i}} \left(\frac{2^{i+1}}{r}\right)^{\frac{2}{\gamma_i}} \cdot$$

$$\cdot \left[\frac{1}{\left(\eta^{-1}\left(\frac{V}{w}; \frac{1}{2C(\frac{\gamma_i}{2})^2}\right)\right)^{Q+2}}\right]^{\frac{1}{\gamma_i}} \|v\|_{L^{\gamma_i}(B_{r_i},w)}.$$

Iteration yields

$$\|v\|_{L^{\infty}(B_r)} \leq C \frac{1}{w(B_{2r})^{\frac{1}{2}}} \prod_{i=0}^{+\infty} \left[\frac{1}{\left(\eta^{-1}\left(\frac{V}{w}; \frac{1}{2C\tau^{2i}}\right)\right)^{Q+2}}\right]^{\frac{1}{\gamma_i}} \|v\|_{L^p(B_{2r},w)}.$$

We underline the fact that

$$\prod_{i=0}^{+\infty}\left[\frac{1}{\left(\eta^{-1}\left(\frac{V}{w};\frac{1}{2C\tau^{2i}}\right)\right)^{Q+2}}\right]^{\frac{1}{\gamma_i}} < +\infty$$

if and only if the series

$$\sum_{i=0}^{+\infty}\frac{1}{2\tau^i}\log\eta^{-1}\left(\frac{V}{w};\frac{1}{2C\tau^{2i}}\right)$$

is convergent. Thus, the desired conclusion follows from Lemma 3.1.

□

By using the same technique, a Harnack-type inequality can be proved.

Theorem 3.4. *Let u be a weak nonnegative solution of equation* (3.8) *in a ball $B_{3r} \subset\subset \Omega$. Assume* (3.2) *and* (3.9). *Then there exists C independent of u, such that for any B_r for which $B_{4r} \subset \Omega$, it holds that*

$$\sup_{B_r} u \le C\left\{\inf_{B_r} u + \eta\left(\frac{f}{w};3r\right) + \left(\sum_{i=1}^{m}\eta\left(\left(\frac{h_i}{w}\right)^2;3r\right)\right)^{1/2}\right\}. \tag{3.13}$$

Proof. The proof follows along the lines of that of the previous Theorem. Setting $v = u + h$, where

$$h = \eta\left(\frac{f}{w};3r\right) + \left(\sum_{i=1}^{m}\eta\left(\left(\frac{h_i}{w}\right)^2;3r\right)\right)^{1/2},$$

and taking $\psi(x) = \varphi^2(x)v^\beta(x)$, where φ is a nonnegative smooth function such that $\operatorname{supp}\varphi(x) \subseteq B_{3r}$ and $\beta \in \mathbb{R}$, (3.10) yields

$$\begin{aligned}
\lambda^{-1}|\beta|\int_{B_{3r}}|Xv|^2\varphi^2 v^{\beta-1}w\,dx &\le 2\lambda\int_{B_{3r}}|Xv|\,|X\varphi|\,\varphi\,v^\beta w\,dx+\\
&+|\beta|\sum_{j=1}^{m}\int_{B_{3r}}|d_j|\,|X_jv|\,\varphi^2v^\beta\,dx + 2\sum_{j=1}^{m}\int_{B_{3r}}|d_j|\,\varphi\,|X_j\varphi|\,v^{\beta+1}\,dx+\\
&+\sum_{i=1}^{m}\int_{B_{3r}}|b_i|\,\eta^2\,|X_iv|\,v^\beta\,dx + \int_{B_{3r}}|c|\,\eta^2v^{\beta+1}\,dx + \int_{B_{3r}}h^{-1}|f|\,\varphi^2v^{\beta+1}\,dx+\\
&+|\beta|\sum_{i=1}^{m}\int_{B_{3r}}h^{-1}|h_i|\,\varphi^2\,|X_iv|\,v^\beta\,dx + 2\sum_{i=1}^{m}\int_{B_{3r}}h^{-1}|h_i|\,\varphi\,|X_i\varphi|\,v^\beta\,dx.
\end{aligned}$$

It follows that

$$\int_{B_{3r}} |Xv|^2\varphi^2 v^{\beta-1} w\,dx \leq$$
$$\leq C(\lambda)\left\{\frac{|\beta|+1}{\beta^2}\int_{B_{3r}} |X\varphi|^2 v^{\beta+1} w\,dx + \left(\frac{|\beta|+1}{\beta}\right)^2 \int_{B_{3r}} V\,\varphi^2 v^{\beta+1}\,dx\right\}, \tag{3.14}$$

where

$$V = \sum_{i=1}^m \frac{|b_i|^2}{w} + |c| + \sum_{i=1}^m \frac{|d_i|^2}{w} + h^{-1}|f| + h^{-2}\sum_{i=1}^m \frac{|h_i|^2}{w}.$$

Setting

$$\mathscr{U}(x) = \begin{cases} v(x)^{\frac{\beta+1}{2}} & \text{if} \quad \beta \neq -1 \\ \log v(x) & \text{if} \quad \beta = -1 \end{cases},$$

by virtue of (5.24) one concludes that

$$\int_{B_{3r}} |X\mathscr{U}|^2\varphi^2 w\,dx \leq C(\lambda)\left\{\frac{(|\beta|+1)^3}{\beta^2}\int_{B_{3r}} |X\varphi|^2\mathscr{U}^2 w\,dx + \right.$$
$$\left. + \left(\frac{|\beta|+1}{\beta}\right)^2 \int_{B_{3r}} V\,\varphi^2\mathscr{U}^2\,dx\right\} \quad \text{if} \quad \beta \neq -1, \tag{3.15}$$

and that

$$\int_{B_{3r}} |X\mathscr{U}|^2\varphi^2 w\,dx \leq C(\lambda)\left\{\int_{B_{3r}} |X\varphi|^2 w\,dx + \right.$$
$$\left. + \int_{B_{3r}} V\,\varphi^2\,dx\right\} \quad \text{if} \quad \beta = -1. \tag{3.16}$$

Consider first (5.25). From Corollary 3.2

$$\int_{B_{3r}} |X\mathscr{U}|^2\varphi^2 w\,dx \leq C\left\{\int_{B_{3r}} |X\varphi|^2 w\,dx + \int_{B_{3r}} \varphi^2 w\,dx\right\}. \tag{3.17}$$

Chooseφ such that $\varphi(x) \equiv 1$ in B_ρ, $\operatorname{supp}\eta \subset B_{2\rho} \subset B_{3r}$, and $|X\varphi| \leq \frac{C}{\rho}$, where B_ρ is an arbitrary open ball contained in B_{2r}. By (3.17) and on account of the doubling property of w, it follows that

$$\left(\frac{1}{w(B_\rho)}\int_{B_\rho} |X\mathscr{U}|^2 w\,dx\right)^{\frac{1}{2}} \leq C\frac{1}{\rho}.$$

Thus, from (3.5)

$$\left(\frac{1}{w(B_\rho)}\int_{B_\rho}|\mathscr{U}-\mathscr{U}_{B_\rho}|^2 w\,dx\right)^{\frac{1}{2}}\le C$$

for every $B_\rho \subseteq B_{2r}$, with C depending on λ, Q, $\eta\left(\frac{V}{w};3r\right)$ and the constant in (3.5). By the John-Nirenberg Lemma for *BMO* spaces (see [1]) there exist two positive constants, p_0 and $\overline{C}$, depending on the same arguments as C, such that

$$\left(\frac{1}{w(B_{2r})}\int_{B_{2r}}e^{p_0\mathscr{U}}w\,dx\right)^{\frac{1}{p_0}}\left(\frac{1}{w(B_{2r})}\int_{B_{2r}}e^{-p_0\mathscr{U}}w\,dx\right)^{\frac{1}{p_0}}\le\overline{C}. \tag{3.18}$$

For any real number $p\neq 0$ and $h>0$ now set

$$\Phi(p,h)=\left(\int_{B_h}v^p w\,dx\right)^{\frac{1}{p}}.$$

By virtue of (3.18) and since $\mathscr{U}=\log v$, one deduces

$$w(B_{2r})^{-\frac{1}{p_0}}\Phi(p_0,2r)\le\overline{C}w(B_{2r})^{\frac{1}{p_0}}\Phi(-p_0,2r). \tag{3.19}$$

Now consider (3.15). Corollary 3.2 implies

$$\int_{B_{3r}}|X\mathscr{U}|^2\varphi^2 w\,dx\le C\Bigg\{(|\beta|+1)^3\left(1+\frac{1}{|\beta|}\right)^2\int_{B_{3r}}|X\varphi|^2\mathscr{U}^2 w\,dx+ \\ +\left[\frac{1}{\eta^{-1}\left(\frac{V}{w};|q|^{-2}(1+\frac{1}{|\beta|})^{-2}\right)}\right]^{Q+2}\int_{B_{3r}}\varphi^2\mathscr{U}^2 w\,dx\Bigg\}.$$

By (3.4),

$$\left(\int_{B_{3r}}|\varphi\mathscr{U}|^{2\tau}w\,dx\right)^{\frac{1}{\tau}}\le \\ \le Cr^2 w(B_{3r})^{\frac{1}{\tau}-1}\Bigg\{(|\beta|+1)^3\left(1+\frac{1}{|\beta|}\right)^2\int_{B_{3r}}|X\varphi|^2\mathscr{U}^2 w\,dx+ \\ +\left[\frac{1}{\eta^{-1}\left(\frac{V}{w};|q|^{-2}(1+\frac{1}{|\beta|})^{-2}\right)}\right]^{Q+2}\int_{B_{3r}}\eta^2\mathscr{U}^2 w\,dx\Bigg\}, \tag{3.20}$$

where C is a positive constant independent of $\mathscr{U}$.

Let r_1 and r_2 be real numbers such that $r < r_1 < r_2 \leq 2r$. Let the function φ be chosen so that $\varphi \equiv 1$ in B_{r_1}, $0 \leq \varphi(x) \leq 1$ in B_{r_2}, $\varphi(x) = 0$ outside B_{r_2} and $|X\varphi(x)| \leq \frac{C}{r_2 - r_1}$. By (5.27) it follows that

$$\left(\int_{B_{r_1}} |\mathscr{U}|^{2\tau} w\,dx \right)^{\frac{1}{\tau}} \leq$$

$$\leq C r^2 w(B_{3r})^{\frac{1}{\tau}-1} \frac{1}{(r_2-r_1)^2} (|\beta|+1)^3 \left(1 + \frac{1}{|\beta|}\right)^2 \cdot$$

$$\cdot \left[\frac{1}{\eta^{-1}\left(\frac{V}{w}; |q|^{-2}(1+\frac{1}{|\beta|})^{-2}\right)} \right]^{Q+2} \int_{B_{r_2}} \mathscr{U}^2 w\,dx. \quad (3.21)$$

Selecting $p = \beta + 1$, taking the p-th root in (5.28) and recalling $\mathscr{U}^2(x) = v^{\beta+1}(x) = v^p(x)$, it is easily derived that

$$\Phi(\tau p, r_1) \leq C^{\frac{1}{p}} r^{\frac{1}{q}} w(B_{3r})^{\frac{1}{p}\left(\frac{1}{\tau}-1\right)} (|\beta|+1)^{\frac{3}{p}} \left(1 + \frac{1}{|\beta|}\right)^{\frac{1}{q}} \cdot$$

$$\cdot \left[\frac{1}{\eta^{-1}\left(\frac{V}{w}; |q|^{-2}(1+\frac{1}{|\beta|})^{-2}\right)} \right]^{\frac{Q+2}{p}} \frac{1}{(r_2-r_1)^{\frac{2}{p}}} \Phi(p, r_2), \quad (3.22)$$

for positive $p \neq 1$, and that

$$\Phi(\tau p, r_1) \geq C^{\frac{1}{p}} r^{\frac{1}{q}} w(B_{3r})^{\frac{1}{p}\left(\frac{1}{\tau}-1\right)} (|\beta|+1)^{\frac{3}{p}} \cdot$$

$$\cdot \left[\frac{1}{\eta^{-1}\left(\frac{V}{w}; |q|^{-2}(1+\frac{1}{|\beta|})^{-2}\right)} \right]^{\frac{Q+2}{p}} \frac{1}{(r_2-r_1)^{\frac{2}{p}}} \Phi(p, r_2), \quad (3.23)$$

for negative p. These inequalities are next iterated in the spirit of [37] and [9] to get

$$\sup_{B_r} v \leq C \inf_{B_r} v$$

where C depends on λ, Q, $\eta\left(\frac{V}{w}; r\right)$, C_1 and τ. Since $v = u + h$, (3.13) follows easily. □

3.3.2 Degenerate equations under natural growth

Let Ω be a bounded domain in $\mathbb{R}^n$ and let $X = (X_1, X_2, \ldots, X_m)$ be a system of locally Lipschitz vector fields in $\mathbb{R}^n$ whose formal adjoints are denoted by X_i^*. Assume $(A3)$ with $p=2$ and let w be a 2-admissible weight. Let $\{a_{ij}(x)\}$ be a symmetric matrix of measurable functions in Ω satisfying the ellipticity condition (3.2). We consider the elliptic quasi linear equation in divergence form

$$-X_j^*(a_{ij}X_iu+d_ju)+\frac{b_0}{\lambda}w|Xu|^2+b_iX_iu+cu=f-X_i^*h_i\,, \tag{3.24}$$

where

$$b_0\in\mathbb{R}\setminus\{0\}\,,\left(\frac{b_i}{w}\right)^2,\frac{c}{w},\left(\frac{d_i}{w}\right)^2,\frac{f}{w},\left(\frac{h_i}{w}\right)^2\in S'(\Omega,w)\,. \tag{3.25}$$

Definition 3.6. A function $u\in W^{1,2}_{loc}(\Omega,w)$ is said to be a local, weak supersolution (subsolution), of (3.24) if for any $\phi\in W_0^{1,2}(\Omega,w)$, $\phi\geq 0$

$$\int_\Omega\left[(a_{ij}X_iu+d_ju)X_j\phi+\left(\frac{b_0}{\lambda}w|Xu|^2+b_iX_iu+cu\right)\phi\right]dx\geq(\leq)$$

$$\int_\Omega(f\phi+h_iX_i\phi)dx\,.$$

$u\in W^{1,2}_{loc}(\Omega,w)$ is a local, weak solution of (3.24) if it is both a local supersolution and a local subsolution.

Our first result is the weak Harnack inequality for local, bounded supersolutions of (3.24). The proof follows along the lines of [33].

Theorem 3.5. *Assume conditions* (3.2) *and* (3.25) *to be satisfied and let u be a weak nonnegative supersolution of equation* (3.24) *in a ball $B_{3r}\subset\subset\Omega$. Let $M>0$ be a constant such that $u\leq M$ in B_{3r}. Then, there exists C depending on Q, M, λ and the weight w, such that*

$$w^{-1}(B_{2r})\int_{B_{2r}}uw\,dx\leq$$

$$\leq C\left\{\inf_{B_r}u+\eta\left(\frac{f}{w};3r\right)+\left(\sum_{i=1}^m\eta\left(\left(\frac{h_i}{w}\right)^2;3r\right)\right)^{1/2}\right\}.$$

Proof. Let $k=\eta\left(\frac{f}{w};3r\right)+\left(\sum_{i=1}^m\eta\left(\left(\frac{h_i}{w}\right)^2;3r\right)\right)^{1/2}$ and $v=u+k$. For $\varphi\in C_0^1(B_{3r})$, $\varphi\geq 0$, set $\phi(x)=\varphi^2(x)v^\beta(x)e^{-|b_0|v(x)}$, $\beta<0$, as a test function in (3.24). Since u is a supersolution in B_{3r} of (3.24), one has

$$\int_{B_{3r}} \Big[2\varphi(a_{ij}X_iu+d_ju-h_j)X_j\varphi v^{\beta}e^{-|b_0|v}+ \\ (-|\beta|v^{\beta-1}-|b_0|v^{\beta})\varphi^2e^{-|b_0|v}(a_{ij}X_iu+d_ju-h_j)X_jv+ \\ \frac{b_0}{\lambda}w|Xu|^2\varphi^2v^{\beta}e^{-|b_0|v}+(b_iX_iu+cu-f)\varphi^2v^{\beta}e^{-|b_0|v}\Big]dx \geq 0$$

and

$$\int_{B_{3r}} \varphi^2e^{-|b_0|v}(b_0v^{\beta}+|\beta|v^{\beta-1})|Xv|^2wdx \leq \\ \int_{B_{3r}} \varphi^2e^{-|b_0|v}(|b_0|v^{\beta}+|\beta|v^{\beta-1})|Xv|^2wdx \leq \\ \lambda\int_{B_{3r}} \varphi^2e^{-|b_0|v}(|b_0|v^{\beta}+|\beta|v^{\beta-1})a_{ij}X_ivX_jvdx \leq \\ \lambda\int_{B_{3r}} \varphi^2e^{-|b_0|v}(|\beta|v^{\beta-1}+|b_0|v^{\beta})(h_j-d_ju)X_jvdx+ \\ 2\lambda\int_{B_{3r}} \varphi(a_{ij}X_iv+d_ju-h_j)X_j\varphi v^{\beta}e^{-|b_0|v}dx+ \\ \int_{B_{3r}} b_0w|Xv|^2\varphi^2v^{\beta}e^{-|b_0|v}dx+ \\ \lambda\int_{B_{3r}} (b_iX_iv+cu-f)\varphi^2v^{\beta}e^{-|b_0|v}dx. \quad (3.26)$$

From (3.26), it follows

$$\int_{B_{3r}} \varphi^2e^{-|b_0|v}|\beta|v^{\beta-1}|Xv|^2wdx \leq \\ \lambda\int_{B_{3r}} \varphi^2e^{-|b_0|v}(|\beta|v^{\beta-1}+|b_0|v^{\beta})(h_j-d_ju)X_jvdx+ \\ 2\lambda\int_{B_{3r}} \varphi(a_{ij}X_iv+d_ju-h_j)X_j\varphi v^{\beta}e^{-|b_0|v}dx+ \\ \lambda\int_{B_{3r}} (b_iX_iv+cu-f)\varphi^2v^{\beta}e^{-|b_0|v}dx.$$

Since v is bounded the exponential can be dropped to obtain

$$\int_{B_{3r}} \varphi^2|\beta|v^{\beta-1}|Xv|^2 w dx \leq$$

$$C(M,b_0)\Big[2\lambda \int_{B_{3r}} \varphi a_{ij}X_i v X_j \varphi v^\beta dx + \lambda|\beta| \int_{B_{3r}} |d_j||X_j v|v^\beta \varphi^2 dx +$$

$$2\lambda \int_{B_{3r}} |d_j|v^{\beta+1}\varphi X_j \varphi dx + 2\lambda \int_{B_{3r}} |h_j|v^\beta \varphi X_j \varphi dx + \lambda \int_{B_{3r}} |b_i||X_i v|\varphi^2 v^\beta +$$

$$\lambda \int_{B_{3r}} |c|\varphi^2 v^{\beta+1} dx + \lambda \int_{B_{3r}} |f|\varphi^2 v^\beta dx +$$

$$\lambda|\beta| \int_{B_{3r}} h_j X_j v v^{\beta-1}\varphi^2 dx + \lambda \int_{B_3 r} |d_j||X_i v|\varphi^2 v^\beta dx\Big].$$

Now, set

$$V = \sum_{i=1}^n \frac{|b_i|^2}{w} + |c| + \sum_{j=1}^n \frac{|d_j|^2}{w} + k^{-1}|f| + k^{-2}\sum_{i=1}^n \frac{|h_i|^2}{w}.$$

A straightforward application of Young's inequality yields

$$\int_{B_{3r}} \varphi^2 v^{\beta-1}|Xv|^2 w dx \leq$$

$$C(M,b_0,\lambda)\left[\frac{|\beta|+1}{\beta^2}\int_{B_{3r}} v^{\beta+1}|X\eta|^2 w dx + \left(\frac{|\beta|+1}{\beta}\right)^2 \int_{B_{3r}} V\varphi^2 v^{\beta+1} dx\right] \leq$$

$$C(M,b_0,\lambda)\left(\frac{|\beta|+1}{\beta}\right)^2\left[\int_{B_{3r}} v^{\beta+1}|X\varphi|^2 w dx + \int_{B_{3r}} V\varphi^2 v^{\beta+1} dx\right]. \quad (3.27)$$

Next, set

$$\mathscr{U}(x) = \begin{cases} v^{\frac{\beta+1}{2}}(x) & \text{if} \quad \beta \neq -1 \\ \log v(x) & \text{if} \quad \beta = -1. \end{cases}$$

By (5.26),

$$\int_{B_{3r}} \varphi^2|X\mathscr{U}|^2 w dx \leq C(\beta+1)^2\left(\frac{|\beta|+1}{\beta}\right)^2\left\{\int_{B_{3r}} |X\varphi|^2\mathscr{U}^2 w dx + \int_{B_{3r}} V\varphi^2\mathscr{U}^2 dx\right\}, \beta \neq -1 \quad (3.28)$$

while

$$\int_{B_{3r}} \varphi^2|X\mathscr{U}|^2 w dx \leq C\left\{\int_{B_{3r}} |X\varphi|^2 w dx + \int_{B_{3r}} V\varphi^2 dx\right\} \quad (3.29)$$

if $\beta = -1$.

Consider first the case $\beta = -1$. Corollary 3.2 yields

$$\int_{B_{3r}} \varphi^2 |X\mathscr{U}|^2 w\,dx \le C\left(\int_{B_{3r}} |X\varphi|^2 w\,dx + \int_{B_{3r}} \varphi^2 w dx\right).$$

Let B_h be a ball contained in B_{2r}. Choosing φ so that $\varphi(x) \equiv 1$ in B_h, $0 \le \varphi \le 1$ in $B_{3r} \setminus B_h$ and such that $|X\varphi| \le \frac{C}{h}$, one obtains

$$\|X\mathscr{U}\|_{L^2(B_h,w)} \le C\frac{w(B_h)^{\frac{1}{2}}}{h}.$$

From (3.5) and on account of the John-Nirenberg lemma (see [1]), it holds that $\mathscr{U}(x) = \log v(x) \in BMO$. Then, there exist two positive constants p_0 and C, such that

$$\left(\fint_{B_{2r}} e^{p_0\mathscr{U}} w\,dx\right)^{\frac{1}{p_0}} \left(\fint_{B_{2r}} e^{-p_0\mathscr{U}} w\,dx\right)^{\frac{1}{p_0}} \le C. \tag{3.30}$$

Consider the following family of seminorms:

$$\Phi(p,h) = \left(\int_{B_h} |v|^p w\,dx\right)^{1/p}, \quad p \ne 0.$$

(3.30) yields

$$\frac{1}{w(B_{2r})^{1/p_0}}\Phi(p_0,2r) \le Cw(B_{2r})^{1/p_0}\Phi(-p_0,2r).$$

Next, consider $\beta \ne -1$ (see inequality (3.28)). Corollary 3.2 implies

$$\int_{B_{3r}} |X\mathscr{U}|^2\varphi^2 w\,dx \le C\left\{\left[\left(\frac{\beta+1}{2}\right)^2 + 1\right]\left(1+\frac{1}{|\beta|}\right)^2 \int_{B_{3r}} |X\varphi|^2\mathscr{U}^2 w\,dx + \right.$$
$$\left. + \left[\frac{1}{\eta^{-1}\left(\frac{V}{w}; \left(\frac{\beta+1}{2}\right)^{-2}\left(1+\frac{1}{|\beta|}\right)^{-2}\right)}\right]^{Q+2} \int_{B_{3r}} \varphi^2\mathscr{U}^2 w\,dx\right\}. \tag{3.31}$$

From (3.4) one has

$$\left(\int_{B_{3r}} |\varphi\mathscr{U}|^{2\tau} w\,dx\right)^{\frac{1}{\tau}} \leq cw(B_{3r})^{\frac{1}{\tau}-1}\left\{\left[\left(\frac{\beta+1}{2}\right)^2+2\right]\left(1+\frac{1}{|\beta|}\right)^2\cdot\right.$$
$$\int_{B_{3r}} |X\varphi|^2\mathscr{U}^2 w\,dx+$$
$$\left.+\left[\frac{1}{\eta^{-1}\left(\frac{V}{w};\left(\frac{\beta+1}{2}\right)^{-2}\left(1+\frac{1}{|\beta|}\right)^{-2}\right)}\right]^{Q+2}\int_{B_{3r}}\varphi^2\mathscr{U}^2 w\,dx\right\} \quad (3.32)$$

where c is a positive constant independent of w.

Choose the function φ. For r_1 and r_2 such that $r \leq r_1 < r_2 \leq 2r$, select φ such that $\varphi(x) \equiv 1$ in B_{r_1}, $0 \leq \varphi(x) \leq 1$ in B_{r_2}, $\varphi(x) = 0$ outside B_{r_2} znd that $|X\varphi| \leq \frac{c}{r_2-r_1}$ for some fixed constant c.

Then,

$$\left(\int_{B_{r_1}} \mathscr{U}^{2\tau} w\,dx\right)^{\frac{1}{\tau}} \leq cw(B_{3r})^{\frac{1}{\tau}-1}\frac{1}{(r_2-r_1)^2}\left[\left(\frac{\beta+1}{2}\right)^2+2\right]\cdot$$
$$\cdot\left(1+\frac{1}{|\beta|}\right)^2\left[\frac{1}{\eta^{-1}\left(\frac{V}{w};\left(\frac{\beta+1}{2}\right)^{-2}\left(1+\frac{1}{|\beta|}\right)^{-2}\right)}\right]^{Q+2}\int_{B_{r_2}}\mathscr{U}^2 w\,dx.$$

Setting $\gamma = \beta+1$ and recalling that $\mathscr{U}(x) = v^{\frac{\beta+1}{2}}(x)$, it can be seen that

$$\Phi(\tau\gamma, r_1) \geq c^{\frac{1}{\gamma}} w(B_{3r})^{\frac{1}{\gamma}(\frac{1}{\tau}-1)}\left[\left(\frac{\beta+1}{2}\right)^2+2\right]^{\frac{1}{\gamma}}\cdot$$
$$\cdot\left[\frac{1}{\eta^{-1}\left(\frac{V}{w};\left(\frac{\beta+1}{2}\right)^{-2}\right)}\right]^{\frac{Q+2}{\gamma}}\frac{1}{(r_2-r_1)^{\frac{2}{\gamma}}}\Phi(\gamma, r_2), \quad (3.33)$$

for negative γ.

Iterate the inequality just obtained. Setting $\gamma_i = \tau^i p_0$ and $r_i = r + \frac{r}{2^i}$, $i = 1, 2, \ldots$, iteration of (3.33) and Lemma 3.1 yield

$$\Phi(-\infty, r) \geq c(\phi_{\frac{V}{w}}, \operatorname{diam}\Omega)w(B_{3r})^{\frac{1}{p_0}}\Phi(-p_0, 2r).$$

Therefore, by Hölder's inequality,

$$\Phi(p_0', 2r) \leq \Phi(p_0, 2r) w(B_{3r})^{\frac{1}{p_0'} - \frac{1}{p_0}}, \quad p_0' \leq p_0,$$

from which it follows that

$$w^{-1}(B_{2r})\Phi(1, 2r) \leq c\Phi(-\infty, r)$$

and hence, the desired claim. □

The following weak Harnack inequality for subsolutions can be obtained in a similar way.

Theorem 3.6. *Let u be a weak nonnegative subsolution of* (3.24) *in $B_{3r} \subset\subset \Omega$. Assume* (3.2) *and* (3.25). *Let $M > 0$ be a constant such that $u \leq M$ in B_{3r}. Then there exists C depending on Q, M, λ and on the weight w, such that*

$$\sup_{B_r} u \leq$$
$$C\left\{ w^{-1}(B_{2r}) \int_{B_{2r}} uw\,dx + \eta\left(\frac{f}{w}; 3r\right) + \left(\sum_{i=1}^m \eta\left(\left(\frac{h_i}{w}\right)^2; 3r\right)\right)^{1/2} \right\}.$$

Now, from our previous results, a Harnack-type inequality for solutions can be derived.

Theorem 3.7. *Assume conditions* (3.2) *and* (3.25) *are satisfied and let u be a weak nonnegative solution of equation* (3.24) *in a ball $B_{3r} \subset\subset \Omega$. Let $M > 0$ be a constant such that $u \leq M$ in B_{3r}. Then, there exists C depending on Q, M, λ and on the weight w, such that*

$$\sup_{B_r} u \leq C\left\{ \inf_{B_r} u + \eta\left(\frac{f}{w}; 3r\right) + \left(\sum_{i=1}^m \eta\left(\left(\frac{h_i}{w}\right)^2; 3r\right)\right)^{1/2} \right\}.$$

3.4 Continuity of solutions for linear degenerate equations

Harnack's inequality implies that weak solutions of (3.8) and (3.24) are continuous (see [17], [19], [18], [8], [10]). Next, the regularity of the solutions of equation (3.8) will be proved.

Theorem 3.8. *Let $u \in W^{1,2}(\Omega, w)$ be a weak solution of* (3.8) *such that the degenerate ellipticity condition* (3.2) *holds true and such that*

$$\left(\frac{b_i}{w}\right)^2, \frac{c}{w}, \left(\frac{d_i}{w}\right)^2, \frac{f}{w}, \left(\frac{h_i}{w}\right)^2 \in S'(\Omega, w)$$

then u is continuous in Ω.

Proof. Let B_r be an arbitrary ball contained in Ω and consider the functions, with $0 < \rho \leq r$

$$M(\rho) = \sup_{B_\rho} u, \quad m(\rho) = \inf_{B_\rho} u \quad \varphi(\rho) = M(\rho) - m(\rho).$$

setting $\overline{u} = M - u = M(r) - u$, $\overline{u}$ is a nonnegative weak solution of the equation

$$-X_j^*(a_{ij}X_iu - d_ju) + b_iX_iu + cu = (Mc - f) - X_i^*(Md_i - h_i)$$

in B_r, with

$$\frac{Mc - f}{w}, \left(\frac{Md_i - h_i}{w}\right)^2 \in S'(\Omega, w).$$

By virtue of Harnack's inequality (3.13) one has

$$\sup_{B_{\frac{r}{3}}} \overline{u} \leq C \left(\inf_{B_{\frac{r}{3}}} \overline{u} + \overline{h}(r) \right)$$

and

$$\overline{h}(r) = \eta\left(\frac{\overline{f}}{w}; r\right) + \left(\sum_{i=1}^{m} \eta\left(\left(\frac{\overline{h}_i}{w}\right)^2; r\right)\right)^{1/2},$$

where $\overline{f} = Mc - f$ and $\overline{h}_i = Md_i - h_i$.

Observe that $\overline{h}(r)$ is a positive non-decreasing function with $\lim_{r \to 0} \overline{h}(r) = 0$.

Then

$$M(r) - m\left(\frac{r}{3}\right) \leq C\left\{M(r) - M\left(\frac{r}{3}\right) + \overline{h}(r)\right\}. \tag{3.34}$$

In the same way, setting $\overline{\overline{u}} = u - m = u - m(r)$, it follows that

$$M\left(\frac{r}{3}\right) - m(r) \leq C\left\{m\left(\frac{r}{3}\right) - m(r) + \overline{h}(r)\right\}. \tag{3.35}$$

Adding (3.34) and (3.35),

$$M\left(\frac{r}{3}\right) - m\left(\frac{r}{3}\right) \leq \frac{C-1}{C+1}[M(r) - m(r)] + \frac{2C}{C+1}\overline{h}(r).$$

Set $\theta = \frac{C-1}{C+1} < 1$, $h(r) = \frac{2C}{C+1}\overline{h}(r)$.
Thus,

$$\varphi\left(\frac{r}{3}\right) \leq \theta\varphi(r) + h(r)$$

and the conclusion follows from Lemma 3.2. □

The next result is a natural consequence of the previous one if one assumes the lower order terms to belong to the Morrey classes M_σ. Specifically,

Corollary 3.3. *Let $u \in W^{1,2}(\Omega, w)$ be a weak solution of* (3.8) *or a local, bounded solution of* (3.24)*, such that the degenerate ellipticity condition holds true and such that*

$$\left(\frac{b_i}{w}\right)^2, \frac{c}{w}, \left(\frac{d_i}{w}\right)^2, \frac{f}{w}, \left(\frac{h_i}{w}\right)^2 \in M_\sigma(\Omega, w).$$

Then u is locally Hölder continuous in Ω.

Using similar techniques, analogous results for local, bounded solution of equation (3.24) can be derived.

Theorem 3.9. *Let $u \in W^{1,2}(\Omega, w)$ be a local, bounded weak solution of* (3.24)*, such that the degenerate ellipticity condition* (3.2) *holds true and that*

$$\left(\frac{b_i}{w}\right)^2, \frac{c}{w}, \left(\frac{d_i}{w}\right)^2, \frac{f}{w}, \left(\frac{h_i}{w}\right)^2 \in S'(\Omega, w).$$

Then u is continuous in Ω.

Moreover, assuming

$$\left(\frac{b_i}{w}\right)^2, \frac{c}{w}, \left(\frac{d_i}{w}\right)^2, \frac{f}{w}, \left(\frac{h_i}{w}\right)^2 \in M_\sigma(\Omega, w),$$

it follows that u is locally Hölder continuous in Ω.

3.5 Smoothness for non linear degenerate elliptic equations

Let Ω be a bounded domain in $\mathbb{R}^n$. Let $X = (X_1, X_2, \ldots, X_m)$ be a system of locally Lipschitz vector fields in $\mathbb{R}^n$. For $i = 1, 2, \ldots, m$, let X_i^* stand for the formal adjoint of the vector fields X_i. Let $1 < p < Q$ and w be a p-admissible weight.

We consider the following quasilinear elliptic equation

$$X_i^* A_i(x,u,Xu) + B(x,u,Xu) = 0, \tag{3.36}$$

where $A : \Omega \times \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$ and $B : \Omega \times \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$ are measurable functions in their domains. Two kinds of structural assumptions on A and B will be considered, referred to as controlled-growth or natural-growth assumptions, respectively. For the controlled-growth assumptions, there exist $a > 0$ and functions b, c, d, e, f and g, such that

$$\begin{cases} |A(x,u,\xi)| \le aw(x)|\xi|^{p-1} + b(x)|u|^{p-1} + e(x) \\ |B(x,u,\xi)| \le c(x)|\xi|^{p-1} + d(x)|u|^{p-1} + f(x) \\ \xi \cdot A(x,u,\xi) \ge w(x)|\xi|^p - d(x)|u|^p - g(x). \end{cases} \tag{3.37}$$

For the natural growth assumptions it is assumed that there exist $a, b_0 > 0$ and functions b, c, d, e, f and g such that

$$\begin{cases} |A(x,u,\xi)| \le aw(x)|\xi|^{p-1} + b(x)|u|^{p-1} + e(x) \\ |B(x,u,\xi)| \le b_0 w(x)|\xi|^p + c(x)|\xi|^{p-1} + d(x)|u|^{p-1} + f(x) \\ \xi \cdot A(x,u,\xi) \ge w(x)|\xi|^p - d(x)|u|^p - g(x). \end{cases} \tag{3.38}$$

Next, we give the definition of weak solutions of the quasilinear equation (3.36).

Definition 3.7. A function $u \in W^{1,p}(\Omega, w)$ is a local weak solution of (3.36) in Ω if

$$\int_\Omega A(x,u(x),Xu(x))\, X\phi(x)\, dx = \int_\Omega B(x,u(x),Xu(x))\ \phi(x)\, dx, \tag{3.39}$$

for any $\phi \in W_0^{1,p}(\Omega, w)$.

Local boundedness will next be proved for solutions of equations (3.36) under controlled growth.

Theorem 3.10. *Let u be a weak solution of* (3.36). *Assume that the structural conditions* (3.37) *hold true with*

$$a \in \mathbb{R},\ \left(\frac{b}{w}\right)^{p/p-1},\left(\frac{c}{w}\right)^p,\ \left(\frac{d}{w}\right),\ \left(\frac{e}{w}\right)^{p/p-1},\ \frac{f}{w}, \frac{g}{w}\ \in S_p'(\Omega, w). \tag{3.40}$$

Then, there exists a positive constant C, independent of u, such that, for any metric ball $B_r = B(x_0, r)$ for which $B(x_0, 4r) \subset \Omega$, it holds that

$$\sup_{B_r} |u| \le C \left\{ \left(\int_{B_{2r}} |u|^p\, w\, dx \right)^{\frac{1}{p}} + h(3r) \right\}, \tag{3.41}$$

where

$$h(3r) = \left[\eta\left(\left(\frac{e}{w}\right)^{\frac{p}{p-1}};3r\right) + \eta\left(\frac{g}{w};3r\right)\right]^{\frac{1}{p}} + \left[\eta\left(\frac{f}{w};3r\right)\right]^{\frac{1}{p-1}}.$$

Proof. Let $B(x_0,4r)$ be a ball in Ω. We provide next a simplified form of the structural assumptions (3.37). Setting $h = h(3r)$ and $\tilde{u} = |u| + h$, one easily gets from (3.37)

$$\begin{cases} |A(x,u,\xi)| \leq & aw(x)|\xi|^{p-1} + b_1|\tilde{u}|^{p-1} \\ |B(x,u,\xi)| \leq & c|\xi|^{p-1} + d_1|\tilde{u}|^{p-1} \\ \xi \cdot A(x,u,\xi) \geq & w(x)|\xi|^p - d_1|\tilde{u}|^p, \end{cases} \tag{3.42}$$

where

$$b_1 = b + h^{1-p}e \quad , \quad d_1 = d + h^{1-p}f + h^{-p}g.$$

The functions $\left(\frac{b_1}{w}\right)^{\frac{p}{p-1}}$ and $\frac{d_1}{w}$ belong to the class $S'_p(B_{4r},w)$ and moreover, for any $0 < \rho < 2r$,

$$\begin{aligned} \eta\left(\left(\frac{b_1}{w}\right)^{\frac{p}{p-1}};\rho\right) &\leq C(p)\left[\eta\left(\left(\frac{b}{w}\right)^{\frac{p}{p-1}};\rho\right) + h^{-p}\eta\left(\left(\frac{e}{w}\right)^{\frac{p}{p-1}};\rho\right)\right] \\ &\leq C(p)\left[\eta\left(\left(\frac{b}{w}\right)^{\frac{p}{p-1}};\rho\right) + 1\right] \end{aligned}$$

$$\begin{aligned} \eta\left(\frac{d_1}{w};\rho\right) &\leq C(p)\left[\eta\left(\frac{d}{w};\rho\right) + h^{1-p}\eta\left(\frac{f}{w};\rho\right) + h^{-p}\eta\left(\frac{g}{w};\rho\right)\right] \\ &\leq C(p)\left[\eta\left(\frac{d}{w};\rho\right) + 2\right]. \end{aligned}$$

This means that, under assumptions (3.40), the reduced structural assumptions (3.42) are of the same kind as the general structural assumptions (3.37). Fix $q \geq 1$, $l > h$ and let

$$F(\tilde{u}) = \begin{cases} \tilde{u}^q & \text{if} \quad h \leq \tilde{u} \leq l \\ ql^{q-1}(\tilde{u} - l) + l^q & \text{if} \quad l \leq \tilde{u}. \end{cases}$$

Set

$$G(u) = \operatorname{sign} u\left(F(\tilde{u})[F'(\tilde{u})]^{p-1} - q^{p-1}h^{\beta}\right) \quad u \in]-\infty,+\infty[\,,$$

where β is such that $pq = p + \beta - 1$.

The test function to be used in (3.39) is $\phi(x) = \varphi^p(x)G(u)$, where $\varphi(x)$ is a smooth function such that $0 \leq \varphi \leq 1$, $\varphi \equiv 1$ in B_r, compactly supported in B_{2r}.

Following the classical pattern in [32] the test function in Definition 3.7 is substituted. So, by using the structural conditions (3.42) it follows that

$$\int_{B_{2r}} \varphi^p |X\mathscr{U}|^p w\,dx \leq ap \int_{B_{2r}} |(X\varphi)\mathscr{U}||\varphi(X\mathscr{U})|^{p-1} w\,dx +$$
$$+ q^{p-1} p \int_{B_{2r}} b_1 |(X\varphi)\mathscr{U}||\varphi\mathscr{U}|^{p-1}\,dx + \int_{B_{2r}} c|\varphi\mathscr{U}||\varphi(X\mathscr{U})|^{p-1}\,dx +$$
$$+ (1+p)q^{p-1} \int_{B_{2r}} d_1 |\varphi\mathscr{U}|^p\,dx,$$

where $\mathscr{U} = \mathscr{U}(x) = F(\tilde{u})$.

With the aid of the elementary inequality

$$ab^{p-1} \leq \frac{1}{p}\varepsilon^{1-p} a^p + \left(1 - \frac{1}{p}\right)\varepsilon b^p \qquad \forall \varepsilon > 0,$$

the previous one can be simplified to get

$$\int_{B_{2r}} \varphi^p |X\mathscr{U}|^p w\,dx \leq$$
$$\leq C(p,a)q^{p-1} \left\{ \int_{B_{2r}} |\mathscr{U}(X\varphi)|^p w\,dx + \int_{B_{2r}} V|\varphi\mathscr{U}|^p\,dx \right\},$$

where

$$V = \frac{b_1^{\frac{p}{p-1}}}{w^{\frac{1}{p-1}}} + \frac{c^p}{w^{p-1}} + d_1\,.$$

The desired clam follows by arguing as in the proof of Theorem 3.3.

□

Theorem 3.11. *Let u be a nonnegative weak solution of* (3.36). *Assume that the structural conditions* (3.37) *hold with*

$$a \in \mathbb{R},\ \left(\frac{b}{w}\right)^{p/p-1},\ \left(\frac{c}{w}\right)^p,\ \frac{d}{w},\ \left(\frac{e}{w}\right)^{p/p-1},\ \frac{f}{w},\ \frac{g}{w},\ \in S'_p(\Omega, w)\,. \quad (3.43)$$

Then, there exists a positive constant C, independent of u, such that, for any $B_r = B(x_0, r) \subset B(x_0, 4r) \subset \Omega$, one has

$$\sup_{B_r} u \leq C\left\{ \inf_{B_r} u + h(3r) \right\},$$

where

$$h(3r) = \left[\eta\left(\left(\frac{e}{w}\right)^{\frac{p}{p-1}};3r\right) + \eta\left(\frac{g}{w};3r\right)\right]^{\frac{1}{p}} + \left[\eta\left(\frac{f}{w};3r\right)\right]^{\frac{1}{p-1}}.$$

Proof. Start as in the proof of Theorem 3.10, setting $\tilde{u} = |u| + h$, where $h = h(3r)$. From this it follows that conditions (3.42) are verified. Now let φ be a nonnegative smooth function compactly supported in B_{3r}. Taking $\varphi^p(x)\tilde{u}^\beta(x)$, $\beta \in \mathbb{R}$ as test function in (3.39), it is obtained that

$$\begin{aligned} \int_{B_{3r}} |X\tilde{u}|^p \varphi^p \tilde{u}^{\beta-1} w\,dx \le \\ \le C_1(p,a)(1+|\beta|^{-1})^p \left\{ \int_{B_{3r}} |X\varphi|^p \tilde{u}^{p+\beta-1} w\,dx + \right. \\ \left. + \int_{B_{3r}} V\varphi^p \tilde{u}^{p+\beta-1}\,dx \right\}, \end{aligned} \tag{3.44}$$

where

$$V = \frac{b_1^{\frac{p}{p-1}}}{w^{\frac{1}{p-1}}} + \frac{c^p}{w^{p-1}} + d_1\,.$$

Setting

$$\mathscr{U}(x) \quad = \quad \begin{cases} \tilde{u}^q(x) & \text{where} \quad pq = p+\beta-1 \quad \text{if} \quad \beta \neq 1-p \\ \log \tilde{u}(x) & \text{if} \quad \beta = 1-p, \end{cases}$$

(3.44) yields

$$\begin{aligned} \int_{B_{3r}} \varphi^p |X\mathscr{U}|^p w\,dx \le C_1 |q|^p (1+|\beta|^{-1})^p \left\{ \int_{B_{3r}} |X\varphi|^p \mathscr{U}^p w\,dx + \right. \\ \left. + \int_{B_{3r}} V\varphi^p \mathscr{U}^p\,dx \right\}, \beta \neq 1-p, \end{aligned}$$

whereas

$$\int_{B_{3r}} \varphi^p |X\mathscr{U}|^p w\,dx \le C_1 \left\{ \int_{B_{3r}} |X\varphi|^p w\,dx + \int_{B_{3r}} V\varphi^p\,dx \right\} \text{ if } \beta = 1-p. \tag{3.45}$$

The result follows by arguing as in the proof of Theorem 3.5. $\square$

Next, the Harnack inequality is stated for nonnegative solutions of equations under natural growth. The proof is similar to the one in the linear case.

Theorem 3.12. *Let u be a local, bounded nonnegative weak solution of* (3.36). *Assume that the structural conditions* (3.38) *hold, with*

$$a \in \mathbb{R}, \left(\frac{b}{w}\right)^{p/p-1}, \left(\frac{c}{w}\right)^{p}, \frac{d}{w}, \left(\frac{e}{w}\right)^{p/p-1}, \frac{f}{w}, \frac{g}{w}, \in S'_p(\Omega, w). \quad (3.46)$$

Then, there exists a positive constant C, independent of u, such that, for any $B_r = B(x_0, r) \subset B(x_0, 4r) \subset \Omega$*, it holds that*

$$\sup_{B_r} u \le C\left\{\inf_{B_r} u + h(3r)\right\},$$

where

$$h(3r) = \left[\eta\left(\left(\frac{e}{w}\right)^{\frac{p}{p-1}}; 3r\right) + \eta\left(\frac{g}{w}; 3r\right)\right]^{\frac{1}{p}} + \left[\eta\left(\frac{f}{w}; 3r\right)\right]^{\frac{1}{p-1}}.$$

As shown next, regularity results are obtained directly from Harnack inequalities.

Theorem 3.13 (Regularity of weak solutions). *Let u be a weak solution of* (3.36). *Let us assume that the conditions* (3.37) *hold true with*

$$a \in \mathbb{R}, \left(\frac{b}{w}\right)^{p/p-1}, \left(\frac{c}{w}\right)^{p}, \left(\frac{d}{w}\right), \left(\frac{e}{w}\right)^{p/p-1}, \frac{f}{w}, \frac{g}{w} \in S'_p(\Omega, w).$$

Then u is continuous in Ω.

Moreover, assuming

$$\left(\frac{b}{w}\right)^{p/p-1}, \left(\frac{c}{w}\right)^{p}, \left(\frac{d}{w}\right), \left(\frac{e}{w}\right)^{p/p-1}, \frac{f}{w}, \frac{g}{w} \in M_\sigma(\Omega, w),$$

u is locally Hölder continuous in Ω.

Theorem 3.14 (Regularity of weak solutions). *Let u be a local, bounded, weak solution of* (3.36). *Assume that the conditions* (3.38) *hold with*

$$a \in \mathbb{R}, \left(\frac{b}{w}\right)^{p/p-1}, \left(\frac{c}{w}\right)^{p}, \left(\frac{d}{w}\right), \left(\frac{e}{w}\right)^{p/p-1}, \frac{f}{w}, \frac{g}{w} \in S'_p(\Omega, w).$$

Then u is continuous in Ω.

Moreover if

$$\left(\frac{b}{w}\right)^{p/p-1}, \left(\frac{c}{w}\right)^{p}, \left(\frac{d}{w}\right), \left(\frac{e}{w}\right)^{p/p-1}, \frac{f}{w}, \frac{g}{w} \in M_\sigma(\Omega, w),$$

then u is locally Hölder continuous in Ω.

Bibliography

[1] S. M. Buckley, *Inequalities of John-Nirenberg type in doubling spaces*, J. Anal. Math. **79** (1999), 215-240.

[2] F. Chiarenza, E. Fabes, N. Garofalo, *Harnack's inequality for Schrödinger operators and the continuity of solutions*, Proc. Amer. Math. Soc., **98** (1986) no. 3, 415-425.

[3] F. Chiarenza, M. Frasca, *A remark on a paper by C. Fefferman*, Proc. Amer. Math. Soc. **108** (1990), 407-409.

[4] D. Danielli, *A Fefferman - Phong inequality and applications to quasilinear subelliptic equations*, Potential Analysis, **11** (1999), 387-413.

[5] D. Danielli, N. Garofalo, Nhieu, *Trace inequalities for Carnot-Carathéodory spaces and applications*, Ann. Scuola Norm. Sup. Pisa, **XXVII** (1998), 195-252.

[6] G. Di Fazio, *Dirichlet problem Characterixation of regularity*, Manuscripta Math., **84**, 47-56 (1994).

[7] G. Di Fazio, M. S. Fanciullo, P. Zamboni, *Harnack inequality and smoothness for quasilinear degenerate elliptic equations* J. Differential Equations **245** (2008), no. 10, 2939-2957.

[8] G. Di Fazio, M. S. Fanciullo, P. Zamboni, *Harnack inequality and regularity for degenerate quasilinear elliptic equations*, Math. Z. **264** (2010), no. 3, 679-695

[9] G. Di Fazio, M. S. Fanciullo, P. Zamboni, *Harnack inequality for a class of strongly degenerate elliptic operators formed by Hörmander vector fields*, Manuscripta Math. **135** (2011), no. 3-4, 361-380.

[10] G. Di Fazio, M. S. Fanciullo, P. Zamboni, *Regularity for a class of strongly degenerate quasilinear operators*, J. Differential Equations **255** (2013), no. 11, 3920-3939.

[11] G. Di Fazio, M. S. Fanciullo, P. Zamboni, *Sum operators and Fefferman - Phong inequalities* , Geometric Methods in PDE's, Springer INdAM Series, Vol. 13 (2015), 111-120.

[12] G. Di Fazio, M. S. Fanciullo, P. Zamboni, *Harnack inequality for degenerate elliptic equations and sum operators*, Comm. on Pure and Applied Analysis **14**, n. 6, (2015), 2363-2376.

[13] G. Di Fazio, M. S. Fanciullo, P. Zamboni, *Harnack inequality and continuity of weak solutions for doubly degenerate elliptic equations*, Mathematische Zeitschrift 292 (2019), 1325-1336.

[14] G. Di Fazio, M. S. Fanciullo, P. Zamboni, *Local regularity for strongly degenerate elliptic equations and weighted sum operators*, Differential and Integral Equations, Volume 32, Numbers 7-8, July/August 2019, 479-492.

[15] G. Di Fazio, M. S. Fanciullo, P. Zamboni, *Local behavior for solutions of quasilinear elliptic equations related to weighted sum operators*, Nonlinear Analysis (2019), DOI: 0.1016/j.na.2019.07.003.

[16] G. Di Fazio, P. Zamboni, *A Fefferman-Poincaré type inequality for Carnot-Carathéodory vector fields*, Proc. Amer. Math. Soc. **130** (2002), no. 9, 2655-2660.

[17] G. Di Fazio, P. Zamboni, *Hölder continuity for quasilinear subelliptic equations in Carnot Carathéodory spaces,* Math. Nachr. **272** (2004), 3-10.

[18] G. Di Fazio, P. Zamboni, *Regularity for quasilinear degenerate elliptic equations*, Math. Z. **253** (2006), no. 4, 787-803.

[19] G. Di Fazio, P. Zamboni, *Local regularity of solutions to quasilinear subelliptic equations in Carnot Carathéodory spaces,* Boll. Unione Mat. Ital. Sex. B Artic. Ric. Mat. (8) 9 (2006), no. 2, 485-504.

[20] E. Fabes, C. Kenig, R. Serapioni, *The local regularity of solutions of degenerate elliptic equations*, Comm. Partial Differential Equations, **7(1)** (1982), 77-116.

[21] B. Franchi, C. Perez, R. L. Wheeden, *A Sum Operator with Applications to Self Improving Properties of Poincaré Inequalities in Metric Spaces*, The Journal of Fourier Analysis and Applications **9** (2003), Issue 5, 511-540.

[22] B. Franchi, R. Serapioni, F. Serra Cassano, *Approximation and imbedding theorems for weighted Sobolev spaces associated with Lipschitx continuous vector fields*, Boll. Un. Mat. Ital. B (7) **11** (1997), no. 1, 83-117.

[23] D. Gilbarg, N.S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin, 1983.

[24] C. E. Gutierrez, *Harnack's inequality for degenerate Schrödinger operators*, Trans. AMS, **312**, (1989), 403-419.

[25] P. Hajlasz, P. Koskela, *Sobolev met Poincaré*, Mem. Amer. Math. Soc. 688, (2000).

[26] J. Heinonen, T. Kilpeläinen, O. Martio, *Nonlinear potential theory of degenerate elliptic equations*. Oxford Mathematical Monographs. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, (1993)

[27] G. Lieberman, *Sharp forms of Estimates for Subsolutions and Supersolutions of Quasilinear Elliptic Equations Involving Measures*, Comm. P.D.E. 18 (1993), 1191-1212.

[28] G. Lu, *On Harnack's inequality for a class of strongly degenerate Schrödinger operators formed by vector fields*, Differential And Integral Equations, **7**, no. 1, (1994), 73-100.

[29] G. Lu, *A Fefferman-Phong type inequality for degenerate vector fields and applications*, Panamer. Math. J., **6** (1996), no. 4, 37-57.

[30] M. A. Ragusa, P. Zamboni, *Local regularity of solutions to quasilinear elliptic equations with general structure*, Commun. Appl. Anal., **3** (1999), no. 1, 131-147.

[31] J. M. Rakotoson - W. P. Ziemer, *Local behavior of solutions of quasilinear elliptic equations with general structure*, Trans. Amer. Math. Soc., **319** (1990), no. 2, 747-764.

[32] J. Serrin, *Local behavior of solutions of quasilinear equations*, Acta Math. **111** (1964), 247-302.

[33] N. S. Trudinger, *On Harnack Type Inequalities and Their Application to Quasilinear Elliptic Equations*, Comm. on pure and applied Mathematics, **20**, (1967), 721-747.

[34] P. Zamboni, *Some function spaces and elliptic partial differential equations*, Le Matematiche (Catania) **42** (1987), no. 1-2, 171-178.

[35] P. Zamboni, *The Harnack inequality for quasilinear elliptic equations under minimal assumptions*, Manuscripta Math., **102** (2000), no.3, 311-323.

[36] P. Zamboni, *Unique continuation for non-negative solutions of quasilinear elliptic equations*, Bull. Austral. Math. Soc. **64** (2001), no. 1, 149-156.

[37] P. Zamboni, *Hölder continuity for solutions of linear degenerate elliptic equations under minimal assumptions*, J. Differential Equations **182** (2002), no. 1, 121-140.

Chapter 4

Sobolev embeddings for spaces of Musielak-Orlicz type

J. Lang

Department of Mathematics, The Ohio State University, Columbus, OH, USA,
E-mail: lang.162@osu.edu

O. Méndez

Department of Mathematical Sciences, University of Texas at El Paso, El Paso, USA,
E-mail: osmendez@utep.edu

We report recent results, examples and countererexamples on compactness of Sobolev embeddings in the context of Musielak-Orlicz spaces.

4.1 Introduction and historical context

Twenty one years after F. Riesz' ground-breaking "Untersuchungen Über Systeme Integrierbarer Funktionen" [22], W. Orlicz, while dealing with a seemingly simple question regarding lacunary sequences, came to the realization of the fact that much of the structure of the L^p spaces introduced by Riesz could be preserved by allowing a variable exponent p. In specific terms, Orlicz considered the class of functions f defined on the interval $[a,b]$ such that

$$\int_a^b |f(x)|^{\alpha(x)} dx < \infty \tag{4.1}$$

and referred to such functions as integrable with respect to $\alpha(x)$. To the authors' best knowledge, his is the first occurrence in the literature of the class of functions that are integrable with respect to variable exponents. The main difficulty inherent to Orlicz' definition may have been the lack of a suitable norm to study the new structure in the light of the by then novel concept of B space. Orlicz himself did not continue studying these spaces immediately. In fact the generalization of the L^p class (today known as the Orlicz' class) that, shortly after the publication of [24], Orlicz introduced in [20], did not include the variable exponent class. It wasn't until 1950 that Nakano [21], while addressing a series of questions not directly related to Orlicz' line of work, came up with an example that generalized the Orlicz' class and did contain the variable exponent case. In 1961, Sharapudinov [24] studied the variable exponent class on $[0,1]$. Notably, he introduced the Luxemburg norm on this space and proved its reflexivity under the assumption that the exponent be bounded away from 1 and ∞. The first systematic treatment of Lebesgue variable exponent spaces was the work by O. Kováčik and J. Rákosník, [35]. In it, questions about the reflexivity and duality of $L^{p(\cdot)}(\Omega)$ (under the Luxemburg norm) were addressed in the case of an open subset $\Omega \subset \mathbb{R}^n$ and an admissible variable exponent

$$p : \Omega \longrightarrow \mathbb{R}.$$

4.2 Modular spaces

Let V be a real or complex vector space over the scalar field $\mathbb{K}$.

Definition 4.1. A convex modular on V is a function

$$\rho : V \longrightarrow [0,\infty]$$

that satisfies the following conditions:

(i) $\rho(x) = 0 \iff x = 0$

(ii) $\rho(\alpha x) = |\alpha|\rho(x)$, for any $x \in V$, $|\alpha| = 1$

(iii) $\rho(\alpha x + (1-\alpha)y) \leq \alpha\rho(x) + (1-\alpha)\rho(y)$, for all $x, y \in V$ and $\alpha \in (0,1]$.

A convex modular ρ on a vector space V is left-(right-) continuous if for any $x \in V$ the map

$$\alpha \longrightarrow \rho(\alpha x)$$

is left-(right-) continuous on $[0,\infty)$ (on $(0,\infty)$ for left continuity); if ρ is both left- and right-continuous we refer to it as a continuous modular.

Example 4.1. If $(\Omega, \mathscr{A}, \mu)$ is a measure space and $p \in [1, \infty)$, then

$$\rho(u) = \int_{\Omega} |u|^p \, d\mu \tag{4.2}$$

defines a convex modular on $L^p(\Omega, d\mu)$.

Example 4.2. Any norm on a real or complex vector space X is a convex modular on X.

It follows from the above axioms that for each fixed $x \in$ the map

$$\alpha \longrightarrow \rho(\alpha x)$$

is non-decreasing on $[0, \infty)$, for if $0 \leq \alpha < \beta$ and $x \in V$, then convexity yields:

$$\rho(\alpha x) = \rho(\alpha \beta^{-1} \beta x) \leq \alpha \beta^{-1} \rho(\beta x) \leq \rho(\beta x).$$

Consequently:

$$\rho(\alpha x) \quad = \rho(|\alpha| x) \leq |\alpha| \rho(x) \text{ if } |\alpha| \leq 1 \tag{4.3}$$

$$\rho(\alpha x) \quad = \rho(|\alpha| x) \geq |\alpha| \rho(x) \text{ if } |\alpha| \geq 1. \tag{4.4}$$

To the effect of characterizing the modular space associated to ρ we prove the following Lemma:

Lemma 4.1. *Let ρ be a modular on a vector space V and $x \in V$. Then conditions (i) and (ii) below are equivalent:*

(i) $\rho(\lambda x) < \infty$ *for some* $\lambda > 0$.

(ii) $\lim\limits_{\lambda \to 0^+} \rho(\lambda x) = 0$.

Proof. If (i) holds for $\lambda > 0$, and $0 < \lambda_j \to 0$ as $j \to \infty$, then there exists $J \in \mathbb{N}$ such that $j \geq J \Rightarrow \lambda_j < \lambda$. Since $\rho(0) = 0$, convexity yields, for $j \geq J$:

$$\rho(\lambda_j x) \leq \lambda_j \lambda^{-1} \rho(\lambda x) \to 0 \text{ as } j \to \infty.$$

The arbitrariness of the sequence (λ_j) yields the implication $(i) \Rightarrow (ii)$. Conversely if (ii) holds, there must be $\theta > 0$ such that $\rho(\lambda) < 1$ for all $\lambda < \theta$, whence (i) holds. $\square$

The preceding Lemma justifies the equality of the two sets in the next Definition:

Definition 4.2. For a modular ρ on a vector space V we set:

$$\begin{aligned} V_\rho \quad &:= \{x \in V : \rho(\lambda x) < \infty \text{ for some } \lambda > 0\} \\ &= \left\{ x \in V : \lim_{\lambda \longrightarrow 0^+} \rho(\lambda x) = 0 \right\}. \end{aligned}$$

Proposition 4.1. *Let ρ be a convex modular on a linear space X and define*

$$\|x\|_\rho := \inf\{\lambda > 0 : \rho(x/\lambda) \leq 1\} \ (x \in V_\rho).$$

Then:

(i) $\left(V_\rho, \|\cdot\|_\rho\right)$ *is a normed linear space.*

(ii) If $\rho(x) \leq 1$, then $\|x\|_\rho \leq 1$.

(iii) If ρ is left-continuous, then $\|x\|_\rho \leq 1$ if and only if $\rho(x) \leq 1$. If ρ is continuous, then $\|x\|_\rho < 1$ if and only if $\rho(x) < 1$; and $\|x\|_\rho = 1$ if and only if $\rho(x) = 1$.

Proof. See [14]. □

Corollary 4.1. *Let ρ be a left-continuous modular on a modular space V. Then:*

(i) *If $\|x\|_\rho \leq 1$, then $\rho(x) \leq \|x\|_\rho$.*

(ii) *If $\|x\|_\rho > 1$, then $\rho(x) \geq \|x\|_\rho$.*

(iii) *For any $x \in V$, $\|x\|_\rho \leq \rho(x) + 1$.*

Proof. For (i), observe that if $0 < \|x\|_\rho \leq 1$ then the convexity of ρ yields

$$\begin{aligned} \rho(x) = \rho\left(\frac{\|x\|_\rho}{\|x\|_\rho} x\right) &\leq \|x\|_\rho \rho\left(\frac{x}{\|x\|_\rho}\right) \\ &\leq \|x\|_\rho; \end{aligned}$$

on the other hand if $1 < \lambda < \|x\|_\rho$ one has

$$\rho\left(\frac{x}{\lambda}\right) > 1$$

and by convexity one gets $1 < \lambda^{-1}\rho(x)$, which forces $\rho(x) \geq \|x\|_\rho$, as claimed. (iii) follows immediately from (i) and (ii). □

We finish this Section with the following Definition:

Definition 4.3. The modular ρ on the vector space V is said to satisfy the Δ_2-condition if there exists a constant $K \geq 1$ such that for any $x \in V$

$$\rho(2x) \leq K\rho(x).$$

4.3 Musielak-Orlicz spaces

Next, the Musielak-Orlicz spaces are introduced as a specific example fitting the modular space theory described in the preceding Section.

Let $\emptyset \neq \Omega \subset \mathbb{R}^n$, $(n \geq 1)$ be a domain (i.e., open and connected). An Orlicz function on Ω is a convex, left-continuous function

$$\varphi : [0,\infty) \longrightarrow [0,\infty)$$

with $\varphi(0) = 0$, $\lim\limits_{x\to\infty} \varphi(x) = \infty$ and $\lim\limits_{x\to 0^+} \varphi(x) = 0$. In particular, any Orlicz function is non-decreasing. A Musielak-Orlicz function on Ω is a function

$$\varphi : \Omega \times [0,\infty) \to [0,\infty)$$

such that

$$\varphi(x,\cdot) : [0,\infty) \to [0,\infty)$$

is an Orlicz function for each fixed $x \in \Omega$ and that

$$\varphi(\cdot,y) : \Omega \to [0,\infty)$$

is Lebesgue measurable for each fixed $y \in \mathbb{R}$.
There is convex modular associated to any Musielak-Orlicz function φ, namely:

$$\rho_\varphi(u) := \int\limits_\Omega \varphi(x,|u(x)|)\,dx$$

The Musielak-Orlicz space $L^\varphi(\Omega)$ is the real-vector space of all extended-real valued, Borel-measurable functions u on Ω for which

$$\rho_\varphi(\lambda u) := \int\limits_\Omega \varphi(x,|u(x)|\lambda)\,dx < \infty \text{ for some } \lambda > 0.$$

$L^\varphi(\Omega)$ is furnished with the norm

$$\|u\|_\varphi = \inf\left\{\lambda > 0 : \int_\Omega \varphi\left(x, \frac{|u(x)|}{\lambda}\right)dx \leq 1\right\}.$$

We refer the reader to ([7, 20, 14]) for the proof of the fact that $L^\varphi(\Omega)$ is a Banach space under $\|u\|_\varphi$.
The Musielak-Orlicz Sobolev space $W^{1,\varphi}(\Omega)$ is the vector space of all functions in $L^\varphi(\Omega)$ whose distributional derivatives are in $L^\varphi(\Omega)$, furnished with the norm

$$\|u\|_{1,\varphi} = \|u\|_\varphi + \||\nabla u|\|_\varphi,$$

Here $\nabla\cdot$ stands for the gradient operator, i.e.,

$$\nabla u = \left(\frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_n}\right)$$

and $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^n$. It is well known (see [7, 20]) that $W^{1,\varphi}(\Omega)$ $W^{1,\varphi}(\Omega)$ is a Banach space under the assumptions

$$\int_K \varphi(x,t)\,dx < \infty \tag{4.5}$$

for any $K \subset \Omega$ with Lebesgue measure $|K| < \infty$ and satisfying the condition

$$\inf_{x\in\Omega} \varphi(x,1) > 0. \tag{4.6}$$

The Sobolev space $W_0^{1,\varphi}(\Omega)$ is defined to be the closure of $C_0^\infty(\Omega)$ in $W^{1,\varphi}(\Omega)$.

Example 4.3. Consider the family of admissible exponent functions

$$\begin{aligned}\mathscr{P} \quad = \quad & \Big\{ p : \Omega \to \mathbb{R},\, p \text{ Borel-measurable,} \\ & \quad 1 < \inf_{x\in\Omega} p(x) = p_- \leq \sup_{x\in\Omega} p(x) p_+ < \infty \Big\}. \end{aligned} \tag{4.7}$$

In this case one can set

$$\varphi(x,t) = t^{p(x)}.$$

It is clear that φ is a Musielak-Orlicz function. In this case it is customary to denote $L^\varphi(\Omega)$ by $L^{p(\cdot)}(\Omega)$ and it is easy to prove that under the restrictions on the exponent p, $L^{p(\cdot)}(\Omega)$ is the set of all real-valued, Borel measurable functions on Ω for which

$$\rho_p(f) := \int_\Omega |f(x)|^{p(x)}\,dx < \infty.$$

As before, the function ρ_p is a convex, monotone, continuous modular on $L^{p(\cdot)}(\Omega)$ and

$$\|u\|_{L^{p(\cdot)}(\Omega)} := \inf\left\{\lambda > 0 : \rho_p\left(\frac{u}{\lambda}\right) \leq 1\right\}$$

defines a norm under which $L^{p(\cdot)}(\Omega)$ is a Banach, reflexive, uniformly convex space (see [7, 11]). It is apparent that the latter coincides with the usual Lebesgue $L^p(\Omega)$ norm when $p(\cdot)$ is constant; accordingly the family of

$L^{p(\cdot)}(\Omega)$ for p as in (4.7) will be referred to as the generalized Lebesgue class in Ω. The generalized Sobolev class in Ω for this particular Musielak-Orlicz function is denoted as

$$W^{1,p(\cdot)}(\Omega) := \left\{ u \in L^{p(\cdot)}(\Omega) : |\nabla u| \in L^{p(\cdot)}(\Omega) \right\}$$

and will be endowed with the norm

$$\|u\|_{W^{1,p(\cdot)}(\Omega)} := \|u\|_{L^{p(\cdot)}(\Omega)} + \||\nabla u|\|_{L^{p(\cdot)}(\Omega)}.$$

The closure of $C_0^\infty(\Omega)$ in $W^{1,p(\cdot)}(\Omega)$ is written as $W_0^{1,p(\cdot)}(\Omega)$.

The reader is referred to [7] and [35] for an exhaustive treatment of variable-exponent Lebesgue-Sobolev spaces.
Under the restrictions posed on the exponent p, the following inequalities hold for any $w \in L^{p(\cdot)}(\Omega)$ ([9]):

$$\min\left\{ \rho_p^{\frac{1}{p+}}(w), \rho_p^{\frac{1}{p-}}(w) \right\} \leq \|w\|_{p(\cdot)} \leq \max\left\{ \rho_p^{\frac{1}{p+}}(w), \rho_p^{\frac{1}{p-}}(w) \right\}. \tag{4.8}$$

4.3.1 Sobolev-type embeddings

It is a result from classical analysis [4, 6] that for $1 < p < \infty$, the natural Sobolev embedding

$$W_0^{1,p}(\Omega) \hookrightarrow L^p(\Omega) \tag{4.9}$$

is compact. Such compactness results are essential in a number of mathematical situations in which one needs to extract a convergent subsequence from a bounded sequence.

We refer the reader to [1, 4, 6, 25] for the following classical embedding theorems:

Theorem 4.1. *Let $\Omega \subset \mathbf{R}^n$ be a bounded domain, $p \in (1,n)$. Then the space $W_0^{1,p}(\Omega)$ is compactly embedded in $L^q(\Omega)$ for any $q \in (1, \frac{np}{n-p})$. If $p > n$, $W_0^{1,p}(\Omega)$ embeds compactly in $C(\overline{\Omega})$ (and hence in $L^q(\Omega)$ for any $q \in (1,\infty)$). Finally, $W_0^{1,n}(\Omega)$ is compactly embedded in $L^q(\Omega)$ for any $q \in [1,\infty)$. Moreover, for any $p \in (1,\infty)$, the embedding*

$$W_0^{1,p}(\Omega) \hookrightarrow L^p(\Omega) \tag{4.10}$$

is compact.

The extension of the compactness of the embedding (4.9) to the variable exponent case was given in [7, 35]. We refer the reader to [10] for a detailed proof. Specifically

Theorem 4.2. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded, Lipschitz domain and $p \in C(\overline{\Omega})$ be an admissible exponent satisfying $1 < p_- \le p_+ < \infty$. Then the embedding*

$$W_0^{1,p(\cdot)}(\Omega) \hookrightarrow L^{p(\cdot)}(\Omega) \tag{4.11}$$

is compact.

A natural question is whether a similar embedding theorem holds true in the general framework of the Musielak-Orlicz spaces.
The following examples, however, show that one can't expect Sobolev-type embeddings to hold even in very simple settings, without imposing some restrictions on the Musielak-Orlicz function.

Example 4.4. We consider Ω to be the Euclidean unit ball in $\mathbb{R}^6$ and for $n \in \mathbb{N}$ let B_n be the ball of radius 2^{-n-2} centered at $x_n = (2^{-n}, 0, 0, 0, 0, 0)$. The ball concentric with B_n of radius 2^{-n-3} is denoted by B_n^- and we set:

$$B_n^+ = B_n \setminus B_n^- .$$

For each $h > 0$, let the function

$$v_h : \mathbb{R}^6 \longrightarrow \mathbb{R}$$

be given as follows:

$$v_h(x) = \begin{cases} 1 & \text{if } |x| \le h \\ 2 - \frac{|x|}{h} & \text{if } |x| \in (h, 2h) \\ 0 & \text{if } |x| \in (2h, \infty). \end{cases}$$

Let w_6 stand for the Lebesgue measure of the unit sphere in $\mathbb{R}^6$ and consider the sequence (u_n) given by

$$u_n(x) = 2^{2(n+3)} (w_6)^{-\frac{1}{3}} v_{2^{-(n+3)}}(x - x_n).$$

For each $n \in \mathbb{N}$ fix $z_n = 2^{3(n+3)} \approx |\nabla u_n(x)|$ and consider

$$\phi_{z_n}(t) = \begin{cases} t^2 & \text{if } t \in (0, z_n) \\ 3s_n^2(t - z_n) + z_n^2 & \text{if } t \in (z_n, s_n) \\ t^3 & \text{if } t \in (s_n, \infty), \end{cases} \tag{4.12}$$

where for each $n \in \mathbb{N}$, $z_n = 2^{3(n+3)}$ and $s_n > z_n$ is selected in such a way that

$$s_n^3 - z_n^2 = 3s_n^2(s_n - z_n).$$

Then the Musielak-Orlicz function φ is set to be:

$$\varphi(x,t) = \begin{cases} t^3 & \text{if } x \in \Omega \setminus \bigcup_n B_n^+ \\ \phi_{z_n}(t) & \text{if } x \in B_n^+. \end{cases}$$

We next prove that the sequence (u_n) is bounded in $W_0^{1,\varphi}(\Omega)$. To this end observe that

$$\begin{aligned} \rho_\varphi(u_n) &= \left(\int_{B_n^-} + \int_{B_n^+} \right) \varphi\left(x, u_n(x)\right) dx = \int_{B_n^-} \varphi(x, 2^{2(n+3)} w_6^{-\frac{1}{3}}) dx \\ &\quad + \int_{B_n^+} \varphi\left(x, 2^{2(n+3)} w_6^{-\frac{1}{3}} \left(2 - 2^{n+3}|x - x_n|\right)\right) dx \\ &\leqq |B_n| \left(2^{2(n+3)} w_6^{-\frac{1}{3}}\right)^3 + \int_{B_n^+} \varphi\left(x, 2^{2(n+3)} w_6^{-\frac{1}{3}}\right) dx \\ &\leq C + |B_n| \left(2^{2(n+3)} w_6^{-\frac{1}{3}}\right)^2 \\ &\leq C_1 + C_2. \end{aligned}$$

In the above statement, C_1 is a positive constant independent of n and $C_2 \to 0$ as $n \to \infty$.

It is clear that $2^{3(n+3)w_6^{-1/3}} > s_n$ for large enough n. A straightforward computation reveals that while $\nabla u_n(x) = 0$ on $\Omega \setminus B_n^+$, for $x \in B_n^+$ one has:

$$|\nabla u_n(x)| \approx 2^{3(n+3)} w_6^{-1/3}.$$

It follows from this statement that

$$\begin{aligned} \rho_\varphi(|\nabla u_n|) &= \left(\int_{B_n^-} + \int_{B_n^+} \right) \varphi\left(x, u_n(x)\right) dx \\ &= \int_{B_n^+} \varphi\left(x, 2^{2(n+3)} w_6^{-\frac{1}{3}} \left(2 - 2^{n+3}|x - x_n|\right)\right) dx \\ &\approx |B_n| \left(2^{2(n+3)} w_6^{-1/3}\right)^3 \\ &\leq C \end{aligned}$$

for a positive constant C, independent of n. In all, it is clear that there are positive constants k_1, k_2 such that for all $n \in \mathbb{N}$, one has:

$$k_1 \leq \|u_n\|_\varphi \leq k_2$$
$$k_1 \leq \||\nabla u_n|\|_\varphi \leq k_2.$$

It follows then that the sequence (u_n) is bounded in $W_0^{1,\varphi}(\Omega)$. On the other hand, each u_n is continuous, any two different functions in the sequence (u_n) have disjoint supports and $\sup\limits_{\Omega} u_n \to \infty$ as $n \to \infty$. Hence, no subsequence of (u_n) converges in $L^\varphi(\Omega)$, and $W_0^{1,\varphi}(\Omega)$ is not compactly embedded in $L^\varphi(\Omega)$.

4.4 The Matuszewska-Orlicz index of a Musielak-Orlicz function

We next introduce a generalized version of the Matuszewska-Orlicz index, which is to play a fundamental role in our further developments. The Matuszewska-Orlicz index of an Orlicz function φ was introduced by Matuszewska and Orlicz in [13].

Definition 4.4. Let φ be a Musielak-Orlicz function. For each $x \in \Omega$, set

$$M(x,t) = \limsup_{u\to\infty} \frac{\varphi(x,tu)}{\varphi(x,u)}. \tag{4.13}$$

The Matuszewska-Orlicz index of φ is defined to be

$$m(x) = \lim_{t\to\infty} \frac{\ln M(x,t)}{\ln t} = \inf_{t>1} \frac{\ln M(x,t)}{\ln t}. \tag{4.14}$$

Definition 4.5. The limit (4.13) is said to be uniform if for each $\kappa > 0$ there exist $s_0 > 1$ and $T > 1$ such that, for all $(x,t) \in \Omega \times [T,\infty)$ and $s \geq s_0$ one has

$$M(x,t) - \kappa < \frac{\varphi(x,ts)}{\varphi(x,s)} < M(x,t) + \kappa.$$

As it will be seen, the behavior of the Matuszewska-Orlicz index in fact determines the validity of the Sobolev embedding.

Example 4.5. Let $\Omega \subseteq \mathbb{R}^n$ be a bounded domain and

$$p : \Omega \longrightarrow (0,\infty)$$

be Borel-measurable. The Musielak-Orlicz function

$$\varphi : \Omega \times [0,\infty) \longrightarrow [0,\infty)$$
$$\varphi(x,t) = t^{p(x)}$$

has Matuszewska index equal to $p(x)$. In this case, the convergence (4.13) is trivially uniform on Ω and the limit (4.14) is clearly uniform.

Example 4.6. Slightly less trivial is the uniform convergence of (4.13) for the Musielak-Orlicz function

$$\varphi(x,t) = t^{p(x)}(1+\log t)^{q(x)}$$

where p is as in Example 4.5 and

$$q : \Omega \longrightarrow \mathbb{R}$$

is a Borel-measurable, bounded function on Ω.

Example 4.7. It is a matter of course to verify that the Musielak-Orlicz function φ in Example 4.4 does not fulfill the Δ_2-condition . Indeed, for any $x \in B_n^+$

$$\varphi(x,z_n) = \varphi(x,2^{3(n+3)}) = \phi_{2^{3(n+3)}}\left(2^{3(n+3)}\right) = 2^{6(n+3)},$$

whereas

$$\begin{aligned}
\varphi(x,2z_n) &= \varphi(x,2^{3n+10}) = \phi_{2^{3(n+3)}}\left(2^{3n+10}\right) \\
&\geq 3s_n^2(2^{3n+10} - 2^{3(n+3)}) + 2^{6(n+3)} \\
&\geq 2^{3(n+3)}(3s_n^2 + 2^{3(n+3)}) \\
&= \varphi(x,z_n)\frac{3s_n^2 + 2^{3(n+3)}}{2^{3n+9}}. \qquad (4.15)
\end{aligned}$$

Since $s_n > 2^{3n+9}$, inequality (4.15) shows that φ fails the Δ_2-condition, as claimed.
On the other hand, it is easy to show that the Matuszewska index of φ is equal to 3 in Ω and that

$$\varphi(x,t) \longrightarrow t^2 \text{ as } x \longrightarrow (0,0,0,0,0,0), \text{ on } B_n^+$$
$$\varphi(x,t) \longrightarrow t^3 \text{ as } x \longrightarrow (0,0,0,0,0,0) \text{ on } \Omega \setminus \bigcup_{j\in\mathbb{N}} B_j^+.$$

Lemma 4.2. *If φ is an Musielak-Orlicz function for which the limits (4.13) and (4.14) are uniform, then there are constants $C_1 > 1$, $C_2 > 1$ and $S_0 > 1$ for which*

$$\varphi(x,C_1 s) \leq C_2\varphi(x,s) \text{ } for \text{ } s \geq S_0.$$

Proof. A straightforward calculation shows that if $\delta > 0$ then there exists a constant $C_1 > 1$ for which $t \geq C_1$ implies

$$M(x,t) < t^{m(x)+\delta};$$

the assumed uniformity of the limit yields the existence of $S_0 > 1$ for which

$$\sup_{s \geq S_0} \frac{\varphi(x,ts)}{\varphi(x,s)} < t^{m(x)+\delta} + \frac{1}{2} C_1^{[\sup_{x\in\Omega} m(x)+\delta]} \tag{4.16}$$

whenever $s \geq S_0$, $t \geq C_1$; in particular, setting $t = C_1$ in (4.16) one easily sees that for $s \geq S_0$ it holds that

$$\varphi(x, C_1 s) \leq \frac{3}{2} C_1^{[\sup_{x\in\Omega} m(x)+\delta]} \varphi(x,s), \tag{4.17}$$

whence the lemma follows immediately. $\square$

4.5 Soblev embedding for spaces of Musielak-Orlicz type

In this Section we state and prove the following version of the Sobolev embedding theorem for spaces of Musielak-Orlicz type.

Theorem 4.3. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain and*

$$\varphi : \Omega \times [0,\infty) \longrightarrow \mathbb{R}$$

be a locally integrable Musielak-Orlicz function. Assume that

(i)

$$\mathbf{essinf}_{x\in\Omega} \varphi(x,1) > 0.$$

(ii) The limits (4.13) and (4.14) are uniform on Ω.

(iii) The Matuszewska-Orlicz index m_φ is the restriction to Ω of a continuous function $\tilde{m}$ defined on the the closure of Ω.

(iv)

$$1 < m_- := \inf_{\Omega} m_\varphi.$$

(v) There exists a function

$$\beta : (0,\infty) \longrightarrow (0,\infty)$$

such that the inequality

$$\varphi(x,t) \leq \beta(t) \tag{4.18}$$

holds uniformly in Ω.

Then the embedding

$$W_0^{1,\varphi}(\Omega) \hookrightarrow L^{\varphi}(\Omega) \tag{4.19}$$

is compact.

Proof. Suppose first that $m_- < n$. Observe that for γ such that

$$0 < \gamma < \frac{m_-(n-m_-)}{2n-m_-},$$

one has:

$$\frac{m_- - \gamma}{n - m_- + \gamma} > \frac{1}{2}\frac{m_-}{n-m_-}.$$

Set

$$\mu := \frac{1}{2}\frac{m_-}{n-m_-}.$$

Let $r > 0$ be so small that the inequality

$$\frac{n(n-r)}{r} > n + 2r \tag{4.20}$$

holds and write

$$\varepsilon < \min\{\mu, 5r\}\ ,\ 0 < \gamma < \min\left\{\frac{\varepsilon}{30}, \frac{m_-(n-m_-)}{2n-m_-}, \frac{m_- - 1}{2}\right\}. \tag{4.21}$$

It is then clear that

$$w(x) := m(x) - \gamma < m(x) < m(x) - \gamma + \frac{\varepsilon}{20} = w(x) + \frac{\varepsilon}{20}.$$

From the uniformity conditions (4.13) and (4.14) it follows the existence of γ satisfying the second inequality in (4.21) and of a constant $T_0 > 1$ such that $t \geq T_0$ implies that the following inequality holds uniformly for all $t \geq T_0$, $x \in \Omega$:

$$t^{m(x)-\gamma} = t^{w(x)} < M(x,t) < t^{m(x)-\gamma+\varepsilon/20} = t^{w(x)+\varepsilon/20}.$$

The uniformity of the infimum (4.14) with respect to t and x yields a positive number S_0, which can be assumed to be greater than 1, such that, for all $(x,t) \in \Omega \times [0,\infty)$, one has, for $t \geq T_0 > 1$ and any δ such that

$$\gamma < \delta < \frac{\varepsilon}{20}:$$

$$M(x,t) - \frac{1}{2}T_0^{m_- - \delta} < \frac{\varphi(x,tS_0)}{\varphi(x,S_0)} < M(x,t) + \frac{1}{2}T_0^{m_- - \delta}.$$

In all, the inequalities

$$\frac{1}{2}t^{w(x)} < \frac{\varphi(x,tS_0)}{\varphi(x,S_0)} < \frac{3}{2}t^{w(x)+\varepsilon/20}$$

hold uniformly in Ω for any $t \geq T_0$. Setting $tS_0 = s$, it is readily seen that for $s \geq T_0 S_0$, one has

$$\frac{1}{2}\varphi(x,S_0)\left(\frac{s}{S_0}\right)^{w(x)} \leq \varphi(x,s) \leq \frac{3}{2}\varphi(x,S_0)\left(\frac{s}{S_0}\right)^{w(x)+\frac{\varepsilon}{20}}. \tag{4.22}$$

Furthermore, assumptions (i) and (v) yield positive constants A, B such that for all $x \in \Omega$ one has

$$(\varphi(x,S_0))^{\frac{1}{w(x)}} \leq \left(\sup_{\Omega} \varphi(x,S_0)\right)^{\frac{1}{w(x)}} \leq B$$

and

$$A \leq \left(\inf_{\Omega} \varphi(x,S_0)\right)^{\frac{1}{w(x)}} \leq (\varphi(x,S_0))^{\frac{1}{w(x)}}.$$

Consequently, (4.22) implies the existence of positive constants c_1, c_2 for which

$$c_1 s^{w(x)} < \varphi(x,s) < c_2 s^{w(x)+\frac{\varepsilon}{20}}, \tag{4.23}$$

valid for all $x \in \Omega$ and $s \geq S_0 > 1$. By assumption and by virtue of Tietze's extension theorem, w is the restriction to Ω of a continuous function

$$p : \mathbb{R}^n \longrightarrow [w_-, w_+].$$

Let $p_1 = w_-$. For $k > 1$ if $p_{k-1} < n$, set

$$p_k = \frac{np_{k-1}}{n - p_{k-1}} - \frac{\varepsilon}{5}.$$

Clearly, if $p_{j-1} < n$, then $p_j > p_1 + (j-1)\frac{4}{5}\varepsilon$ and $p_{j-1} < p_j$. Let J be the first subindex for which $p_J > n - \frac{r}{2}$, where r is as in (4.20). Let $I = [w_-, w_+]$ and

$$\Omega_1 := p^{-1}\left(\left(\frac{w_- + 1}{2}, \frac{np_1}{n-p_1} - \frac{\varepsilon}{10}\right) \cap I\right).$$

For $1 < k \leq J-1$ set

$$\Omega_k := p^{-1}\left(\left(p_k, \frac{np_k}{n-p_k} - \frac{\varepsilon}{10}\right) \cap I\right);$$

furthermore, define Ω_J and Ω_{J+1} as

$$\Omega_J := p^{-1}\left((n-r, n+r) \cap I\right)$$
$$\Omega_{J+1} := p^{-1}\left(\left(n + \frac{r}{2}, \infty\right) \cap I\right)$$

and let $(\chi_k)_{1 \leq k \leq J+1}$ be a partition of unity subordinated to the cover $(\Omega_k)_k$ of Ω. A straightforward argument shows that if $v \in C_0^\infty(\Omega)$ (which can be considered extended by 0 to $\mathbb{R}^n$) then, for each $k : 1 \leq k \leq J+1$, $v\chi_k \in C_0^\infty(\Omega \cap \Omega_k)$. It follows from this observation that if $v \in W_0^{1,w}(\Omega)$, then for each fixed k,

$$v\chi_k \in W_0^{1,w}(\Omega \cap \Omega_k).$$

Fix a sequence (u_j) bounded in $W_0^{1,\varphi}(\Omega)$; then inequalities (4.23) in concert with a simple calculation imply that (u_j) is bounded in $W_0^{1,w}(\Omega)$. We contend that $(u_j\chi_k)_j$ is bounded in $W_0^{1,p_k}(\Omega)$, for any subindex $k : 1 \leq k \leq J-1$. Denote the indicator function of any set A by I_A. Then by construction

$$w_k := wI_{\Omega\cap\Omega_k} \geq p_k I_{\Omega\cap\Omega_k}$$

so the embedding

$$W_0^{1,w_k}(\Omega) \hookrightarrow W_0^{1,p_k I_{\Omega\cap\Omega_k}}(\Omega) \tag{4.24}$$

is continuous, that is, for some positive constant C

$$\|u_j\chi_k\|_{W_0^{1,p_k I_{\Omega\cap\Omega_k}}(\Omega)} \leq C\|u_j\chi_k\|_{W_0^{1,w_k}(\Omega)}.$$

On the other hand, if F_{kj} stands for any of the functions $u_j\chi_k$, $(\nabla u_j)\chi_k$ or $u_j\nabla\chi_k$, it is clear that

$$1 = \int_\Omega \left|\frac{|F_{kj}|}{\|F_{kj}\|_{p_k}}\right|^{p_k} = \int_\Omega \left|\frac{|F_{kj}|}{\|F_{kj}\|_{p_k}}\right|^{p_k I_{\Omega\cap\Omega_k}}$$

and with the same token, that

$$1 = \int_{\Omega} \left| \frac{|F_{kj}|}{\|F_{kj}\|_{w_k}} \right|^{w_k} = \int_{\Omega} \left| \frac{|F_{kj}|}{\|F_{kj}\|_{w_k}} \right|^{w}. \tag{4.25}$$

The two preceding observations and (4.24) yield

$$\|F_{kj}\|_{p_k} = \|F_{kj}\|_{p_k I_{\Omega\cap\Omega_k}} \leq C\|F_{kj}\|_{w_k} = \|F_{kj}\|_{w},$$

whence the contention follows.

Hence, by vitue of Theorem 4.1 there is no loss of generality in assuming that $(u_j\chi_k)_j$ converges in $L^{\frac{np_k}{n-p_k}-\frac{\varepsilon}{20}}(\Omega)$. For simlicity, let q_k be the right-endpoint of $p(\Omega_k)$ for $1 \leq k \leq J$. Next, if $1 \leq j \leq J$, set

$$d_j := \left(q_j + \frac{\varepsilon}{20}\right) I_{\Omega_j} + \left(w_+ + \frac{\varepsilon}{20}\right) I_{\Omega\setminus\Omega_j}.$$

Then $d_j \geq w + \frac{\varepsilon}{20}$ for all $x \in \Omega$ and one has the continuous embedding

$$L^{d_j}(\Omega) \hookrightarrow L^{w+\frac{\varepsilon}{20}}(\Omega). \tag{4.26}$$

For any function $u \in W_0^{1,\varphi}(\Omega)$ and $1 \leq k \leq J-1$:

$$\int_{\Omega} |u\chi_k|^{d_k} = \int_{\Omega} |u\chi_k|^{d_k I_{\Omega_k}} = \int_{\Omega} |u\chi_k|^{q_k+\frac{\varepsilon}{20}} = \int_{\Omega} |u\chi_k|^{\frac{np_k}{n-p_k}-\frac{\varepsilon}{20}}. \tag{4.27}$$

The preceding string of inequalities yields the following observation:

If $(u_j\chi_k)_j$ is a Cauchy sequence in $L^{\frac{np_k}{n-p_k}-\frac{\varepsilon}{20}}(\Omega)$, $1 \leq k \leq J-1$ then it is convergent in $L^{d_k}(\Omega)$ and by virtue of (4.26), $(u_j\chi_k)_j$ converges in $L^{w+\frac{\varepsilon}{20}}(\Omega)$.

We claim that the latter observation yields the convergence of $(u_j\chi_k)_j$ in $L^{\varphi}(\Omega)$ for $1 \leq k \leq J-1$. Indeed, there is no loss of generality by assuming that $(u_j\chi_k)_j$ converges pointwise $a.e$; on the other hand:

$$\int_{\Omega} \varphi(x, |u_j(x) - u_i(x)|\chi_k(x))\,dx = \tag{4.28}$$

$$\int_{\{x:|u_j(x)-u_i(x)|\chi_k(x)\leq S_0\}} \varphi(x, |u_j(x) - u_i(x)|\chi_k(x))\,dx +$$

$$\int_{\{x:|u_j(x)-u_i(x)|\chi_k(x)>S_0\}} \varphi(x, |u_j(x) - u_i(x)|\chi_k(x))\,dx.$$

Since for any fixed $x \in \Omega$, the function $\varphi(x,\cdot)$ is nondecreasing, the integrand in the first term above satisfies the inequality

$$\varphi(x,|u_j(x)-u_i(x)|\chi_k(x)) \leq \varphi(x,S_0).$$

The assumption of local integrability on φ in conjunction with a straightforward application of Lebesgue's dominated convergence yields

$$\lim_{i,j\to\infty} \int\limits_{\{x:|u_j(x)-u_i(x)|\chi_k(x)\leq S_0\}} \varphi(x,|u_j(x)-u_i(x)|\chi_k(x))\,dx = 0.$$

Since $S_0 > 1$, the second integral in (4.28) is dominated by

$$\int\limits_{\{x:|u_j(x)-u_i(x)|\chi_k(x)>S_0\}} |u_j(x)-u_i(x)|^{w(x)+\frac{\varepsilon}{20}}\chi_k(x)\,dx.$$

In all, $\left(\rho_\varphi((u_j-u_i)\chi_k)_j\right) \longrightarrow 0$ as $i,j \longrightarrow \infty$. Next, we observe that for C_1 as in the statement of Lemma 4.2 one has:

$$\begin{aligned}\rho_\varphi(C_1(u_i-u_j)\chi_k) = \\ &\int\limits_{\{x:|u_j(x)-u_i(x)|\chi_k(x)\leq S_0\}} \varphi(x,C_1|u_j-u_i|\chi_k))\,dx \\ &+ \int\limits_{\{x:|u_j(x)-u_i(x)|\chi_k(x)>S_0\}} \varphi(x,C_1|u_j-u_i|\chi_k))\,dx;\end{aligned}$$

a straightforward application of Lebesgue's dominated convergence theorem on the first integral and the consideration of Lemma 4.2 in the second one easily yield

$$\rho_\varphi(C_1(u_i-u_j)\chi_k) \longrightarrow 0 \ as \ i,j \longrightarrow \infty.$$

It follows automatically by induction that for any $l \in \mathbb{N}$ one has

$$\rho_\varphi(C_1^l(u_i-u_j)\chi_k) \longrightarrow 0 \ as \ i,j \longrightarrow \infty,$$

and it is concluded from here that the sequence $(u_j\chi_k)_j$ is Cauchy in $L^\varphi(\Omega)$, as claimed.

The remaining intervals in the covering are handled similarly: define

$$w_J := wI_{\Omega\cap\Omega_J} \geq (n-r)I_{\Omega\cap\Omega_J};$$

then the embedding

$$W_0^{1,w_J}(\Omega) \hookrightarrow W_0^{1,n-r}(\Omega)$$

is bounded. Retaining the notation of the above discussion, the sequence $(u_j\chi_J)_j$ is bounded in $W_0^{1,n-r}(\Omega)$; without loss of generality it can be considered convergent in $L^{\frac{n(n-r)}{r}}(\Omega)$, which by the choice of r in (4.20) is continuously embedded in $L^{n+2r}(\Omega)$. Setting

$$h = (n+2r)I_{\Omega_J} + \left(w_+ + \frac{\varepsilon}{20}\right) I_{\Omega\setminus\Omega_J}$$

it is clear that $L^h(\Omega)$ is continuously embedded in $L^{w+\frac{\varepsilon}{20}}(\Omega)$. It follows immediately that $(u_j\chi_J)_j$ is Cauchy in the latter space. Theorem 4.3 ensures now that $(u_j\chi_J)_j$ is convergent in $L^\varphi(\Omega)$.

Finally, via the continuous embeddings

$$W_0^{1,\varphi}(\Omega) \hookrightarrow W_0^{1,w}(\Omega) \hookrightarrow W_0^{1,(n+\frac{r}{2})I_{\Omega_{J+1}}+w_-I_{\Omega\setminus\Omega J+1}}(\Omega)$$

the boundedness of $(u_j\chi_{J+1})_j$ in $W_0^{1,\varphi}(\Omega)$ yields its boundedness in $W_0^{1,n+\frac{r}{2}}(\Omega)$ and by way of Theorem 4.1 it is readily concluded that $(u_j\chi_{J+1})_j$ can be considered convergent in $C(\overline{\Omega})$, hence convergent in $L^\varphi(\Omega)$.

In all, for $m_- < n$ any bounded sequence $(u_j)_j \subset W_0^{1,\varphi}(\Omega)$ has a subsequence that converges in $L^\varphi(\Omega)$.

The case $n \leq m_-$ follows similarly; we only sketch the proof in this case.

For r as in (4.20), there exists $T_0 > 1$ such that uniformly on Ω and for all $t \geq T_0$ it holds that

$$t^{m(x)-\frac{r}{4}} < M(x,t) < t^{m(x)+r}.$$

It follows as earlier that given the conditions on the index, there are positive constants $c_1 > 1$, $c_2 > 1$ and $T > 1$ for which

$$c_1 t^{n-\frac{r}{4}} \leq \varphi(x,t) < t^{m(x)+r} \tag{4.29}$$

uniformly in Ω, for all $t \geq T$. Consider a partition of unity (χ_1, χ_2) subordinated to the cover of Ω that consists of the open sets

$$\Omega_1 = p^{-1}\left((n-r, n+r)\cap I\right) \ , \ \Omega_2 = p^{-1}\left(\left(n+\frac{r}{2}, \infty\right)\cap I\right).$$

If $(u_j)_j$ is a bounded sequence in $W_0^{1,\varphi}(\Omega)$ (hence in $W_0^{1,n-\frac{r}{4}}(\Omega)$) one can set

$$q = \left(n - \frac{r}{4}\right) I_{\Omega_1} + m_- I_{\Omega\setminus\Omega_1}$$

and along the same lines as in (4.24)-(4.25), conclude that $(u_j\chi_1)_j$ is bounded in $W_0^{1,n-\frac{r}{4}}(\Omega)$. Via Theorem 4.1 and on account of the choice (4.20) it follows that $(u_j\chi_1)_j$ has a subsequence that converges in $L^{n+2r}(\Omega)$. If

$$t := (n+2r) I_{\Omega\cap\Omega_1} + (m_+ + 2r) I_{\Omega\setminus\Omega\cap\Omega_1},$$

then the obvious equality

$$\int_\Omega \left|(u_i - u_j)\chi_1\right|^{n+2r} = \int_\Omega \left|(u_i - u_j)\chi_1\right|^t$$

implies that the subsequence also converges in $L^t(\Omega)$ and since $m + r < t$ in Ω, it converges also in $L^{m+r}(\Omega)$, hence in $L^\varphi(\Omega)$ via the right-hand inequality in (4.29). Still denoting this subsequence by $(u_j\chi_1)_j$, it is easy to see, that $(u_j\chi_2)_j$ is bounded in $W_0^{1,n+\frac{r}{2}}(\Omega)$; therefore from Theorem 4.1 it is clear that it has a subsequence (still denoted by $(u_j\chi_2)_j$) that converges in $L^{m+2r}(\Omega)$. The right-hand inequality in (4.22) yields the convergence of $(u_j\chi_2)_j$ in $L^\varphi(\Omega)$.
A straightforward computation reveals that the above conclusion implies the compactness of the embedding (4.19) in all cases. $\square$

It is apparent from the proof of the preceding Theorem that functions in $W_0^{1,\varphi}(\Omega)$ belong to a higher order integrability space than just $L^\varphi(\Omega)$. We state this important fact as a separate corollary:

Corollary 4.2. *For a Musielak-Orlicz function φ on Ω that fulfills the conditions of Theorem 4.3, the embedding*

$$W_0^{1,\varphi}(\Omega) \hookrightarrow L^{m(x)+\frac{\varepsilon}{20}}(\Omega) \subsetneqq L^\varphi(\Omega)$$

is compact.

The following Corollary generalizes the Poincaré's inequality to the setting of Musielak-Orlicz spaces.

Corollary 4.3 (Poincaré's inequality). *For φ satisfying the conditions of Theorem 4.3, there exists a positive constant C depending only on n, Ω, φ, such that for any $u \in W_0^{1,\varphi}(\Omega)$*

$$\|u\|_\varphi \leq C \||\nabla u|\|_\varphi.$$

Proof. If not, it would be an elementary matter to construct a sequence

$$(v_k) \subset W_0^{1,\varphi}(\Omega)$$

with

$$\|v_k\|_{1,\varphi} = 1 \geq \|v_k\|_\varphi \geq k\||\nabla v_k|\|_\varphi \text{ for } k \in \mathbb{N}.$$

Clearly,

$$|\nabla v_k| \longrightarrow 0 \text{ in } L^\varphi(\Omega) \tag{4.30}$$

as $k \longrightarrow \infty$ and the compactness of the Sobolev embedding yields the existence of $v \in L^\varphi(\Omega)$ for which

$$v_k \longrightarrow v \text{ in } L^\varphi(\Omega).$$

Necessarily then,

$$\|v_k - v_j\|_{1,\varphi} = \|v_k - v_j\|_\varphi + \|\nabla(v_k - v_j)\|_\varphi \longrightarrow 0 \text{ as } k, j \longrightarrow \infty;$$

it follows that $(v_k)_k$ converges in $W_0^{1,\varphi}(\Omega)$ and it is obvious that the limit must be v. On the other hand, (4.30) forces $\nabla v = 0$ and hence $v = 0$, which is a contradiction.

□

4.6 Applications

In this Section some applications of the preceding compactness results are studied. Throughout this Section, $\Omega \subset \mathbb{R}^n$ denotes a bounded domain and φ stands for a Musielak-Orlicz function on Ω that satisfies all the conditions of Theorem 4.3

We set

$$B_r := \left\{ u \in W_0^{1,\varphi}(\Omega) : \rho_\varphi(|\nabla u|) \leq r \right\}$$

Theorem 4.4. *Let φ be an Musielak-Orlicz function on Ω; assume that φ satisfies the conditions of Theorem 4.3; in particular, φ satisfies the Δ_2-condition, i.e., for some $K > 0$, $S_0 > 0$ it holds that*

$$\varphi(x, 2s) \leq K\varphi(x, s) \textit{ for all } s \geq S_0, x \in \Omega. \tag{4.31}$$

Then,

$$\sup\left\{\rho_\varphi(u) : \rho_\varphi(|\nabla u|) \leq r\right\} < \infty. \tag{4.32}$$

The following Lemma is vitally important in what follows.

Lemma 4.3. *For $r > 0$, the modular ball*

$$B_r := \left\{ u \in W_0^{1,\varphi}(\Omega) : \rho_\varphi(|\nabla u|) \leq r \right\}$$

is weakly closed.

Proof. See [14]. □

Theorem 4.5. *Let $r > 0$. Then there exists a function $u_r \in W_0^{1,\varphi}(\Omega)$ that is maximal in the following sense:*

$$\rho_\varphi(u_r) = \mathbf{S}_r = \sup \left\{ \rho_\varphi(u) : \rho_\varphi(|\nabla u|) \leq r \right\}$$

Proof. For each $n \in \mathbb{N}$, select a function u_n u_n in B_r with

$$\mathbf{S}_r - \frac{1}{n} < \rho_\varphi(u_n).$$

The sequence (u_n) is bounded in $W_0^{1,\varphi}(\Omega)$; indeed, Theorem 4.4 guarantees that the numerical sequence $(\rho_\varphi(u_n))$ is bounded and it follows from here that either $\|u_n\|_\varphi \leq 1$ or

$$1 = \rho_\varphi\left(u_n/\|u_n\|_\varphi\right) \leq \|u_n\|_\varphi^{-1} \rho_\varphi(u_n), \tag{4.33}$$

so that the boundedness of the sequence $(\|u_n\|_\varphi)$ follows from that of the sequence $(\rho_\varphi(u_n))$ (4.32). Likewise, since by definition, $(\rho_\varphi(|\nabla u_n|))$ is bounded, either $\||\nabla u_n|\|_\varphi \leq 1$ or

$$1 = \rho_\varphi\left(|\nabla u_n|/\||\nabla u_n|\|_\varphi\right) \leq \||\nabla u_n|\|_\varphi^{-1} \rho_\varphi(|\nabla u_n|). \tag{4.34}$$

Inequalities (4.33) and (4.34) together with the discussions preceding them show the claimed boundedness of (u_n).

On account of the theorem of Banach-Alaoglu and of the reflexivity of $W_0^{1,\varphi}(\Omega)$, no generality is lost by assuming that (u_n) converges weakly in $W_0^{1,\varphi}(\Omega)$; let

$$u_n \overset{W_0^{1,\varphi}(\Omega)}{\rightharpoonup} u \in W_0^{1,\varphi}(\Omega). \tag{4.35}$$

In particular, in the light of Theorem 4.3, statement (4.35) ensures that $u \in B_r$.

By virtue of the compactness of the Sobolev embedding (Theorem 4.3) one has the strong convergence

$$u_n \xrightarrow{L^\varphi(\Omega)} u$$

One can hence assume that $u_n \longrightarrow u$ *a.e.* in Ω. In fact, u is the sought-for maximal function. To see this, we notice that *a.e.* in Ω,

$$\varphi(x,|u_n(x)|) \longrightarrow \varphi(x,|u(x)|)$$

and that on account of convexity, for any $n \in \mathbb{N}$:

$$\varphi(x,|u_n(x)|) \leq \frac{1}{2}\varphi(x,2|u_n(x)-u(x)|)+\frac{1}{2}\varphi(x,2|u(x)|). \tag{4.36}$$

Select n large enough so that $2\|u-u_n\|_\varphi < 1$; for such n (4.36) yields:

$$\varphi(x,|u_n(x)|) \leq \|u_n-u\|_\varphi \varphi\left(x, \frac{|u_n(x)-u(x)|}{\|u-u_n\|_\varphi}\right)+\frac{1}{2}\varphi(x,2|u(x)|). \tag{4.37}$$

Denote the left-hand side and the right-hand side of (4.37) by v_n and w_n respectively. Then the following conditions hold:

(i) $v_n(x) \to v(x) = \varphi(x,|u(x)|) \in L^1(\Omega)$ *a.e.* in Ω

(ii) $w_n(x) \to w(x) = \frac{1}{2}\varphi(x,2|u(x)|) \in L^1(\Omega)$ *a.e.* in Ω

(iii) $v_n, w_n \in L^1(\Omega)$ for any $n \in \mathbb{N}$

(iv) $\int\limits_\Omega w_n dx \to \int\limits_\Omega \frac{1}{2}\varphi(x,2|u|)dx = \frac{1}{2}\rho_\varphi(2u)$.

Since $w-v \geq 0$ *a.e* in Ω, Fatou's Lemma leads to:

$$\begin{aligned}\int\limits_\Omega (w-v)dx &\leq \int\limits_\Omega w\,dx + \liminf_n \int\limits_\Omega (-v_n)\,dx \\ &= \int\limits_\Omega w\,dx - \limsup_n \int\limits_\Omega v_n\,dx\end{aligned}$$

and

$$\int\limits_\Omega (w+v)dx \leq \int\limits_\Omega w\,dx + \liminf_n \int\limits_\Omega v_n\,dx.$$

The two last statements yield

$$\lim_{n\to\infty}\int\limits_\Omega \varphi(x,|u_n(x)|)dx = \int\limits_\Omega \varphi(x,|u(x)|)dx$$

or, equivalently

$$\rho_\varphi(u_n) \longrightarrow \rho_\varphi(u) \text{ as } n \to \infty. \tag{4.38}$$

By construction $\rho_\varphi(u_n) \longrightarrow \mathbf{S}_r$; (4.38) is therefore the desired result. □

Theorem 4.6. *Let*

$$\varphi : \Omega \times [0,\infty) \longrightarrow [0,\infty)$$

be a Musielak-Orlicz function on a bounded domain $\Omega \subset \mathbb{R}^n$. Assume that φ satisfies the conditions of Theorem 4.3, that φ is an N-function and that

$$\frac{\partial \varphi}{\partial t}(x,t) > 0 \text{ for } t > 0 \text{ and a.e. } x \in \Omega. \tag{4.39}$$

Then there exists at least one (weak) solution (λ_0, u_0) to the problem

$$div \left(|\nabla u(\cdot)|^{-1} \nabla u \frac{\partial}{\partial t} \varphi(\cdot, |\nabla u(\cdot)|) \right) = \lambda |u(\cdot)|^{-1} u \frac{\partial}{\partial t} \varphi(\cdot, |u(\cdot)|).$$

To facilitate the proof of Theorem 4.6 we digress on Lagrange-multipliers. Let F be the functional defined by

$$F : L^{\varphi}(\Omega) \to [0,\infty) \ F(v) = \rho_{\varphi(v)} \tag{4.40}$$

and set $G(v) = F(|\nabla v|)$ for $v \in W_0^{1,\varphi}(\Omega)$.

For any $r > 0$, we recall that Theorem 4.5 yields the existence of at least $u_r \in W_0^{1,\varphi}(\Omega)$ such that

$$F(u_r) = \rho_{\varphi}(u_r) = \mathbf{S}_r = \sup \left\{ \rho_{\varphi}(u) : \rho_{\varphi}(|\nabla u|) = G(u) \le r \right\}.$$

Lemma 4.4. *It is easy to conclude (see [14]) that*

$$\langle G'(u_r), u_r \rangle > 0. \tag{4.41}$$

Proof. It is clear that

$$\langle G'(u_r), u_r \rangle = \int_{\Omega} \frac{\partial \varphi}{\partial s}(x, |\nabla u_r(x)|) |\nabla u_r(x)| dx.$$

On the other hand, $G(u_r) = \int_{\Omega} \varphi(x, |\nabla u_r(x)|) dx = r > 0$. In all, one must have $\varphi(x, |\nabla u_r(x)|) > 0$ *a.e.* $x \in \Omega$. Thus, $|\nabla u_r(x)| > 0$ *a.e.* $\in \Omega$. The claim now follows on account of assumption (4.39). □

Lemma 4.5. *Under the assumptions of Lemma 4.4, it follows that*

$$W_0^{1,\varphi}(\Omega) = \ker G'(u_r) \oplus \langle \{u_r\} \rangle. \tag{4.42}$$

Proof. The proof is elementary: Any $v \in W_0^{1,\varphi}(\Omega)$ can be written as

$$v = v - \left(\langle G'(u_r), v \rangle / \langle G'(u_r), u_r \rangle \right) u_r + \left(\langle G'(u_r), v \rangle / \langle G'(u_r), u_r \rangle \right) u_r$$

and a straightforward calculation shows this decomposition to be unique. □

Next, set

$$\omega : \ker G'(u_r) \oplus \mathbb{R} \longrightarrow [0, \infty)$$
$$\omega(h,t) = G((1+t)u_r + h) - r.$$

By definition, it is immediate that $w(0,0) = 0$.
We claim that w is differentiable in both variables, that $\frac{\partial w}{\partial t}(0,0) > 0$ and that $\frac{\partial w}{\partial h}(0,0) = 0$. The differentiability of w is clear from the differentiability of G; it follows that

$$\frac{\partial w}{\partial t}(0,0) = \langle G'(u_r), u_r \rangle > 0.$$

The last assertion follows by direct computation, namely, for any $h \in \ker G'$ (i.e., $\langle G'(u_r), h \rangle = 0$) one has:

$$\begin{aligned} \frac{w(h,0) - w(0,0)}{\|h\|_{1,\varphi}} &= w(h,0)/\|h\|_{1,\varphi} = \frac{G(u_r + h) - G(u_r)}{\|h\|_{1,\varphi}} \\ &= \frac{G(u_r + h) - G(u_r) - \langle G'(u_r), h \rangle}{\|h\|_{1,\varphi}} \\ &\longrightarrow 0 \text{ as } \|h\|_{1,\varphi} \to 0. \end{aligned}$$

Lemma 4.6. *For w as above, the function*

$$\frac{\partial w}{\partial h} : \ker G'(u_r) \oplus \mathbb{R} \longrightarrow \left(\ker G'(u_r)\right)^*$$

is continuous.

Proof. By definition, one has, for $\eta \in \ker G'(u_r)$:

$$\begin{aligned} \frac{\partial w}{\partial h}(h_0, t_0)(\eta) &= G^{'}(h_0 + (1+t_0)u_r)(\eta) \\ &\int_\Omega \frac{\partial \varphi}{\partial s}(x, |\nabla u_0(x)|) \frac{\nabla u_0(x) \nabla \eta(x)}{|\nabla u_0(x)|} dx. \end{aligned}$$

The continuity claim follows immediately from Theorem 2.7.2 in [14], and through repeated applications of Hölder's inequality and Lebesgue's dominated convergence theorem. □

Therefore, the implicit function theorem applies to w. There exist thus a neighborhood of zero U in $W_0^{1,\varphi}(\Omega)$, $\varepsilon > 0$ and a differentiable function

$$\psi : U \cap \ker G'(u_r) \longrightarrow (-\varepsilon, \varepsilon)$$

such that for any $(a,b) \in U \cap \ker G'(u_r) \times (-\varepsilon, \varepsilon)$ one has

$$w(a,b) = 0 \iff b = \psi(a).$$

Select $h \in \ker G'$, let $\delta > 0$ be so small that $th \in U \cap \ker G'(u_r)$ for $|t| < \delta$ and put

$$\begin{aligned} &\mathbf{b} : (-\delta, \delta) \longrightarrow W_0^{1,\varphi}(\Omega) \\ &\mathbf{b}(t) = u_r + th + \psi(th). \end{aligned}$$

Then $\mathbf{b}(0) = u_r$, $\mathbf{b}$ is differentiable at 0 (since so is ψ) and $\mathbf{b}'(0) = h$, since $\psi'(0) = 0$. Then the function

$$\begin{aligned} &\xi : (-\delta, \delta) \longrightarrow [0.\infty) \\ &\xi(t) = F(u_r + th + \psi(th)) \end{aligned}$$

is differentiable and attains a maximum at $t = 0$. It follows then that

$$\xi'(0) = F'(u_r)(h) = 0$$

and from here, by assumption on h, that

$$\ker G' \subseteq \ker F'.$$

Since both, $\ker G'$ and $\ker F'$ have codimension one and neither functional is identically zero, it is concluded that

$$\ker G'(u_r) \subseteq \ker F'(u_r),$$

and thus that there must exist a constant $\lambda \in \mathbb{R}$ such that

$$G'(u_r) = \lambda F'(u_r). \tag{4.43}$$

4.6.1 The eigenvalue problem for the p-Laplacian

Fix a bounded domain $\Omega \subset \mathbb{R}^n$ and a function $p \in C(\overline{\Omega})$ with $1 < p_- \leq p_+ < \infty$ on Ω. It is well known [35] that $W_0^{1,p}(\Omega)$ is reflexive, whereas Theorem 4.3 asserts that the Sobolev embedding

$$E : W_0^{1,p(\cdot)}(\Omega) \hookrightarrow L^{p(\cdot)}(\Omega) \tag{4.44}$$

is in fact, compact. Theorem 2.7.8 in [14] gives an expression for the Fréchet derivative of the norm $\|\cdot\|_p$. Elementary considerations show that the action of the Fréchet derivative of the norm

$$W_0^{1,p}(\Omega) \ni u \to \|u\|_p^{(1)} = \||\nabla u|\|_p$$

on $h \in W_0^{1,p}(\Omega)$ is given by the following:

$$h \longrightarrow \int_\Omega \frac{p\|\nabla u\|_p^{-p}\|\nabla u\|^{p-1}\nabla u \nabla h}{\int_\Omega p\|\nabla u\|_p^{-p-1}\|\nabla u\|^p dx} dz. \tag{4.45}$$

Remarkably, when p is constant on Ω, the operator given by (4.45) takes up the form

$$-\||\nabla u|\|_p^{1-p}\mathrm{div}\left(|\nabla u|^{p-2}\nabla u\right) = -\||\nabla u|\|_p^{1-p}\Delta_p(u),$$

where $\Delta_p(u) := \mathrm{div}\left(|\nabla u|^{p-2}\nabla u\right)$ is the much-studied p-Laplacian operator, which emerged in the treatment of some applications to fluid mechanics.

On the other hand, any extremal function u_0 for the Sobolev embedding (4.44) is a solution of the eigenvalue problem

$$(\mathrm{grad}\,\|u\|_p) = \lambda(\mathrm{grad}\,\|u\|_p^{(1)}), \tag{4.46}$$

with $\lambda = \|u_0\|_p / \||\nabla u_0|\|_p$ ([3]).

It is shown in [3] that if $(\lambda, u) \in (0,\infty) \times W_0^{1,p}(\Omega)$ is a solution of Problem (4.46), then necessarily $\lambda \leq \|E\|$.

Problem (4.46) becomes

$$-\lambda\||\nabla u|\|_p^{1-p}\mathrm{div}\left(|\nabla u|^{p-1}\nabla u\right) = \|u\|_p^{1-p}|u|^{p-2}u,$$

which can be equivalently written as

$$-\mathrm{div}\left(|\nabla u|^{p-1}\nabla u\right) = \lambda^{-1}\left(\||\nabla u|\|_p / \|u\|\right)_p^{p-1}|u|^{p-2}u. \tag{4.47}$$

Any $\lambda \in \mathbb{R}$ satisfying equality (4.47) for some $u \in W_0^{1,p}(\Omega)$ satisfies the inequality

$$\lambda^{-1} = \||\nabla u\|_p / \|u\|_p \geq \|E\|^{-1};$$

moreover, if $(\lambda, u) \in (0,\infty) \times W_0^{1,p}(\Omega)$ solves Problem (4.47), then the pair

$$(\lambda^{-1}\left(\||\nabla u|\|_p / \|u\|\right)_p^{p-1}, u)$$

is a solution of the eigenvalue problem for the p-Laplacian, namely

$$-\Delta_p u = \gamma|u|^{p-2}u \tag{4.48}$$

and conversely if (γ, v) is any solution of (4.48), then

$$(\gamma^{-1}\,(\|v\|_p/\||\nabla v|\|_p)^{p-1}\,, v)$$

is a solution of (4.47). It is apparent from the preceding argument that under the assumption of constant p, the smallest eigenvalue of the p-Laplacian p-Laplacian is $\|E\|^p$, the p-th power of the norm of the Sobolev embedding; also its corresponding first eigenfunction (which was shown in [2] to be unique up to multiplication times constants, see also [7] for a simpler proof) is extremal for the Sobolev embedding E. The situation changes radically when p is non-constant on Ω. In this case, the tempting natural generalization that results from replacing p with a function in problem (4.48), while an eigenvalue problem, is not related to problem (4.46) in any useful way, and both problems, (4.46) and (4.48) must be studied separately (see [10] and the references therein).

Corollary 4.4. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain and $p \in C(\overline{\Omega})$. Then there exists at least a solution $(\lambda, u) \in (0,\infty) \times W_0^{1,p}(\Omega)$ of the modular eigenvalue problem*

$$-\Delta_p u = \lambda |u|^{p-2} u. \tag{4.49}$$

Proof. By virtue of Theorem 4.6, for each $r > 0$ there exists an eigenfunction u_r satisfying the exremality condition of Theorem 4.5 and a corresponding eigenvalue λ_r for which (4.49) holds. $\square$

Bibliography

[1] R. Adams, *Sobolev Spaces*, Academic Press, 1975.

[2] A. Anane, *Simplicité et isolation de la premiére valeur propre du p-Laplacien avec poids*, C.R. Acad. Sci. Paris, t. 305, Série I, (1987), 725-728.

[3] C. Bennewitz, *Approximation numbers=singular values,* J. Comp. Appl. Math. **208**, (1) (2007), 102-110.

[4] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations,* Universitext, Springer, 2011.

[5] L. Diening, P. Harjulehto, P. Hästö, M. Rûžiĉka, *Lebesgue and Sobolev Spaces with Variable Exponents*, Lecture Note in Mathematics 2017, Springer, Berlin 2011.

[6] D. Gilbarg, N. Trudinger, *Elliptic Partial Differential Equations of Second Order,* Springer, Berlin-Heidelberg, New York, Tokyo, 1983.

[7] B. Kawohl, P. Lindqvist, *Positive eigenfunctions for the p-Laplace operator revisited,* Analysis **26** No 4, (2006), 545-550.

[8] O. Kováčik, J. Rákosník, *On spaces* $L^{p(x)}$ *and* $W^{1,p(x)}$, Czechoslovak Math. J., **41** (1991), 592-618.

[9] J. Lang, O. Méndez, *Modular eigenvalues of the Dirichlet* $p(\cdot)$*-Laplacian and their stability,* Spectral theory, function spaces and inequalities, 125-137, Oper. Theory Adv. Appl., 219, Birkhäuser/Springer Basel AG, Basel, 2012.

[10] J. Lang, O. Méndez, *Extension of a result by Lindquist to Lebesgue spaces with variable exponents,* J. Differential Equations **259**, no. 2, (2015), 562-595.

[11] J. Lukeš, L. Pick, D. Pokorný, *On geometric properties of the spaces* $L^{p(x)}(\Omega)$, Rev. Math. Complut.**24** (2011), 115-130.

[12] K. Mania, *Sopra un esempio di Lavrentieff*, Boll. Un. Mat. Ifal. **13** (1934) 14-153.

[13] W. Matuszewska, W. Orlicz, *On certain properties of φ-functions*, Bull. Acad. Polon. Sci., Ser. Sci. Math. Astronom. Phys. **8** (1960) 430-443.

[14] O. Méndez, J. Lang, *Analysis on Function Spaces of Musielak Orlicz Type*, CRC Press, Taylor and Francis group, 2018.

[15] J. Musielak, *Orlicz spaces and modular spaces*, Lecture Notes in Mathematics, Vol.1034, Springer-Verlag, Berlin/Heidelberg/New York/Tokyo, 1983.

[16] H. Nakano, *Modulared Semi-ordered Linear Spaces*, Maruzen Co., Tokyo, 1950

[17] H. Nakano, *Topology of linear topological spaces*, Maruzen Co. Ltd., Tokyo, 1951.

[18] H. Nakano, *Modulared sequence spaces*, Proc. Japan Acad., **27**(1951), 508-512.

[19] W. Orlicz, *Über konjugierte Exponentenfolgen*, Studia Math., **3**(1931), 200-211.

[20] W. Orlicz, *Über eine gewisse Klasse von Räumen vom Typus B*, Bull. Int. Acad. Polon. Sci. **A**, **8/9**(1932), 207-220.

[21] K. Rajagopal, M. Ružička, *On the modeling of electrorheological materials*, Mech. Research Comm., **23**(1996), 401-407.

[22] Riesz, F., *Über Systeme integrierbarer Funktionen,* Math. Ann. **69** (1910), 449-497.

[23] M. Ružička, *Electrorheological fluids: modeling and mathematical theory*, Lecture Notes in Mathematics 1748, Springer-Verlag, Berlin 2000.

[24] L. Sharapudinov, *The topology of the space $L^{p(t)}([0,1])$,* (Russian) Mat. Zametki **26:4** (1979), 613-632; Math. Notes, 26:4 (1979), 796-806.

[25] W. Ziemer, *Weakly Differentiable Functions,* Springer-Verlag, New York, 1989.

Chapter 5

The wave equation with non-standard linearities

S. A. Messaoudi
Department of Mathematics, University of Sharjah, Sharjah, UAE,
E-mail: smessaoudi@sharjah.ac.ae

A. A. Talahmeh
Department of Mathematics, Birzeit University, West Bank, Palestine,
E-mail: atalahmeh@birzeit.edu

In this Chapter, a brief overview of some results related to the variable-exponent Lebesgue and Sobolev spaces is presented. This is followed by a brief discussion of some important and recent results related to the nonexistence and blow-up for wave equations with non-standard nonlinearities (nonlinearities involving variable exponents), as well as some decay and stability results for classical nonlinear wave equations. Finally we present an exponential decay result for a strongly damped wave equation with a nonstandard damping.

5.1 Definitions and preliminaries

5.1.1 The variable-exponent Lebesgue spaces

In this Subsection some preliminary facts about the Lebesgue spaces with variable exponents are summarized.

Definition 5.1. Let X be a $\mathbb{K}-$vector space. A function $\rho : X \longrightarrow [0,\infty]$ is said to be *left-continuous* if the mapping $\lambda \longmapsto \rho(\lambda x)$ is left-continuous on $[0,\infty)$, for every $x \in X$; that is,

$$\lim_{\lambda \to 1^-} \rho(\lambda x) = \rho(x), \quad \forall\, x \in X.$$

Definition 5.2. Let X be a $\mathbb{K}-$vector space. A function $\rho : X \longrightarrow [0,\infty]$ is called a *semimodular* on X if the following properties hold:

(a) $\rho(0) = 0$

(b) $\rho(\lambda x) = \rho(x)$, for all $x \in X$ and $\lambda \in \mathbb{K}$, with $|\lambda| = 1$

(c) ρ is convex

(d) ρ is left-continuous

(e) $\rho(\lambda x) = 0$, for all $\lambda > 0$ implies $x = 0$

A semimodular is called *modular* if

(f) $\rho(x) = 0$ implies $x = 0$

A semimodular is called *continuous* if

(g) the mapping $\lambda \longmapsto \rho(\lambda x)$ is continuous on $[0,\infty)$ for all $x \in X$

Example 5.1. Let $L^0(\Omega)$ be the set of all Lebesgue-measurable functions defined on Ω. If $1 \le p < +\infty$, then

$$\rho_p(f) := \int_\Omega |f(x)|^p \mathrm{dx}$$

defines a continuous modular on $L^0(\Omega)$.

Theorem 5.1. *[25] Let ρ be a semimodular on X. Then, the mapping $\lambda \to \rho(\lambda x)$ is non-decreasing on $[0,\infty)$ for every $x \in X$. Moreover*

$$\begin{aligned} \rho(\lambda x) &= \rho(|\lambda| x) \leq |\lambda| \rho(x) \quad \text{for all } |\lambda| \leq 1, \\ \rho(\lambda x) &= \rho(|\lambda| x) \geq |\lambda| \rho(x) \quad \text{for all } |\lambda| \geq 1. \end{aligned} \tag{5.1}$$

Definition 5.3. [25] Let (A, Σ, μ) be a $\sigma-$finite, complete measure space. Let $\mathscr{P}(A,\mu)$ be the set of all $\mu-$measurable functions $p : \Omega \to [1,\infty]$. The function $p \in \mathscr{P}(A,\mu)$ is called a variable exponent on A. Set

$$p_1 := \operatorname{essinf}_{y\in A} p(y) \ \text{ and } \ p_2 := \operatorname{esssup}_{y\in A} p(y).$$

If $p_2 < +\infty$, then p is said to be a bounded variable exponent. If $p \in \mathscr{P}(A,\mu)$, then we define $p^\star \in \mathscr{P}(A,\mu)$ by

$$\frac{1}{p(y)} + \frac{1}{p^\star(y)} = 1, \ \text{ where } \ \frac{1}{\infty} := 0.$$

The function $p^\star$ is called the dual variable exponent of p. In the special case when μ is the $n-$ dimensional Lebesgue measure and Ω is an open subset of $\mathbb{R}^n$, denote $\mathscr{P}(\Omega) := \mathscr{P}(\Omega,\mu)$.

Definition 5.4. [25] We define the Lebesgue space with a variable exponent $p(\cdot)$ by

$$L^{p(\cdot)}(\Omega) := \Big\{ u : \Omega \to \mathbb{R};\ \text{measurable in}\, \Omega :\ \rho_{p(\cdot)}(\lambda u) < \infty, \text{ for some}\, \lambda > 0 \Big\},$$

where

$$\rho_{p(\cdot)}(u) = \int_\Omega |u(x)|^{p(x)} dx$$

is easily seen to be a modular. $L^{p(\cdot)}(\Omega)$ is endowed with the following Luxemburg-type norm

$$\|u\|_{p(\cdot)} := \inf \Big\{ \lambda > 0 : \int_\Omega \Big| \frac{u(x)}{\lambda} \Big|^{p(x)} dx \leq 1 \Big\}.$$

Example 5.2. Let $p(x) = x$ on $\Omega = (1,2)$. Then $\|1\|_{p(\cdot)} = 1$. Indeed,

$$\rho_{p(\cdot)}(1/\lambda) = \int_1^2 \lambda^{-x} dx = \frac{\lambda - 1}{\lambda^2 \ln \lambda}.$$

Since $\rho_{p(\cdot)}(1) = 1$, then, by definition of $\|1\|_{p(\cdot)}$, we have $\|1\|_{p(\cdot)} \leq 1$. On the other hand, it is easy to check that $\rho_{p(\cdot)}(1/\lambda) > 1$, for $0 < \lambda < 1$. This gives $\|1\|_{p(\cdot)} \geq 1$. Hence, we conclude that $\|1\|_{p(\cdot)} = 1$.

Lemma 5.1. *If $p(x) \equiv p$, where p is constant. Then,*

$$\|u\|_{p(\cdot)} = \lambda_0 = \Big(\int_\Omega |u|^p \Big)^{\frac{1}{p}}. \tag{5.2}$$

Proof. Since $\rho_{p(\cdot)}(u/\lambda_0) = 1$, then

$$\|u\|_{p(\cdot)} \leq \lambda_0. \tag{5.3}$$

Next, using the property of the infimum, it is easy to see that there exists a sequence $\{\lambda_k\}_{k=1}^{\infty}$ such that $\lambda_k \geq \|u\|_{p(\cdot)}$, with

$$\rho_{p(\cdot)}(u/\lambda_k) \leq 1 \text{ and } \lambda_k \to \|u\|_{p(\cdot)}.$$

Since, $\rho_{p(\cdot)}(u/\lambda_k) = \frac{1}{(\lambda_k)^p} \int_\Omega |u|^p \leq 1$, it follows

$$\lambda_0 \leq \|u\|_{p(\cdot)}. \tag{5.4}$$

(5.2) follows by combining (5.3) and (5.4). $\square$

Definition 5.5. [25] We say that a function $q : \Omega \to \mathbb{R}$ is log-Hölder continuous on Ω, if there exist $A > 0$ and $0 < \delta < 1$ such that

$$|q(x) - q(y)| \leq -\frac{A}{\log|x-y|}, \text{ for all } x, y \in \Omega, \text{ with } |x-y| < \delta. \tag{5.5}$$

Lemma 5.2. *Let Ω be a domain of $\mathbb{R}^n$. If $p : \Omega \to \mathbb{R}$ is a Lipchitz function, then it is log-Hölder continuous on Ω.*

Proof. Let $x, y \in \Omega$, with $|x-y| < \delta$ and $0 < \delta < 1$. Then, since p is Lipschitz, there exists $L > 0$ such that

$$\begin{aligned} |p(x) - p(y)| &\leq L|x-y| \\ &\leq -\frac{L}{\log|x-y|}\big(-|x-y|\log|x-y|\big). \end{aligned} \tag{5.6}$$

It is easy to check that $g(s) = -s\log s$ is continuous on $[0,1]$ and hence is bounded. So we have $0 \leq -s\log s \leq M$. Therefore, (5.6) becomes

$$|p(x) - p(y)| \leq -\frac{A}{\log|x-y|}, \tag{5.7}$$

where $A = LM > 0$. Hence, p is log-Hölder continuous. $\square$

Example 5.3. Let $p(x,y) = x^2 + 1$ be defined on the unit ball $\Omega = B(0,1)$. Then, by the previous lemma $p : \Omega \to \mathbb{R}$ is log-Hölder continuous on Ω.

Lemma 5.3. *[25][**Unit ball property**] Let $p \in \mathscr{P}(A,\mu)$ and $f \in L^{p(\cdot)}(A,\mu)$. Then*

(i) $\|f\|_{p(\cdot)} \leq 1$ *if and only if* $\rho_{p(\cdot)}(f) \leq 1$

(ii) *If* $\|f\|_{p(\cdot)} \leq 1$, *then* $\rho_{p(\cdot)}(f) \leq \|f\|_{p(\cdot)}$

(iii) *If* $\|f\|_{p(\cdot)} \geq 1$, *then* $\|f\|_{p(\cdot)} \leq \rho_{p(\cdot)}(f)$

(iv) $\|f\|_{p(\cdot)} \leq 1+\rho_{p(\cdot)}(f)$

The following results from [25] are mentioned without proof.

Lemma 5.4. *If* $1 < p_1 \leq p(x) \leq p_2 < +\infty$ *holds, then*

$$\min\left\{\|u\|_{p(\cdot)}^{p_1}, \|u\|_{p(\cdot)}^{p_2}\right\} \leq \rho_{p(\cdot)}(u) \leq \max\left\{\|u\|_{p(\cdot)}^{p_1}, \|u\|_{p(\cdot)}^{p_2}\right\},$$

for any $u \in L^{p(\cdot)}(\Omega)$.

Theorem 5.2. *If* $p \in \mathscr{P}(A,\mu)$, *then* $L^{p(\cdot)}(A,\mu)$ *is a Banach space.*

Lemma 5.5. *If* $p : \Omega \to [1,\infty)$ *is a measurable function with* $p_2 < +\infty$, *then* $C_0^\infty(\Omega)$ *is dense in* $L^{p(\cdot)}(\Omega)$.

Lemma 5.6 (Young's Inequality). *Let* $p,q,s \in \mathscr{P}(\Omega)$ *such that*

$$\frac{1}{s(y)} = \frac{1}{p(y)} + \frac{1}{q(y)}, \text{ for a.e } y \in \Omega.$$

Then for all $a,b \geq 0$,

$$\frac{(ab)^{s(\cdot)}}{s(\cdot)} \quad \leq \quad \frac{(a)^{p(\cdot)}}{p(\cdot)} + \frac{(b)^{q(\cdot)}}{q(\cdot)}. \tag{5.8}$$

By taking $s = 1$, *and* $1 < p,q < +\infty$, *it follows that for any* $\varepsilon > 0$,

$$ab \leq \varepsilon a^p + C_\varepsilon b^q, \quad \forall a,\, b \geq 0,$$

where $C_\varepsilon = \dfrac{1}{q(\varepsilon p)^{\frac{q}{p}}}$. *For* $p = q = 2$, ,

$$ab \leq \varepsilon a^2 + \frac{b^2}{4\varepsilon}.$$

Lemma 5.7 (Hölder's Inequality). *Let* $p,q,s \in \mathscr{P}(\Omega)$ *such that*

$$\frac{1}{s(y)} = \frac{1}{p(y)} + \frac{1}{q(y)}, \text{ for a.e } y \in \Omega.$$

If $f \in L^{p(\cdot)}(\Omega)$ *and* $g \in L^{q(\cdot)}(\Omega)$, *then* $fg \in L^{s(\cdot)}(\Omega)$ *and*

$$\|fg\|_{s(\cdot)} \leq \; 2\, \|f\|_{p(\cdot)} \|g\|_{q(\cdot)}.$$

Taking $p = q = 2$, *yields the **Cauchy-Schwarz inequality**.*

5.1.2 The variable-exponent Sobolev spaces

In this Subsection some functional analysis-type properties of Sobolev spaces with variable exponents are studied. Recall the definition of weak derivative.

Definition 5.6. ***(weak derivative).*** *Let $\Omega \subset \mathbb{R}^n$ be a domain. Assume that $u \in L^1_{loc}(\Omega)$. Let $\alpha := (\alpha_1, ..., \alpha_n) \in \mathbb{N}^n$ be a multi-index and let $|\alpha| = \alpha_1 + ... + \alpha_n$. If there exists $g \in L^1_{loc}(\Omega)$ such that*

$$\int_\Omega u \frac{\partial^{|\alpha|}\psi}{\partial^{\alpha_1}x_1 .....\partial^{\alpha_n}x_n} dx = (-1)^{|\alpha|} \int_\Omega \psi g \, dx,$$

for all $\psi \in C_0^\infty(\Omega)$, then g is called a weak partial derivative *of u of order α. The function g is denoted by $\partial_\alpha u$ or $\frac{\partial^{|\alpha|}u}{\partial^{\alpha_1}x_1 .....\partial^{\alpha_n}x_n}$.*

Definition 5.7. *Let $k \in \mathbb{N}$. The Sobolev space $W^{k,p(\cdot)}(\Omega)$ is defined as*

$$W^{k,p(\cdot)}(\Omega) := \Big\{ u \in L^{p(\cdot)}(\Omega) \text{ such that } \partial_\alpha u \in L^{p(\cdot)}(\Omega), \ \forall \, |\alpha| \le k \Big\}.$$

A semimodular on $W^{k,p(\cdot)}(\Omega)$ is defined by

$$\rho_{W^{k,p(\cdot)}(\Omega)}(u) = \sum_{0 \le |\alpha| \le k} \rho_{L^{p(\cdot)}(\Omega)}(\partial_\alpha u).$$

This induces a norm given by

$$\|u\|_{W^{k,p(\cdot)}(\Omega)} := \inf\Big\{ \lambda > 0 : \rho_{W^{k,p(\cdot)}(\Omega)}\Big(\frac{u}{\lambda}\Big) \le 1 \Big\} := \sum_{0 \le |\alpha| \le k} \|\partial_\alpha u\|_{p(\cdot)}.$$

For $k \in \mathbb{N}$, the space $W^{k,p(\cdot)}(\Omega)$ is called Sobolev space and its elements are called Sobolev functions. *Clearly $W^{0,p(\cdot)}(\Omega) = L^{p(\cdot)}(\Omega)$ and*

$$W^{1,p(\cdot)}(\Omega) = \Big\{ u \in L^{p(\cdot)}(\Omega) \ \text{ such that } \nabla u \text{ exists and } |\nabla u| \in L^{p(\cdot)}(\Omega) \Big\},$$

equipped with the norm

$$\|u\|_{W^{1,p(\cdot)}(\Omega)} = \|u\|_{p(\cdot)} + \||\nabla u|\|_{p(\cdot)}.$$

Theorem 5.3. *Let $p \in \mathscr{P}(\Omega)$. The space $W^{k,p(\cdot)}(\Omega)$ is a Banach space, which is separable if p is bounded, and reflexive if $1 < p_1 \le p_2 < +\infty$.*

Definition 5.8. *Let $p \in \mathscr{P}(\Omega)$ and $k \in \mathbb{N}$. The Sobolev space $W_0^{k,p(\cdot)}(\Omega)$ "with zero boundary trace" is the closure in $W^{k,p(\cdot)}(\Omega)$ of the set of $W^{k,p(\cdot)}(\Omega)$–functions with compact support, i.e.,*

$$W_0^{k,p(\cdot)}(\Omega) = \overline{\Big\{ u \in W^{k,p(\cdot)}(\Omega) : u = u\chi_K \ for \ a \ compact \ K \subset \Omega \Big\}}.$$

Remark 5.1. [25] Let $p \in \mathscr{P}(\Omega)$ and $k \in \mathbb{N}$. Then

(i) The space $H_0^{k,p(\cdot)}(\Omega)$ is defined as the closure of $C_0^\infty(\Omega)$ in $W^{k,p(\cdot)}(\Omega)$.

(ii) $H_0^{k,p(\cdot)}(\Omega) \subset W_0^{k,p(\cdot)}(\Omega)$.

(iii) If p is log-Hölder continuous on Ω, then $W_0^{k,p(\cdot)}(\Omega) = H_0^{k,p(\cdot)}(\Omega)$.

(iv) The dual of $W_0^{1,p(\cdot)}(\Omega)$ is defined as $W^{-1,p'(\cdot)}(\Omega)$, in the same way as the usual Sobolev spaces, where $\frac{1}{p(\cdot)} + \frac{1}{p'(\cdot)} = 1$.

Theorem 5.4. *Let $p \in \mathscr{P}(\Omega)$. The space $W_0^{k,p(\cdot)}(\Omega)$ is a Banach space, which is separable if p is bounded, and reflexive if $1 < p_1 \leq p_2 < +\infty$.*

Theorem 5.5. (**Poincaré's inequality**). *Let Ω be a bounded domain of $\mathbb{R}^n$ and let $p(\cdot)$ satisfy the Log-Hölder continuity property. Then*

$$\|u\|_{p(\cdot)} \leq C\|\nabla u\|_{p(\cdot)}, \quad \textit{for all } u \in W_0^{1,p(\cdot)}(\Omega),$$

where the positive constant C depends on $p(\cdot)$ and Ω only. In particular, the space $W_0^{1,p(\cdot)}(\Omega)$ has an equivalent norm given by $\|u\|_{W_0^{1,p(\cdot)}(\Omega)} = \||\nabla u|\|_{p(\cdot)}$. If $p = 2$ we set $H_0^1(\Omega) = W_0^{1,2}(\Omega)$.

Remark 5.2. Contrary to the constant-exponent case, there is no modular-type Poincaré's inequality. The following example shows that Poincaré's inequality does not, in general, hold in a modular form.

Example 5.4. [25]Let $p : (-2,2) \longrightarrow [2,3]$ be a Lipschitz continuous exponent defined by

$$p(x) = \begin{cases} 3, & if\ x \in (-2,-1) \cup (1,2) \\ 2, & if\ x \in (-\frac{1}{2}, \frac{1}{2}) \\ -2x+1, & if\ x \in [-1, -\frac{1}{2}] \\ 2x+1, & if\ x \in [\frac{1}{2}, 1]. \end{cases}$$

Let u_μ be the xLipschitz function defined by

$$u_\mu(x) = \begin{cases} \mu x + 2\mu, & if\ x \in (-2,-1] \\ \mu, & if\ x \in (-1,1) \\ -\mu x + 2\mu, & if\ x \in [1,2). \end{cases}$$

Then

$$\frac{\rho(u_\mu)}{\rho(u'_\mu)} = \frac{\int_{-2}^{2} |u_\mu|^{p(x)}\, dx}{\int_{-2}^{2} |u'_\mu|^{p(x)}\, dx} \geq \frac{\int_{-\frac{1}{2}}^{\frac{1}{2}} \mu^2\, dx}{2\int_{-2}^{-1} \mu^3\, dx} = \frac{1}{2\mu} \to \infty$$

as $\mu \to 0^+$.

This Subsection is concluded with some essential embedding results. See [25].

Lemma 5.8. *Let Ω be a bounded domain in $\mathbb{R}^n$ with a smooth boundary $\partial\Omega$. Assume that $p : \Omega \to (1,\infty)$ is a measurable function such that*

$$1 < p_1 \leq p(x) \leq p_2 < +\infty, \text{ for a.e. } x \in \Omega.$$

Assume that $p(x), q(x) \in C(\overline{\Omega})$ and that $q(x) < p^*(x)$ in $\overline{\Omega}$ with $p^*(x) = \begin{cases} \frac{np(x)}{n-p(x)}, & \text{if } p_2 < n \\ \infty, & \text{if } p_2 \geq n. \end{cases}$
Then the embedding $W_0^{1,p(\cdot)}(\Omega) \hookrightarrow L^{q(\cdot)}(\Omega)$ is continuous and compact.

As a special case, we underline the following:

Corollary 5.1. *Let Ω be a bounded domain in $\mathbb{R}^n$ with a smooth boundary $\partial\Omega$. Assume that $p : \overline{\Omega} \to (1,\infty)$ is a continuous function such that*

$$2 \leq p_1 \leq p(x) \leq p_2 < \frac{2n}{n-2}, \quad n \geq 3. \tag{5.9}$$

Then the embedding $H_0^1(\Omega) \hookrightarrow L^{p(\cdot)}(\Omega)$ is continuous and compact.

5.1.3 Blow-up for the wave equation with variable-exponent nonlinearity

In recent years, a great deal of attention has been paid to the investigation of nonlinear models of hyperbolic, parabolic and elliptic equations with nonlinearities involving variable exponents. Such probems appear, for instance, in some models for physical phenomena like flows of electrorheological fluids or fluids with temperature-dependent viscosity, filtration processes in a porous media, nonlinear viscoelasticity, and image processing. More details on this subject can be found in [2] and [3]. Interestingly, only few works have appeared on hyperbolic problems with nonlinearities of variable-exponent type. For instance, Antontsev [5] considered the equation

$$u_{tt} - \operatorname{div}(a(x,t)|\nabla u|^{p(x,t)-2}\nabla u) - \alpha\Delta u_t = b(x,t)u|u|^{\sigma(x,t)-2}$$

in a bounded domain $\Omega \subset \mathbb{R}^n$, where $\alpha > 0$ is a constant and a, b, p, σ are given functions. For specific conditions on a, b, p, σ, he proved some blow-up results, for certain solutions with non positive initial energy. He also discussed the case when $\alpha = 0$ and established a blow-up result. Subsequently, the same author [4] discussed the same equation and proved local and a global existence of some weak solutions under certain hypotheses on the functions a, b, p, σ. He also established some blow-up results for certain solutions having non positive initial energy. Guo and Gao [19] looked into the same problem [5] and established several blow-up results for certain solutions associated with negative initial energy. Precisely, they took $\sigma(x,t) = \sigma > 2$, a constant, and established a blow-up result in finite-time. For the case $\sigma(x,t) = \sigma(x)$, they claimed the same blow-up result, but no proof has been given. This work is considered to be an improvement over [5]. In [47], Sun et al. looked into the following equation:

$$u_{tt} - \operatorname{div}(a(x,t)\nabla u) + c(x,t)u_t|u_t|^{q(x,t)-1} = b(x,t)u|u|^{p(x,t)-1}$$

in a bounded domain, with Dirichlet-boundary conditions, and established a blow-up result for solutions with positive initial energy. They also gave lower and upper bounds for the blow-up time and provided numerical illustrations for their result. Recently, Messaoudi and Talahmeh [31] studied

$$u_{tt} - \operatorname{div}(|\nabla u|^{m(x)-2}\nabla u) + \mu u_t = |u|^{p(x)-2}u, \tag{5.10}$$

with Dirichlet-boundary conditions and for $\mu \geq 0$. They proved a blow-up result for certain solutions with arbitrary positive-initial-energy. This result generalized that of Korpusov [24] established for (5.10), with m and p constants. This latter result was later extended by the same authors in [32] to an equation of the form

$$u_{tt} - \operatorname{div}\left(|\nabla u|^{r(\cdot)-2}\nabla u\right) + a|u_t|^{m(\cdot)-2}u_t = b|u|^{p(\cdot)-2}u,$$

where $a,\ b > 0$ are constants and the exponents of nonlinearity $m,\ p$ and r are given functions satisfying specific conditions. They proved a finite-time blow-up result for the solutions with negative initial energy and for certain solutions with positive energy. Very recently, Messaoudi et al. [33] studied the problem

$$u_{tt} - \Delta u + au_t|u_t|^{m(\cdot)-2} = bu|u|^{p(\cdot)-2}, \tag{5.11}$$

where $a,\ b$ are positive constants. They established the existence of a unique local weak solution by using the Faedo-Galerkin method under suitable assumptions on the variable exponents m and p. They also proved the finite-time blow-up of solutions and gave a two-dimensional numerical example to illustrate the blow up result. Yunzhu Gao and Wenjie Gao [15] studied a nonlinear viscoelastic equation with variable exponents and proved the existence of weak solutions by using the Faedo-Galerkin method under suitable assumptions. Autuori et al. [8] looked into a nonlinear Kirchhoff system in the presence of the $\overrightarrow{p}(x,t)-$Laplace operator, a nonlinear force $f(t,x,u)$ and a nonlinear damping term $Q = Q(t,x,u,u_t)$. They established a global nonexistence result under suitable conditions on f, Q, p. For more results concerning the blow-up of hyperbolic problems, we refer the reader to Antontsev and Ferreira [6] and the book by Antontsev and Shmarev [7].

5.1.4 Stability of the wave equation with variable-exponent nonlinearity

There is an extensive literature on the stabilization of the wave equation by internal or boundary feedbacks. Zuazua [51] proved the exponential stability of the energy for the wave equation by a locally distributed internal feedback depending linearly on the velocity. Komornik [22] and Nakoa [42] extended the result of Zuazua by considering the case of a nonlinear damping term with a polynomial growth near the origin. Martinez [27] studied a damped wave equation and used the piecewise multiplier technique combined with some nonlinear integral inequalities to establish explicit decay rate estimates. These decay estimates are not optimal for some cases including the case of the polynomial growth. Many authors considered the following initial boundary value problem of the Kirchhoff equation with a general dissipation of the form

$$\begin{cases} u_{tt} - \phi\Big(\int_\Omega |\nabla u|^2\Big)\Delta u + \sigma(t)g(u_t) = 0, & \text{in } \Omega\times[0,+\infty) \\ u(x,t) = 0, & \text{on } \partial\Omega\times[0,+\infty) \\ u(x,0) = u_0(x), u_t(x,0) = u_1(x), & \text{in } \Omega, \end{cases} \tag{5.12}$$

where Ω is a bounded domain $\mathbb{R}^n$ $(n\geq 1)$ with a smooth boundary $\partial\Omega$ and ϕ,σ and g are given functions and the functions (u_0,u_1) are the given initial data. For instance, in the case when $g=\sigma=0$, the one-dimensional case of (5.12) was first introduced by Kirchhoff [20] in 1876, and was called the Kirchhoff string thereafter. When $\sigma=1, \phi(r)=r^\alpha$ $(\alpha\geq 1)$ and $g(x)=\tau x$ $(\tau>0)$, problem (5.12) was treated by Nishihara and Yamada [43]. They proved the existence and uniqueness of a global solution and the polynomial decay for small data $(u_0,u_1)\in\Big(H_0^1(\Omega)\cap H^2(\Omega)\Big)\times H_0^1(\Omega)$ with $u_0\neq 0$. In [44], Ono extended the work [43] to the case where $\phi(r)=r$ and $\sigma(t)\equiv(1+t)^{-\delta}$, $\delta<\frac{1}{3}$ by using the decay lemma of Nakao [39]. In [11], Benaissa and Guesmia extended the results obtained by Ono [44] and proved an existence and uniqueness theorem in Sobolev spaces, of a global solution to the problem (5.12) when $\phi(r)=r$, $g(v)=v$ and for general functions σ. Also, they obtained an explicit and general decay rate, depending on σ, g and ϕ, for the energy of solutions of (5.12), without any growth assumption on g and ϕ at the origin, or on σ at infinity. Also, the following problem

$$\begin{cases} u_{tt} - \Delta u + g(u_t) + f(u) = 0, & \text{in } \Omega\times(0,+\infty) \\ u(x,t) = 0, & \text{on } \partial\Omega\times[0,+\infty) \\ u(x,0) = u_0(x), u_t(x,0) = u_1(x), & \text{in } \Omega, \end{cases} \tag{5.13}$$

where Ω is a bounded region in $\mathbb{R}^n$ $(n\geq 1)$, with a smooth boundary $\partial\Omega$, was considered by many authors. For instance, in the case when $f(u)=$

$|u|^{p-2}u$, $g(u_t) = |u_t|^{m-2}u_t$, $m, p > 2$, Nakao [40] showed that (5.13) has a unique global weak solution if $0 \leq p-2 \leq 2/(n-2), n \geq 3$ and a global unique strong solution if $p-2 > 2/(n-2), n \geq 3$. In addition to global existence, the issue of the decay rate was also addressed. In both cases it has been shown that the energy of the solution decays algebraically if $m > 2$ and decays exponentially if $m = 2$. This improved an earlier result in [38], where Nakao studied the problem in an abstract setting and established a theorem concerning decay of the solution energy only for the case $m-2 \leq 2/(n-2), n \geq 3$. Also in a joint work, Nakao and Ono [41] extended this result to the Cauchy problem

$$\begin{cases} u_{tt} - \Delta u + \lambda^2(x)u + \rho(u_t) + f(u) = 0, & \text{in } \mathbb{R}^n \times (0, +\infty) \\ u(x,0) = u_0(x), u_t(x,0) = u_1(x), & \text{in } \mathbb{R}^n, \end{cases} \tag{5.14}$$

where $\rho(u_t)$ behaves like $|u_t|^\beta u_t$ and $f(u)$ behaves like $-bu|u|^\alpha$. In this case the authors required that the initial data be small enough in the $H^1 \times L^2$ norm and with compact supports. In [28], Messaoudi considered problem (5.13) in the case $f(u) = bu|u|^{p-2}$, $g(u_t) = a(1+|u_t|^{m-2})u_t$, $a, b > 0, p, m > 2$, and showed that, for any initial data $(u_0, u_1) \in H_0^1(\Omega) \times L^2(\Omega)$, the problem has a unique global solution with energy decaying exponentially. Benaissa and Messaoudi [9] studied (5.13), for $f(u) = -bu|u|^{p-2}$, and $g(u_t) = a(1+|u_t|^{m-2})u_t$, and showed that, for suitably chosen initial data, the problem possesses a global weak solution which decays exponentially even if $m > 2$. In [17], Guesmia looked into the following problem

$$\begin{cases} u_{tt} - \Delta u + h(\nabla u) + g(u_t) + f(u) = 0, & \text{in } \Omega \times (0, +\infty) \\ u(x,t) = 0, & \text{on } \partial\Omega \times [0, +\infty) \\ u(x,0) = u_0(x), u_t(x,0) = u_1(x), & \text{in } \Omega, \end{cases} \tag{5.15}$$

where Ω is a bounded open domain in $\mathbb{R}^n$ $(n \geq 1)$, with a smooth boundary $\partial\Omega$ and $f, g : \mathbb{R} \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}$ are continuous, nonlinear functions satisfying some general properties. He obtained uniform decay of strong and weak solutions under weak growth assumptions on the feedback function and without any control of the sign of the derivative of the energy related to the above equation. Guesmia and Messaoudi [18] considered (5.15) with $h(\nabla u) = -\nabla\phi \cdot \nabla u$, where $\phi \in W^{1,\infty}(\Omega)$, proved local and global existence results and showed that weak solutions decay either algebraically or exponentially depending on the rate of growth of g. Pucci and Serrin [46] discussed the stability of the following problem

$$\begin{cases} u_{tt} - \Delta u + Q(x,t,u,u_t) + f(x,u) = 0, & \text{in } \Omega \times (0, +\infty) \\ u(x,t) = 0, & \text{on } \partial\Omega \times [0, +\infty) \\ u(x,0) = u_0(x), u_t(x,0) = u_1(x), & \text{in } \Omega, \end{cases} \tag{5.16}$$

and proved that the energy of the solution is a Liaponov function. Although they did not discuss the issue of the decay rate, they did show that in general the energy goes to zero as t approaches infinity. They also considered an important special case of (5.16), namely when $Q(x,t,u,u_t) = a(t)t^{\alpha}u_t$ and $f(x,u) = V(x)u$, and showed that the behavior of the solution depends crucially on the parameter α. Specifically, they showed that if $|\alpha| \leq 1$ then the rest field is asymptotically stable. On the other hand, when $|\alpha| > 1$ there are solutions that do not approach zero or approach a nonzero function $\phi(x)$ as $t \to \infty$. In [16], Guesmia studied the following elasticity system

$$\begin{cases} \partial_{tt}u_i - \sigma_{ij,j} + \ell_i(x, \partial_t u_i) = 0, & \text{in } \Omega \times (0, +\infty) \\ u(x,t) = 0, & \text{on } \partial\Omega \times [0, +\infty) \\ u_i(0) = u_i^0, \partial_t u_i(0) = u_i^1, & \text{in } \Omega, \end{cases} \tag{5.17}$$

where $\ell_i(x, \partial_t u_i) = b_i(x) g_i(\partial_t u_i)$, b_i's $\in L^\infty(\Omega)$, are bounded nonnegative functions and g_i's are non-decreasing continuous real-valued functions satisfying some conditions. He proved some precise decay estimates of the energy for the system (5.17) with some localized dissipations. Zuazua [52] considered the following damped semilinear wave equation

$$u_{tt} - \Delta u + \alpha u + f(u) + a(x)u_t = 0 \text{ in } \mathbb{R}^n \times (0, \infty),$$

with $\alpha > 0$. He proved the exponential decay of the energy of the solution under suitable conditions on the functions f and a. In [10], Benaissa and Mokeddem looked into the following equation

$$u_{tt} - \operatorname{div}\left(|\nabla u|^{p-2}\nabla u\right) - \sigma(t)\operatorname{div}\left(|\nabla u_t|^{m-2}\nabla u_t\right) = 0,$$

where σ is a positive function, $p, m \geq 2$ and Ω is a bounded domain in $\mathbb{R}^n$ $(n \geq 1)$ with a regular boundary. They gave an energy-decay estimate for the solutions and extended the results of Yang [50] and Messaoudi [29]. Cavalcanti and Guesmia [12] looked into the following problem

$$\begin{cases} u_{tt} - \Delta u + F(x,t,u,\nabla u) = 0, & \text{in } \Omega \times (0, +\infty) \\ u(x,t) = 0, & \text{on } \partial\Gamma_0 \times (0, +\infty) \\ u + \int_0^t g(t-s) \frac{\partial u}{\partial \nu}(s) ds = 0, & \text{on } \partial\Gamma_1 \times (0, +\infty) \\ u(x,0) = u_0(x), u_t(x,0) = u_1(x), & \text{in } \Omega, \end{cases} \tag{5.18}$$

where Ω is a bounded region in $\mathbb{R}^n$ whose boundary is partitioned into disjoint sets Γ_0, Γ_1, under some assumptions on the relaxation function g. They proved that the dissipation given by the memory term is strong enough to assure exponential (or polynomial) decay provided that the relaxation function also decays exponentially (or polynomially). In both cases the solution decays at the same rate as that of the relaxation function. This result was later generalized by Messaoudi and Soufyane [30], where relaxation functions of general-decay type were considered. Alabau-Boussouira [1] used some weighted integral inequalities and convexity arguments and proved a semi-explicit formula which leads to decay rates of the energy in terms of the behavior of the nonlinear feedback near the origin, from which the optimal exponential and polynomial decay rate estimates are only special cases. The following problem has been widely studied in the literature:

$$\begin{cases} u_{tt} - \Delta u + \alpha(t) g(u_t) = 0, & \text{in } \Omega \times (0, +\infty) \\ u(x,t) = 0, & \text{on } \partial\Omega \times (0, +\infty), \end{cases} \tag{5.19}$$

where Ω is a bounded domain of $\mathbb{R}^n$ with a smooth boundary $\partial\Omega$ and g, α are specific functions. For instance, when $\alpha \equiv 1$ and g satisfies

$$c_1 \min\{|s|, |s|^q\} \leq |g(s)| \leq c_2 \max\{|s|, |s|^{1/q}\},$$

where $c_1, c_2 > 0$ are constants and $q > 1$, it was proved that

$$E(t) \leq C\big(E(0)\big) t^{-2/(q-1)}, \ \forall t > 0,$$

and for $q = 1$ the decay rate is exponential (see [21]). In the presence of a weak frictional damping, Benaissa and Messaoudi [9] treated problem (5.19), for g having a polynomial growth near the origin, and established energy decay results. Stabilization of wave equations with dampings of arbitrary growth have been considered for the first time in the work of Lasiecka and Tataru [26]. They showed that the energy decays as fast as the solution of an associated differential equation whose coefficients depend on the damping term. Mustafa and Messaoudi [37] considered (5.19) and established an explicit and general decay rate result, using some properties of convex functions. Their result was obtained without imposing any restrictive growth assumption on the frictional damping term. Wu and Xue [49] studied the following quasilinear hyperbolic equation

$$u_{tt} - \psi(t) \operatorname{div}\big(|\nabla u_t|^{p-2} \nabla u_t\big) - \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \sigma_i(u_{x_i}) + \mu |u_t|^{\alpha} u_t = 0,$$

where $\mu, \alpha \geq 0$, and $p \geq 2$ are constants, the functions σ_i $(i = 1, 2, ..., n)$ and ψ are nonlinear and the domain Ω is bounded in $\mathbb{R}^n$ $(n \geq 1)$ and has a regular boundary. By using multiplier methods they investigated the stability of weak solutions and obtained an explicit estimate for the rate of decay. In 2015, Mokeddem and Mansour [36] revisited the problem considered in [10] with some modifications. Specifically, they treated the equation

$$u_{tt} - \operatorname{div}\big(|\nabla u|^{p-2} \nabla u\big) - \sigma(t)\big(u_t - \operatorname{div}\big(|\nabla u_t|^{m-2} \nabla u_t\big)\big) = 0,$$

and gave the same decay result. Recently, Cavalcanti et al. [13] treated the following damped wave problem

$$\begin{cases} u_{tt} - \Delta u + a(x)u_t - \operatorname{div}(b(x)\nabla u_t) = 0, & \text{in } \Omega \times (0,+\infty) \\ u(x,t) = 0, & \text{on } \partial\Omega \times (0,+\infty) \\ u(x,0) = u_0(x), u_t(x,0) = u_1(x), & \text{in } \Omega, \end{cases} \tag{5.20}$$

where Ω is a bounded open domain in $\mathbb{R}^n$ $(n \geq 1)$, with a smooth boundary $\partial\Omega$ and $a, b : \Omega \to \mathbb{R}^+$ are nonnegative functions satisfying specific conditions. Under appropriate assumptions on the coefficients and on the initial data (u_0, u_1), they proved the stabilization of the problem (5.20). Taniguchi [48] studied the following problem with nonlinear boundary conditions:

$$\begin{cases} u_{tt}(t) - \rho(t)\Delta u(t) + b(x)u_t(t) = f(u(t)), & \text{on } \Omega \times (0,T), \\ u(t) = 0, & \text{on } \Gamma_0 \times (0,T), \\ \frac{\partial u(t)}{\partial \nu} + \gamma(u_t(t)) = 0, & \text{on } \Gamma_1 \times (0,T) \\ u(0) = u_0, u_t(0) = u_1, & \text{in } \Omega, \end{cases} \tag{5.21}$$

where $\Omega \subset \mathbb{R}^n$ is a bounded domain with a smooth boundary $\partial\Omega = \Gamma_0 \cup \Gamma_1$ and $\overline{\Gamma_0} \cap \overline{\Gamma_1} = \phi$. Under some conditions on $\|u_0\|$ and $E(0)$, the global existence and exponential decay of the energy $E(t)$ of weak solutions of (5.21) were established.

To the best of our knowledge, there aren't many stability results for wave problems involving nonstandard nonlinearities. The only works in this direction the authors are aware of, are those by Ferreira and Messaoudi [14] and by Yunzhu Gao and Wenjie Gao [15]. In [14] the authors studied a nonlinear viscoelastic plate equation with a lower order perturbation of a $\overrightarrow{p}(x,t)-$Laplacian operator of the form

$$u_{tt} + \Delta^2 u - \Delta_{\overrightarrow{p}(x,t)} u + \int_0^t h(t-s)\Delta u(s)ds - \Delta u_t + f(u) = 0,$$

where

$$\Delta_{\overrightarrow{p}(x,t)} u = \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\Big(\Big|\frac{\partial u}{\partial x_i}\Big|^{p_i(x,t)-2} \frac{\partial u}{\partial x_i}\Big), \quad \overrightarrow{p} = (p_1, p_2, \ldots\ldots, p_n)^T,$$

$h \geq 0$ is a memory kernel that decays at a general rate and f is a nonlinear function. They proved a general decay result under appropriate assumptions on h, f, and the variable exponent $\overrightarrow{p}(x,t)-$Laplacian operator. Yunzhu Gao and Wenjie Gao [15] considered the following nonlinear viscoelastic hyperbolic problem:

$$\begin{cases} \mathscr{L}(u)(x,t) = |u|^{p(x)-2}u, & \text{in } \Omega \times (0,T), \\ u(x,t) = 0, & \text{on } \partial\Omega \times [0,T) \\ u(x,0) = u_0(x),\ u_t(x,0) = u_1(x), & \text{in } \Omega, \end{cases} \tag{5.22}$$

where $\mathscr{L}(u)(x,t) = u_{tt} - \Delta u - \Delta u_{tt} + \int_0^t g(t-\tau)\Delta u(\tau)d\tau + |u_t|^{m(x)-2}u_t$, with $m(x),\ p(x)$ being continuous functions in Ω such that

$$1 < \inf_{x\in\Omega} m(x) \le m(x) \le \sup_{x\in\Omega} m(x) < +\infty,\ \ 1 < \inf_{x\in\Omega} p(x) \le p(x) \le \sup_{x\in\Omega} p(x) < +\infty$$

and

$$\forall z,\xi \in \Omega,\ |z-\xi| < 1,\ |m(z) - m(\xi)| + |p(z) - p(\xi)| \le \omega\big(|z-\xi|\big),$$

where

$$\limsup_{t\to 0^+} \omega(\tau) \ln(\frac{1}{\tau}) = C < +\infty.$$

They also assumed that

(i) $g : \mathbb{R}^+ \to \mathbb{R}^+$ is a C^1 function satisfying

$$g(0) > 0, \quad 1 - \int_0^{+\infty} g(s)ds = \ell > 0;$$

(ii) There exists $\eta > 0$ such that

$$g'(t) \le -\eta g(t), \quad t \ge 0$$

and proved the existence of a weak solution to problem (5.22). In the recent work of Messaoudi et al. [35], the authors considered the following nonlinear damped wave equation:

$$u_{tt} - \operatorname{div}\big(|\nabla u|^{r(\cdot)-2}\nabla u\big) + |u_t|^{m(\cdot)-2}u_t = 0.$$

By using a lemma by Komornik [23], they obtained decay estimates for the solution, under suitable assumptions on the variable exponents m, r and on the initial data. They also gave two numerical applications to illustrate their theoretical results.

5.2 A viscoelastic wave equation

This Section is devoted to the study of the existence and decay of solutions of the following viscoelastic or strongly damped wave problem

$$\begin{cases} \mathscr{F}(u)(x,t) = 0, & \text{in } \Omega \times (0,+\infty) \\ u(x,t) = 0, & \text{on } \partial\Omega \times [0,+\infty) \\ u(x,0) = u_0(x), u_t(x,0) = u_1(x), & \text{in } \Omega, \end{cases} \tag{5.23}$$

where $\mathscr{F}(u)(x,t) = u_{tt} - \Delta u - \Delta u_t + a|u_t|^{m(x)-2}u_t + b|u|^{p(x)-2}u$ and Ω is a bounded domain with smooth boundary, a, b are positive constants and $m(\cdot)$, $p(\cdot)$ are continuous variable exponents defined in Ω and satisfy

$$\begin{cases} 2 \leq m_1 \leq m(x) \leq m_2 < \frac{2(n-1)}{n-2}, & \forall n \geq 3; \\ 2 \leq m(x) < +\infty, & n = 1,2 \\ 2 \leq p_1 \leq p(x) \leq p_2 < \frac{2(n-1)}{n-2}, & \forall n \geq 3; \\ 2 \leq p(x) < +\infty, & n = 1,2. \end{cases} \tag{5.24}$$

This problem is a generalization of a problem considered by Messaoudi and Benaissa [9] for m and p constants. The equation in (5.23) can be regarded as a Kelvin-Voight model for a viscoelastic material in the presence of nonlinear damping and forcing terms. The reader is referred to [45] for the application of such models.

5.2.1 Existence

In this Subsection an existence Theorem is stated, which can be established by repeating the steps of the proof of Theorem 3.1 [34]. See also [4] and [32].

Theorem 5.6. *Let $(u_0,u_1) \in H_0^1(\Omega) \times L^2(\Omega)$ and let (5.24) be satisfied. Then problem (5.23) has a global unique solution such that*

$$\begin{cases} u \in L^\infty\big((0,+\infty), H_0^1(\Omega)\big), \\ u_t \in L^\infty\big((0,+\infty), L^2(\Omega)\big) \cap L^{m(\cdot)}\big(\Omega \times (0,+\infty)\big) \cap L^2\big((0,+\infty), H_0^1(\Omega)\big), \\ u_{tt} \in L^2\big((0,+\infty), H^{-1}(\Omega)\big). \end{cases}$$

The energy of the solution of (5.23) is defined by

$$E(t) := \frac{1}{2}\int_\Omega \left[u_t^2 + |\nabla u|^2\right] dx + b\int_\Omega \frac{|u|^{p(x)}}{p(x)} dx. \tag{5.25}$$

By multiplying the equation in (5.23) by u_t and integrating over Ω, one can easily see that

$$E'(t) = -\int_\Omega |\nabla u_t|^2 dx - a\int_\Omega |u_t|^{m(x)} dx. \qquad (5.26)$$

5.2.2 Decay of the solution

This Subsection is devoted to the proof of the fact that the solution of (5.23) decays exponentially under conditions (5.24).

Theorem 5.7. *Let $(u_0, u_1) \in H_0^1(\Omega) \times L^2(\Omega)$ be given. Assume that conditions (5.24) hold. Then there exist two constants $\Gamma, \lambda > 0$ such that*

$$E(t) \le \Gamma e^{-\lambda t}, \ \forall t \ge 0.$$

Proof. Define the following functional

$$H(t) = E(t) + \varepsilon \int_\Omega u u_t,$$

for $\varepsilon > 0$. It is standard to verify that, for $\varepsilon \le \varepsilon_0$, $H \sim E$. Direct differentiation using (5.23) leads to

$$\begin{aligned}
H'(t) \quad &= \quad -\int_\Omega |\nabla u_t|^2 - a\int_\Omega |u_t|^{m(x)} + \varepsilon\int_\Omega u_t^2 - \varepsilon\int_\Omega |\nabla u|^2 - \varepsilon\int_\Omega \nabla u\cdot\nabla u_t \\
&- \quad a\varepsilon\int_\Omega u u_t|u_t|^{m(x)-2} - \varepsilon b\int_\Omega |u|^{p(x)} \\
&\le \quad -\int_\Omega |\nabla u_t|^2 - a\int_\Omega |u_t|^{m(x)} + \varepsilon\int_\Omega u_t^2 - \varepsilon\int_\Omega |\nabla u|^2 - \varepsilon\int_\Omega \nabla u\cdot\nabla u_t \\
&- \quad a\varepsilon\int_\Omega u u_t|u_t|^{m(x)-2} - \varepsilon p_1 b\int_\Omega \frac{|u|^{p(x)}}{p(x)}.
\end{aligned}$$

By using (5.25), we get

$$\begin{aligned}
H'(t) \quad \le \quad & -\int_\Omega |\nabla u_t|^2 - a\int_\Omega |u_t|^{m(x)} + 2\varepsilon\int_\Omega u_t^2 \\
& -2\varepsilon E(t) - \varepsilon b(p_1 - 2)\int_\Omega \frac{|u|^{p(x)}}{p(x)} v \\
& -\varepsilon\int_\Omega \nabla u\cdot\nabla u_t - a\varepsilon\int_\Omega u u_t|u_t|^{m(x)-2}. \qquad (5.27)
\end{aligned}$$

Now, the last two terms of (5.27) can be estimated as follows:

- $\displaystyle\int_\Omega \nabla u\cdot\nabla u_t \le \frac{1}{2}\int_\Omega |\nabla u|^2 + \frac{1}{2}\int_\Omega |\nabla u_t|^2 \le E(t) + \frac{1}{2}\int_\Omega |\nabla u_t|^2.$

• For the last term, Young's inequality with $m(x)$ and its conjugate $m^\star(x) = \frac{m(x)}{m(x)-1}$ yields

$$|u_t|^{m(x)-1}|u| \leq \delta|u|^{m(x)} + c_\delta(x)|u_t|^{m(x)},$$

where $\delta > 0$ is any constant and

$$c_\delta(x) = \delta^{1-m(x)}(m(x))^{-m(x)}(m(x)-1)^{m(x)-1}.$$

It is thus clear that

$$\int_\Omega uu_t|u_t|^{m(x)-2} \leq \delta \int_\Omega |u|^{m(x)} + \int_\Omega c_\delta(x)|u_t|^{m(x)}. \tag{5.28}$$

The first term in (5.28) can be estimated as follows:

$$\int_\Omega |u|^{m(x)}dx = \int_{\Omega_+} |u|^{m(x)}dx + \int_{\Omega_-} |u|^{m(x)}dx,$$

where

$$\Omega_+ = \left\{x \in \Omega \;/\; |u(x,t)| \geq 1\right\} \quad \text{and} \quad \Omega_- = \left\{x \in \Omega \;/\; |u(x,t)| < 1\right\}.$$

So, we have

$$\int_\Omega |u|^{m(x)} \leq \int_\Omega |u|^{m_2} + \int_\Omega |u|^{m_1}.$$

By using the embeddings $H_0^1(\Omega) \hookrightarrow L^{m_1}(\Omega)$ and $H_0^1(\Omega) \hookrightarrow L^{m_2}(\Omega)$, we arrive at

$$\begin{aligned}
\int_\Omega |u|^{m(x)} &\leq c_0\left[\|\nabla u\|_2^{m_1} + \|\nabla u\|_2^{m_2}\right] \\
&\leq c_1\left[\big(E(t)\big)^{\frac{m_1-2}{2}} + \big(E(t)\big)^{\frac{m_2-2}{2}}\right]\|\nabla u\|_2^2 \\
&\leq c_1\left[\big(E(0)\big)^{\frac{m_1-2}{2}} + \big(E(0)\big)^{\frac{m_2-2}{2}}\right]\|\nabla u\|_2^2 \\
&\leq \widetilde{c}E(t).
\end{aligned}$$

Therefore, (5.28) yields

$$\int_\Omega uu_t|u_t|^{m(x)-2} \leq \delta\widetilde{c}E(t) + \int_\Omega c_\delta(x)|u_t|^{m(x)}, \;\; \forall\delta > 0,$$

and, hence, (5.27) takes the form

$$
\begin{aligned}
H'(t) \quad \leq \quad & -\left(1-\frac{\varepsilon}{2}-2\varepsilon c_p\right)\int_\Omega |\nabla u_t|^2 - a\int_\Omega (1-\varepsilon c_\delta(x))|u_t|^{m(x)} \\
& -\varepsilon(1-\delta\widetilde{c})E(t)-\varepsilon b(p_1-2)\int_\Omega \frac{|u|^{p(x)}}{p(x)}, \qquad (5.29)
\end{aligned}
$$

where c_p is the Poincaré constant.
At this point, fix $\delta > 0$ so small that $1-\delta\widetilde{c} > 0$. Once δ is fixed, use the boundedness of m to easily notice that $c_\delta(x) \leq M$. Choose $\varepsilon \leq \varepsilon_0$ so small that

$$1-\frac{\varepsilon}{2}-2\varepsilon c_p > 0 \text{ and } 1-\varepsilon M > 0.$$

Consequently, recalling that $p_1 \geq 2$, it follows from (5.29) that

$$H'(t) \leq -\gamma E(t).$$

By using the fact that $H \sim E$, it is clear that

$$H'(t) \leq -\lambda H(t), \ \forall t \geq 0.$$

A simple integration over $(0,t)$ yields

$$H(t) \leq H(0)e^{-\lambda t}.$$

This gives the desired result.

□

Bibliography

[1] F. Alabau-Boussouira, *Convexity and weighted integral inequalities for energy decay rates of nonlinear dissipative hyperbolic systems*, Applied Mathematics and Optimization 51 no. 1 (2005), 61-105.

[2] S. Antontsev, V. Zhikov, *Higher integrability for parabolic equations of $p(x,t)$-Laplacian type*, Advances in Differential Equations 10 no. 9 (2005), 1053-1080.

[3] S. Antontsev, S. Shmarev, *Blow-up of solutions to parabolic equations with nonstandard growth conditions*, Journal of computational and applied mathematics 234 no. 9 (2010), 2633-2645.

[4] S. Antontsev, *Wave equation with $p(x,t)-$Laplacian and damping term: Existence and blow-up*, Differ. Equ. Appl 3 no. 4 (2011), 503-525.

[5] S. Antontsev, *Wave equation with $p(x,t)-$Laplacian and damping term: Blow up of solutions*, Comptes Rendus Mecanique 339 no. 12 (2011), 751-755.

[6] S. Antontsev, J. Ferreira, *Existence, uniqueness and blow-up for hyperbolic equations with nonstandard growth conditions*, Nonlinear Analysis: Theory, Methods and Applications 93 (2013), 62-77.

[7] S. Antontsev, S. Shmarev, *Evolution PDEs with nonstandard growth conditions: Existence, Uniqueness, Localization, Blow up*, Atlantis Studies in Differential Equations, Volume 4, Atlantis Press 2015.

[8] G. Autuori, P. Pucci, M. Salvatori, *Global nonexistence for nonlinear Kirchhoff systems*, Arch. Rational Mech. Anal. 196 no. 2 (2010), 489-516.

[9] A. Benaissa, S.A. Messaoudi, *Exponential decay of solutions of a nonlinearly damped wave equation*, NoDEA: Nonlinear Differential Equations and Applications 12 no. 4 (2006), 391-399.

[10] A. Benaissa, S. Mokeddem, *Decay estimates for the wave equation of $p-$Laplacian type with dissipation of $m-$Laplacian type*, Mathematical methods in the applied sciences 30 no. 2 (2007), 237-247.

[11] A. Benaissa, A. Guesmia, *Global existence and general decay estimates of solutions for degenerate or nondegenerate Kirchhoff equation with general dissipation*, Diff. Inte. Equa., Preprint.

[12] M. Cavalcanti, A. Guesmia, *General decay rates of solutions to a nonlinear wave equation with boundary condition of memory type*, Differential and Integral Equations 18 no. 5 (2005), 583-600.

[13] M. Cavalcanti, V. D. Cavalcanti, L. Tebou, *Stabilization of the wave equation with localized compensating frictional and Kelvin-Voigt dissipating mechanisms*, Electronic Journal of Differential Equations 2017, no. 83 (2017), 1-18.

[14] J. Ferreira, S. A. Messaoudi, *On the general decay of a nonlinear viscoelastic plate equation with a strong damping and $\overrightarrow{p}(x,t)-$Laplacian*, Nonlinear Analysis: Theory, Methods and Applications 104 (2014), 40-49.

[15] Y. Gao, W. Gao, *Existence of weak solutions for viscoelastic hyperbolic equations with variable exponents*, Boundary Value Problems 2013 no. 1 (2013), 1-8.

[16] A. Guesmia, *On the decay estimates for elasticity systems with some localized dissipations*, Asymptotic Analysis 22, no. 1 (2000), 1-13.

[17] A. Guesmia, *A new approach of stabilization of nondissipative distributed systems*, SIAM journal on control and optimization 42, no. 1 (2003), 24-52.

[18] A. Guesmia, S. A. Messaoudi, *Decay estimates of solutions of a nonlinearly damped semilinear wave equation*, In Annales Polonici Mathematici, vol. 85, 25-36. Instytut Matematyczny Polskiej Akademii Nauk, 2005.

[19] B. Guo, W. Gao, *Blow up of solutions to quasilinear hyperbolic equations with $p(x,t)-$Laplacian and positive initial energy*, Comptes Rendus Mecanique 342 no. 9 (2014), 513-519.

[20] G. Kirchhoff, *Vorlesungen uber Mechanik*, Leipzig, Teubner, (1883).

[21] V. Komornik, E. Zuazua, *A direct method for the boundary stabilization of the wave equation*, J. Math. Pures Appl. no. 69 (1990), 33-54.

[22] V. Komornik, *Decay estimates for the wave equation with internal damping*, International Series of Numerical Mathematics no. 118 (1994), 253-266.

[23] V. Komornik, *Exact Controllability and Stabilization*, The Multiplier Method, Masson-John Wiley, Paris (1994).

[24] M.O. Korpusov, *Non-existence of global solutions to generalized dissipative Klein-Gordon equations with positive energy*, Electronic Journal of Differential Equations 2012 no. 119 (2012), 1-10.

[25] D. Lars, P. Harjulehto, P. Hasto, M. Ruzicka, *Lebesgue and Sobolev spaces with variable exponents*, Lecture Notes in Mathematics, vol. 2017 (2011).

[26] I. Lasiecka, D. Tataru, *Uniform boundary stabilization of semilinear wave equations with nonlinear boundary damping*, Differential Integral Equations 6 no. 3 (1993), 507-533.

[27] P. Martinez, *A new method to obtain decay rate estimates for dissipative systems with localized damping*, Rev.Mat. Complut. 12 no. 1 (1999), 251-283.

[28] S. A. Messaoudi, *Decay of the solution energy for a nonlinearly damped wave equation*, Arabian Journal for Science and Engineering 26 no. 1; Part A (2001), 63-68.

[29] S. A. Messaoudi, *On the decay of solutions for a class of quasilinear hyperbolic equations with non-linear damping and source terms*, Mathematical methods in the applied sciences 28 no. 15 (2005), 1819-1828.

[30] S. A. Messaoudi, A. Soufyane, *General decay of solutions of a wave equation with a boundary control of memory type*, Nonlinear Analysis: Real World Applications 11 no. 4 (2010), 2896-2904.

[31] S. A. Messaoudi, A. A. Talahmeh, *A result for a nonlinear wave equation with variable-exponent nonlinearities*, Applicable Analysis 96 no. 9 (2017), 1509-1515.

[32] S. A. Messaoudi, A. A. Talahmeh, *Blow up in solutions of a quasilinear wave equation with variable-exponent nonlinearities*, Mathematical Methods in the Applied Sciences 40 no. 18 (2017), 6976-6986.

[33] S. A. Messaoudi, A. A. Talahmeh, H.J. Al-Smail, *Nonlinear Damped Wave Equation: Existence and Blow up*, Computers and Mathematics with Applications 74 (2017), 3024-3041.

[34] S. A. Messaoudi, A. A. Talahmeh, *On wave equation: review and recent results*, Arabian J. of Math. 7 (2018), 113-154.

[35] S. A. Messaoudi, J. H. Al-Smail, A. A. Talahmeh, *Deacay for solutions of a nonlinear damped wave equation with variable-exponent nonlinearities*, Computers and Math. with appl., DOI: 10.1016/j.camwa.2018.07.035

[36] S. Mokeddem, K. B. W. Mansour, *The Rate at Which the Energy of Solutions for a Class of-Laplacian Wave Equation Decays*, International Journal of Differential Equations 2015 (2015), Art. ID 721503, 5 pages.

[37] M.I. Mustafa, S.A. Messaoudi, *General energy decay rates for a weakly damped wave equation*, Communications in Mathematical Analysis 9 no. 2 (2010), 67-76.

[38] M. Nakao, *Decay of solutions of some nonlinear evolution equations*, J. Math. Anal. Appl. 60 (1977), 542-549.

[39] M. Nakao, *A difference inequality and its applications to nonlinear evolution equations*, J. Math. Soc. Japan 30 (1978), 747-762.

[40] M. Nakao, *Remarks on the existence and uniqueness of global decaying solutions of the nonlinear dissipative wave equations*, Math Z. 206 (1991),265-275.

[41] M. Nakao, K. Ono, *Global existence to the Cauchy problem of the semilinear wave equation with a nonlinear dissipation*, Funkcial Ekvacioj 38 (1995), 417-431.

[42] M. Nakao, *Decay of solutions of the wave equation with a local nonlinear dissipation*, Math. Ann. 305 no. 3 (1996), 403-417.

[43] K. Nishihara, Y. Yamada, *On global solutions of some degenerate quasilinear hyperbolic equations with dissipative terms*, Funkcial. Ekvac. 33, no. 1 (1990), 151-159.

[44] K. Ono, *Global solvability for degenerate Kirchhoff equations with weak dissipation*, Math. Japon. 50, no. 3 (1999), 409-413.

[45] N. Ozkaya, M. Nordin, *Mechanical Properties of Biological Tissues*, Fundamentals of Biomechanics Equilibrium (Springer, 2012), Chap. 9, 195-218.

[46] P. Pucci, J. Serrin, *Asymptotic stability for nonautonomous dissipative wave system*, Comm. Pure Appl. Math 49,(1996), 177-216.

[47] L. Sun, Y. Ren, W. Gao, *Lower and upper bounds for the blow up time for nonlinear wave equation with variable sources*, Computers and Mathematics with Applications 71 no. 1 (2016), 267-277.

[48] T. Taniguchi, *Exponential boundary stabilization for nonlinear wave equations with localized damping and nonlinear boundary condition*, Communications on Pure and Applied Analysis 16 no. 5 (2017), 1571-1585.

[49] Y. Wu, X. Xue, *Decay rate estimates for a class of quasilinear hyperbolic equations with damping terms involving $p-$Laplacian*, Journal of Mathematical Physics 55(121504):2014, DOI: 10.1063/1.4904484.

[50] Z. Yang, *Existence and asymptotic behaviour of solutions for a class of quasi-linear evolution equations with non-linear damping and source terms*, Mathematical Methods in the Applied Sciences 25 no. 10 (2002), 795-814.

[51] E. Zuazua, *Exponential decay for the semi-linear wave equation with locally distributed damping,* Comm. P.D.E no. 15 (1990), 205-235.

[52] E. Zuazua, *Exponential decay for the semilinear wave equation with localized damping in unbounded domains*, J. Math. Pures Appl 70, no. 4 (1991), 513-529.

Chapter 6

Recent advances on geometrical properties of the variable exponent spaces $\ell_{p(.)}$

M. Kbiri Alaoui
Department of Mathematics, King Khalid University, Abha, Saudi Arabia,
E-mail: mka_la@yahoo.fr

M. A. Khamsi
Department of Mathematical Sciences, University of Texas at El Paso, El Paso, USA, and Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia,
E-mail: mohamed@utep.edu

Variable exponent sequences spaces, $\ell_{p(\cdot)}$, emerged naturally in 1931 as W. Orlicz [24] used them to comment on a previous work by S. Banach on lacunary trigonometric series. $\ell_{p(\cdot)}$ spaces are a particular case of the Musielak-Orlicz class introduced by Nakano in 1950.

We devote this chapter to the investigation of some recently discovered modular geometric properties of $\ell_{p(\cdot)}$ that went unnoticed for many decades. Our research is triggered by the idea of uniform convexity: norm-uniform convexity for $\ell_{p(\cdot)}$ turns out to be very cumbersome to deal with, due mainly to the rather complicated relationship between the modular and the norm defined on $\ell_{p(\cdot)}$. This difficulty is not visible in the case of a constant exponent p: in this instance the modular is simply the p^{th} power of the norm. Moreover it is well known that in the endpoint cases, namely when either $p = 1$ or $p = \infty$ one cannot expect uniform convexity in the classical sense.

We prove modular-type uniform convexity properties of $\ell_{p(\cdot)}$ that to some extent seem to be at least as natural as norm-modular convexity and remarkably, allow us to include endpoint cases.

Our results are not merely abstract constructions: we provide some applications of these new modular geometric properties. In particular we obtain a modular analogue of the classical fixed point theorem by Kirk.

6.1 Introduction

In a celebrated work by Orlicz [24], he introduced the vector space of sequences

$$\ell_{p(\cdot)} = \left\{ \{x_n\} \subset \mathbb{R}^{\mathbb{N}};\ \sum_{n=0}^{\infty} |\lambda\ x_n|^{p(n)} < \infty \text{ for some } \lambda > 0 \right\},$$

where $\{p(n)\} \subset [1,\infty)$. Though Orlicz' main interest in the cited work was to provide some input on a previously published article by S. Banach on lacunary series, it was soon realized that these spaces not only constituted a mere ad-hoc tool for that specific purpose but were part of a much more general theory that was insinuated by Orlicz himself in an example ([24, p. 207]). From the modern point of view, variable exponent spaces are a particular case of Musielak-Orlicz spaces, first introduced by Nakano [23].
As is the case with any normed space one unavoidably encounters questions on the uniform convexity of $\ell_{p(\cdot)}$.
Some problems about the geometry and topological properties of the vector space $\ell_{p(\cdot)}$ were investigated in [15, 23, 27, 28].
In particular, it is well known that $\ell_{p(\cdot)}$ is uniformly convex if and only if the exponent $p(\cdot)$ is bounded away from 1 and infinity. The geometry of $\ell_{p(\cdot)}$ when either $\inf_n p(n) = 1$ or $\sup_n p(n) = \infty$ remains largely ill-understood. In the aforementioned works by Nakano [21, 22] the formal definition of a modular that captured the essence of the definition of $\ell_{p(\cdot)}$ was given. It remains a remarkable fact that in [24] it was noted that $\ell_{p(\cdot)}$ was a special case of what is today known as a variable exponent space [7]. Well into the 20^{th} century it was realized that variable exponent spaces had tangible applications, especially in the field of material science and in fluid dynamics. Since then, the area has expanded and continues to evolve at an ever increasing rate. The seminal work by Kováčik and Rákosník [35] was, to the best of our knowledge, the first systematic treatment of the continuous Lebesgue's variable exponent spaces; Rajagopal and Ružička [25, 26] initiated a systematic mathematical study of the hydrodynamics of electrorheological fluids. This application remains one of the main driving forces in fueling the interest in the field of spaces of variable exponent. The behavior of the non-Newtonian fluids introduced in the Rajagopal-Ružička model is described by means of partial differential equations with non-standard growth. The natural habitat for the solutions of such equations are Sobolev spaces of variable exponents. Electrorheological fluids are currently being used in the defense industry, seismology, civil engineering and medicine.

Uniform convexity plays a central role in the study of the geometry of Banach spaces. In particular it is a crucial issue in fixed point theory. On the other hand, it is extremely cumbersome to handle in the absence of linearity, in particular in the modular context [20]. This makes the analysis of uniform convexity of Musielak-Orlicz spaces particularly problematic. We set about to discuss some recent research results pertaining to the investigation of some hitherto unknown properties of $\ell_{p(\cdot)}$ connected to the classical notion of modular uniform convexity.

Much of the material covered in this chapter is related to metric fixed point theory. A handy standard reference for the concepts, notation and terminology used hereafter are the books by Khamsi and Kirk [10] and by Khamsi and Kozlowski [12].

6.2 Modular vector spaces $\ell_{p(.)}$

We start by considering the sequence spaces $\ell_{p(\cdot)}$:

Definition 6.1. [24] If $p : \mathbb{N} \to [1,\infty)$, let

$$\ell_{p(\cdot)} = \Big\{ \{x_n\} \subset \mathbb{R}^{\mathbb{N}};\ \sum_{n=0}^{\infty} \frac{1}{p(n)}\ |\lambda\, x_n|^{p(n)} < \infty \text{ for some } \lambda > 0 \Big\}.$$

As mentioned in the introduction, $\ell_{p(\cdot)}$ is in particular a variable exponent space, as discussed in [24]. Motivated by these spaces, Nakano [21, 22, 23] introduced the concept of the modular vector structure, namely:

Proposition 6.1. *[15, 21, 27] Let p be as in Definition 6.1. Then, the function* $\rho : \ell_{p(\cdot)} \to [0,\infty]$ *defined by*

$$\rho(x) = \rho((x_n)) = \sum_{n=0}^{\infty} \frac{1}{p(n)}\ |x_n|^{p(n)}$$

is a convex modular on $\ell_{p(\cdot)}$*. That is to say, it satisfies the following axioms:*

(i) $\rho(x) = 0$ *if and only if* $x = 0$*,*

(ii) $\rho(\gamma x) = \rho(x)$*, if* $|\gamma| = 1$*,*

(iii) $\rho(tx+(1-t)y) \leq t\rho(x)+(1-t)\rho(y)$, *for any* $t \in [0,1]$,

for any $x, y \in \ell_{p(\cdot)}$.

We remark the fact that ρ is left-continuous, i.e., $\lim\limits_{\alpha \to 1-} \rho(\alpha x) = \rho(x)$, for any $x \in \ell_{p(\cdot)}$. Associated to the modular function ρ, there is a modular topology that captures the essence of the classical metric topology.

Definition 6.2. [11]

(a) A sequence $\{x_n\} \subset \ell_{p(\cdot)}$ is said to be ρ-convergent to $x \in \ell_{p(\cdot)}$ if and only if $\rho(x_n - x) \to 0$ as $n \to \infty$. Note that the ρ-limit is unique if it exists.

(b) A sequence $\{x_n\} \subset \ell_{p(\cdot)}$ is called ρ-Cauchy if $\rho(x_n - x_m) \to 0$ as $n, m \to \infty$.

(c) A set $C \subset \ell_{p(\cdot)}$ is called ρ-closed if for any sequence $\{x_n\} \subset C$ that ρ-converges to x, one has $x \in C$.

(d) A set $C \subset \ell_{p(\cdot)}$ is defined to be ρ-bounded if $\delta_\rho(C) = \sup\{\rho(x - y); x, y \in C\} < \infty$.

Remark 6.1. It is easy to show that $\ell_{p(\cdot)}$ is ρ-complete, i.e., any ρ-Cauchy sequence in $\ell_{p(\cdot)}$ is ρ-convergent to some element in $\ell_{p(\cdot)}$.

To exploit the analogy with the metric terminology, we introduce the following notation. For any $x \in \ell_{p(\cdot)}$ and $r \geq 0$ the x-centered modular ρ-ball of radius r is defined as $B_\rho(x,r) = \{y \in \ell_{p(\cdot)} : \rho(x-y) \leq r\}$. In the interest of a self-contained exposition, we recall the following standard definition:

Definition 6.3. A modular ρ on a metric space M is said to satisfy the Fatou property if for any sequence $\{y_n\} \subseteq M$ that ρ-converges to $y \in M$ and any $x \in M$ one has

$$\rho(x-y) \leq \liminf_{n\to\infty} \rho(x-y_n).$$

We remark the fact that the modular ρ defined in 6.1 satisfies the Fatou property in $\ell_{p(\cdot)}$. Fatou's property, in particular, implies that the ρ-balls are ρ-closed. The next property, called the Δ_2-condition, is central in modular space theory:

Definition 6.4. We say that ρ satisfies the Δ_2-condition if there exists $K \geq 0$ such that

$$\rho(2x) \leq K\,\rho(x),$$

for any $x \in \ell_{p(\cdot)}$.

For an deeper discussion on the Δ_2-condition in its several forms, we refer the reader to [12, 17, 20]. We underline the equivalence of the Δ_2 condition for the modular ρ on $\ell_{p(\cdot)}$ and the condition $\limsup_{n\to\infty} p(n) < +\infty$ [15, 21, 27].
The Luxemburg norm on $\ell_{p(\cdot)}$ is defined as the Minkowski's functional of the modular unit ball; in other words:

$$\|x\|_\rho = \inf\left\{\lambda > 0; \rho\left(\frac{1}{\lambda}x\right) \leq 1\right\}.$$

It is well known that $(\ell_{p(\cdot)}, \|.\|_\rho)$ is a Banach space.

6.3 Modular uniform convexity

The uniform convexity in $(\ell_{p(\cdot)}, \|.\|_\rho)$ was well investigated since the emergence of the theory. For example, it is well known that $(\ell_{p(\cdot)}, \|.\|_\rho)$ is uniformly convex if and only if $1 < \liminf_{n\to\infty} p(n) \leq \limsup_{n\to\infty} p(n) < \infty$ [27]. The latter implies the superreflexivity of $(\ell_{p(\cdot)}, \|.\|_\rho)$ [4, 8].
Since uniform convexity fails for the classical sequence-spaces ℓ_1 and ℓ_∞, it is not surprising that the classical definition of uniform convexity poses a challenge when either $\limsup_{n\to\infty} p(n) = \infty$ or $\liminf_{n\to\infty} p(n) = 1$. In this regard, a new modular uniform convexity-type property of $\ell_{p(\cdot)}$ was recently discovered in [2]. Before we dive into the new modular geometric properties satisfied by $\ell_{p(\cdot)}$, we need the following definitions. Recall that modular uniform convexity was introduced in general vector spaces by Nakano [22]. Its study in Orlicz function spaces was carried out in [7, 20].

Definition 6.5. [1, 7, 20] We define the following *uniform convexity type* properties of the modular ρ:

(a) [22] Let $r > 0$ and $\varepsilon > 0$. Define

$$D_1(r,\varepsilon) = \left\{(x,y);\ x,y \in \ell_{p(\cdot)}, \rho(x) \leq r, \rho(y) \leq r, \rho(x-y) \geq \varepsilon r\right\}.$$

If $D_1(r,\varepsilon) \neq \emptyset$, let

$$\delta_1(r,\varepsilon) = \inf\left\{1 - \frac{1}{r}\,\rho\left(\frac{x+y}{2}\right);\ (x,y) \in D_1(r,\varepsilon)\right\}.$$

If $D_1(r,\varepsilon) = \emptyset$, we set $\delta_1(r,\varepsilon) = 1$. We say that ρ satisfies the uniform convexity (UC) if for every $r > 0$ and $\varepsilon > 0$, we have $\delta_1(r,\varepsilon) > 0$. Note, that for every $r > 0$, $D_1(r,\varepsilon) \neq \emptyset$, for $\varepsilon > 0$ small enough.

(b) [12] We say that ρ satisfies (UUC) if for every $s \geq 0$ and $\varepsilon > 0$, there exists $\eta_1(s,\varepsilon) > 0$ depending on s and ε such that

$$\delta_1(r,\varepsilon) > \eta_1(s,\varepsilon) > 0 \text{ for } r > s.$$

(c) [12] Let $r > 0$ and $\varepsilon > 0$. Define

$$D_2(r,\varepsilon) = \left\{(x,y);\ x,y \in \ell_{p(\cdot)}, \rho(x) \leq r, \rho(y) \leq r, \rho\left(\frac{x-y}{2}\right) \geq \varepsilon r\right\}.$$

If $D_2(r,\varepsilon) \neq \emptyset$, let

$$\delta_2(r,\varepsilon) = \inf\left\{1 - \frac{1}{r}\,\rho\left(\frac{x+y}{2}\right);\ (x,y) \in D_2(r,\varepsilon)\right\}.$$

If $D_2(r,\varepsilon) = \emptyset$, we set $\delta_2(r,\varepsilon) = 1$. We say that ρ satisfies $(UC2)$ if for every $r > 0$ and $\varepsilon > 0$, we have $\delta_2(r,\varepsilon) > 0$. Note, that for every $r > 0$, $D_2(r,\varepsilon) \neq \emptyset$, for $\varepsilon > 0$ small enough.

(d) [12] We say that ρ satisfies $(UUC2)$ if for every $s \geq 0$ and $\varepsilon > 0$, there exists $\eta_2(s,\varepsilon) > 0$ depending on s and ε such that

$$\delta_2(r,\varepsilon) > \eta_2(s,\varepsilon) > 0 \; for \; r > s.$$

(e) [22] ρ is said to be uniformly convex in every direction (in short (UCED)) if and only if for any $z_1, z_2 \in \ell_{p(\cdot)}$ such that $z_1 \neq z_2$ and $R > 0$, there exists $\delta = \delta(z_1,z_2,R) > 0$ such that

$$\left\{\begin{array}{lcl} \rho(x-z_1) & \leq & R \\ \rho(x-z_2) & \leq & R \end{array}\right. \implies \rho\left(x - \frac{z_1+z_2}{2}\right) \leq R(1-\delta),$$

for any $x \in \ell_{p(\cdot)}$. ρ is said to be (UUCED) if $\delta(z_1,z_2,R) \geq \delta(z_1,z_2,\bar{R})$, whenever $R \leq \bar{R}$.

(f) [22] We say that ρ is strictly convex, (SC), if for every $x,y \in \ell_{p(\cdot)}$ such that $\rho(x) = \rho(y)$ and

$$\rho\left(\frac{x+y}{2}\right) = \frac{\rho(x)+\rho(y)}{2},$$

we have $x = y$.

Note that (UC) and $(UC2)$ are equivalent if ρ satisfies the Δ_2-condition [12]. In this case, we must have $\sup_{n\in\mathbb{N}} p(n) < \infty$. Moreover, it is easy to show that $(UUC2)$ implies $(UUCED)$. Since $(UUCED)$ implies (SC), we conclude that $(UUC2)$ also implies (SC).

6.4 Modular uniform convexity in $\ell_{p(.)}$

The following technical result will be of central importance in the sequel:

Lemma 6.1. *The following inequalities are valid:*

(i) [5] If $p \geq 2$, then we have

$$\left|\frac{a+b}{2}\right|^p + \left|\frac{a-b}{2}\right|^p \leq \frac{1}{2}\Big(|a|^p + |b|^p\Big),$$

for any $a, b \in \mathbb{R}$.

(ii) [27] If $1 < p \leq 2$, then we have

$$\left|\frac{a+b}{2}\right|^p + \frac{p(p-1)}{2}\ \left|\frac{a-b}{|a|+|b|}\right|^{2-p} \left|\frac{a-b}{2}\right|^p \leq \frac{1}{2}\Big(|a|^p + |b|^p\Big),$$

for any $a, b \in \mathbb{R}$ such that $|a| + |b| \neq 0$.

Before we state the main result of this work, we agree on the following notation:

$$\rho_K(x) = \rho_K((x_n)) = \sum_{n\in K} \frac{1}{p(n)}\ |x_n|^{p(n)},$$

for any $K \subset \mathbb{N}$ and any $x \in \ell_{p(\cdot)}$. If $K = \emptyset$, we set $\rho_K(x) = 0$.

We also agree on further terminology: the function $p(\cdot)$ will be said to satisfy the (AO) condition if the set $\{n \in \mathbb{N}; p(n) = 1\}$ has at most one element.

Theorem 6.1. *[3] Consider the vector space $\ell_{p(\cdot)}$. The following statements are equivalent:*

(i) $p(\cdot)$ satisfies the condition (AO);

(ii) the modular ρ is (UUCED);

(iii) the modular ρ is (SC).

Proof. Note that (ii) easily implies (iii). We next prove that (iii) implies (i). Assume that ρ is (SC) and that $p(\cdot)$ fails the condition (AO). Hence there exist $i, j \in \mathbb{N}$ such that $i \neq j$ and $p(i) = p(j) = 1$. Set $x = (x_n)$, where $x_n = 0$ if $n \neq i$ and $x_i = 1$, and $y = (y_n)$, where $y_n = 0$ if $n \neq j$ and $y_j = 1$. Then:

$$\rho(x) = \rho(y) = \rho\left(\frac{x+y}{2}\right) = 1$$

and $x \neq y$. This will contradict our assumption that ρ is (SC). In order to complete the proof of Theorem 6.1, we only need to prove that (i) implies (ii). Assume that $p(\cdot)$ satisfies the condition (AO). Let us prove that ρ is $(UUCED)$. Let $z_1 = (z_n^1)$, $z_2 = (z_n^2) \in \ell_{p(\cdot)}$ such that $z_1 \neq z_2$. Let $R > 0$ and $x \in \ell_{p(\cdot)}$ such that

$$\rho(x - z_1) \leq R \ \text{and} \ \ \rho(x - z_2) \leq R.$$

Set $K = \{n \in \mathbb{N};\ z_n^1 \neq z_n^2\}$. Since $z_1 \neq z_2$, K is not empty. We have $K = K_1 \cup K_2 \cup K_3$, where $K_1 = \{n \in K;\ p(n) \geq 2\}$, $K_2 = \{n \in K;\ 1 < p(n) < 2\}$ and $K_3 = \{n \in K;\ p(n) = 1\}$. Since K is not empty, one of the subsets K_1, K_2 or K_3 is not empty.

First case: assume K_1 is not empty. Using Lemma 6.1, we have:

$$\left|x_i - \frac{z_i^1 + z_i^2}{2}\right|^{p(i)} + \left|\frac{z_i^1 - z_i^2}{2}\right|^{p(i)} \leq \frac{1}{2}\Big(|x_i - z_i^1|^{p(i)} + |x_i - z_i^2|^{p(i)}\Big),$$

for any $i \in K_1$. Moreover, we have:

$$\left|x_n - \frac{z_n^1 + z_n^2}{2}\right|^{p(n)} \leq \frac{1}{2}\Big(|x_n - z_n^1|^{p(n)} + |x_n - z_n^2|^{p(n)}\Big),$$

for any $n \notin K_1$, which implies:

$$\rho\left(x - \frac{z_1 + z_2}{2}\right) + \sum_{i \in K_1} \frac{1}{p(i)} \left|\frac{z_i^1 - z_i^2}{2}\right|^{p(i)} \leq \frac{\rho(x - z_1) + \rho(x - z_2)}{2} \leq R.$$

In this case, we take

$$\delta(z_1, z_2, R) = \frac{1}{R} \sum_{i \in K_1} \frac{1}{p(i)} \left|\frac{z_i^1 - z_i^2}{2}\right|^{p(i)} > 0.$$

Second case: assume K_2 is not empty. Using Lemma 6.1, we have

$$\left|x_i - \frac{z_i^1 + z_i^2}{2}\right|^{p(i)} + A_i \leq \frac{1}{2}\Big(|x_i - z_i^1|^{p(i)} + |x_i - z_i^2|^{p(i)}\Big),$$

where

$$\begin{aligned} A_i &= \frac{p(i)(p(i)-1)}{2} \left|\frac{z_i^1 - z_i^2}{|x_i - z_i^1| + |x_i - z_i^2|}\right|^{2-p(i)} \left|\frac{z_i^1 - z_i^2}{2}\right|^{p(i)} \\ &= \frac{p(i)(p(i)-1)}{2^{1+p(i)}} \frac{1}{\Big(|x_i - z_i^1| + |x_i - z_i^2|\Big)^{2-p(i)}} |z_i^1 - z_i^2|^2, \end{aligned}$$

for any $i \in K_2$. On other hand, we use the inequalities $\rho(x - z_1) \leq R$ and $\rho(x - z_2) \leq R$ to get:

$$\frac{1}{p(i)}|x_i - z_i^1|^{p(i)} \leq R \;\; and \;\; \frac{1}{p(i)}|x_i - z_i^2|^{p(i)} \leq R,$$

which implies:

$$\begin{aligned} |x_i - z_i^1| + |x_i - z_i^2| &\leq 2\Big(p(i)\, R\Big)^{1/p(i)} \\ &\leq 2\,\Big(2\, R\Big)^{1/p(i)} \\ &= 2^{1+1/p(i)}\, R^{1/p(i)} \\ &\leq 2^2\, R^{1/p(i)}, \end{aligned}$$

for any $i \in K_2$. Hence,

$$\Big(|x_i - z_i^1| + |x_i - z_i^2|\Big)^{2-p(i)} \leq \Big(2^2\, R^{1/p(i)}\Big)^{2-p(i)} \leq 2^2\, R^{(2-p(i))/p(i)};$$

this yields:

$$\begin{aligned} \frac{1}{2^{1+p(i)}} \frac{1}{\Big(|x_i - z_i^1| + |x_i - z_i^2|\Big)^{2-p(i)}} &\geq \frac{1}{2^{1+p(i)}} \frac{1}{2^2\, R^{(2-p(i))/p(i)}} \\ &\geq \frac{1}{2^5\, R^{(2-p(i))/p(i)}}. \end{aligned}$$

In all,

$$A_i \geq \frac{p(i)(p(i)-1)}{2^5\, R^{(2-p(i))/p(i)}} |z_i^1 - z_i^2|^2,$$

for any $i \in K_2$. Therefore, we have:

$$\sum_{i\in K_2} \frac{(p(i)-1)}{2^5\, R^{(2-p(i))/p(i)}} |z_i^1 - z_i^2|^2 \leq \sum_{i\in K_2} A_i,$$

from which it follows immediately that

$$\rho\left(x-\frac{z_1+z_2}{2}\right)+\sum_{i\in K_2}\frac{(p(i)-1)}{2^5\ R^{(2-p(i))/p(i)}}\ |z_i^1-z_i^2|^2\leq\frac{\rho(x-z_1)+\rho(x-z_2)}{2}\leq R.$$

In this case, we take

$$\delta(z_1,z_2,R)=\sum_{i\in K_2}\frac{(p(i)-1)}{2^5\ R^{1+(2-p(i))/p(i)}}\ |z_i^1-z_i^2|^2=\sum_{i\in K_2}\frac{(p(i)-1)}{2^5\ R^{2/p(i)}}\ |z_i^1-z_i^2|^2>0.$$

Third case: assume $K_1=K_2=\emptyset$ and K_3 is not empty. Since $p(\cdot)$ satisfies the condition (AO), then $K=K_3$ is a singleton, i.e., $K=\{i\}$. Our assumptions on x,z_1 and z_2 imply $\rho_{K^c}(x-z_1)=\rho_{K^c}(x-z_2)=\rho_{K^c}(x-(z_1+z_2)/2)=R(x)$ and

$$\begin{cases}\rho_K(x-z_1) & = & |x_i-z_i^1|\leq R-R(x)\\ \rho_K(x-z_2) & = & |x_i-z_i^2|\leq R-R(x).\end{cases}$$

We have:

$$\left|x_i-\frac{z_i^1+z_i^2}{2}\right|^2+\left|\frac{z_i^1-z_i^2}{2}\right|^2=\frac{|x_i-z_i^1|^2+|x_i-z_i^2|^2}{2}\leq(R-R(x))^2$$

and from here one concludes that

$$\left|x_i-\frac{z_i^1+z_i^2}{2}\right|^2\leq(R-R(x))^2\left(1-\frac{1}{4(R-R(x))^2}|z_i^1-z_i^2|^2\right),$$

i.e.,

$$\rho_K\left(x-\frac{z_1+z_2}{2}\right)\leq(R-R(x))\left(1-\frac{1}{4(R-R(x))^2}|z_i^1-z_i^2|^2\right)^{1/2}.$$

Hence

$$\rho\left(x-\frac{z_1+z_2}{2}\right)\leq(R-R(x))\left(1-\frac{1}{4(R-R(x))^2}|z_i^1-z_i^2|^2\right)^{1/2}+R(x).$$

If we set $\Delta=1-\sqrt{1-\frac{1}{4(R-R(x))^2}|z_i^1-z_i^2|^2}$, we have

$$\rho\left(x-\frac{z_1+z_2}{2}\right)\leq R\left(1-\frac{R-R(x)}{R}\ \Delta\right).$$

On the other hand, note that

$$|z_i^1-z_i^2|\leq|z_i^1-x|+|x-z_i^2|\leq 2(R-R(x))\leq 2R.$$

Hence

$$1-\sqrt{1-\frac{1}{4(R-R(x))^2}|z_i^1-z_i^2|^2} \geq 1-\sqrt{1-\frac{1}{4R^2}|z_i^1-z_i^2|^2},$$

which implies that

$$\begin{aligned} \frac{R-R(x)}{R}\,\Delta &= \frac{R-R(x)}{R}\left(1-\sqrt{1-\frac{1}{4(R-R(x))^2}|z_i^1-z_i^2|^2}\right) \\ &\geq \frac{|z_i^1-z_i^2|}{2R}\left(1-\sqrt{1-\frac{1}{4R^2}|z_i^1-z_i^2|^2}\right). \end{aligned}$$

Set $\delta(z_1,z_2,R)=\dfrac{|z_i^1-z_i^2|}{2R}\left(1-\sqrt{1-\dfrac{1}{4R^2}|z_i^1-z_i^2|^2}\right)$. Therefore, we have:

$$\rho\left(x-\frac{z_1+z_2}{2}\right) \leq R(1-\delta(z_1,z_2,R)).$$

Note that $\delta(z_1,z_2,R)>0$. In all cases, we have $\delta(z_1,z_2,R)\geq\delta(z_1,z_2,\bar{R})$, whenever $R\leq\bar{R}$, i.e., the modular ρ is $(UUCED)$ as claimed.

□

Next we discuss the modular uniform convexity in $\ell_{p(\cdot)}$.

Theorem 6.2. *[2] Consider the vector space $\ell_{p(\cdot)}$. If $\inf_{n\in\mathbb{N}} p(n)>1$, then the modular ρ is $(UUC2)$.*

Proof. Assume $A=\inf_{n\in\mathbb{N}} p(n)>1$. Let $r>0$ and $\varepsilon>0$. Take $x,y\in\ell_{p(\cdot)}$ such that

$$\rho(x)\leq r,\ \ \rho(y)\leq r\ \ \textit{and}\ \ \rho\left(\frac{x-y}{2}\right)\geq r\,\varepsilon.$$

Since ρ is convex, it holds that:

$$r\,\varepsilon\leq\rho\left(\frac{x-y}{2}\right)\leq\frac{\rho(x)+\rho(y)}{2}\leq r,$$

which implies that $\varepsilon\leq 1$. Next, set $I=\{n\in\mathbb{N}; p(n)\geq 2\}$ and $J=\{n\in\mathbb{N}; p(n)<2\}=\mathbb{N}\setminus I$. Note that $\rho(z)=\rho_I(z)+\rho_J(z)$, for any $z\in\ell_{p(\cdot)}$. From our assumptions, we have either $\rho_I((x-y)/2)\geq r\,\varepsilon/2$ or $\rho_J((x-y)/2)\geq$

$r\,\varepsilon/2$.
Assume first $\rho_I((x-y)/2) \geq r\,\varepsilon/2$. Using Lemma 6.1, we conclude that

$$\rho_I\left(\frac{x+y}{2}\right)+\rho_I\left(\frac{x-y}{2}\right) \leq \frac{\rho_I(x)+\rho_I(y)}{2},$$

which implies:

$$\rho_I\left(\frac{x+y}{2}\right) \leq \frac{\rho_I(x)+\rho_I(y)}{2}-\frac{r\,\varepsilon}{2}.$$

Since

$$\rho_J\left(\frac{x+y}{2}\right) \leq \frac{\rho_J(x)+\rho_J(y)}{2},$$

we get:

$$\rho\left(\frac{x+y}{2}\right) \leq \frac{\rho(x)+\rho(y)}{2}-\frac{r\,\varepsilon}{2} \leq r\left(1-\frac{\varepsilon}{2}\right).$$

For the second case, assume $\rho_J((x-y)/2) \geq r\,\varepsilon/2$. Set $C = \varepsilon/4$,

$$J_1 = \Big\{n \in J;\ |x_n - y_n| \leq C(|x_n|+|y_n|)\Big\} \ \textit{and}\ \ J_2 = J \setminus J_1.$$

We have:

$$\rho_{J_1}\left(\frac{x-y}{2}\right) \leq \sum_{n\in J_1} \frac{C^{p(n)}}{p(n)}\ \left|\frac{|x_n|+|y_n|}{2}\right|^{p(n)} \leq \frac{C}{2}\sum_{n\in J_1}\frac{|x_n|^{p(n)}+|y_n|^{p(n)}}{p(n)},$$

because $C \leq 1$ and the power function is convex. Hence,

$$\rho_{J_1}\left(\frac{x-y}{2}\right) \leq \frac{C}{2}\Big(\rho_{J_1}(x)+\rho_{J_1}(y)\Big) \leq \frac{C}{2}\Big(\rho(x)+\rho(y)\Big) \leq C\,r.$$

Since $\rho_J((x-y)/2) \geq r\,\varepsilon/2$, we get:

$$\rho_{J_2}\left(\frac{x-y}{2}\right) = \rho_J\left(\frac{x-y}{2}\right)-\rho_{J_1}\left(\frac{x-y}{2}\right) \geq \frac{r\,\varepsilon}{2}-C\,r.$$

For any $n \in J_2$, we have:

$$A-1 \leq p(n)(p(n)-1)\ \ \textit{and}\ \ C \leq C^{2-p(n)} \leq \left|\frac{x_n-y_n}{|x_n|+|y_n|}\right|^{2-p(n)},$$

which implies by Lemma 6.1:

$$\left|\frac{x_n+y_n}{2}\right|^{p(n)}+\frac{(A-1)C}{2}\ \left|\frac{x_n-y_n}{2}\right|^{p(n)} \leq \frac{1}{2}\Big(|x_n|^{p(n)}+|y_n|^{p(n)}\Big).$$

Hence,

$$\rho_{J_2}\left(\frac{x+y}{2}\right) + \frac{(A-1)C}{2}\,\rho_{J_2}\left(\frac{x-y}{2}\right) \leq \frac{\rho_{J_2}(x)+\rho_{J_2}(y)}{2}.$$

The latter implies

$$\rho_{J_2}\left(\frac{x+y}{2}\right) \leq \frac{\rho_{J_2}(x)+\rho_{J_2}(y)}{2} - r\,\frac{(A-1)\varepsilon^2}{8},$$

since $C = \varepsilon/4$. Therefore, we have:

$$\rho\left(\frac{x+y}{2}\right) \leq r - r\,\frac{(A-1)\varepsilon^2}{8} = r\left(1 - \frac{(A-1)\varepsilon^2}{8}\right).$$

Using the definition of $\delta_2(r,\varepsilon)$, we conclude that

$$\delta_2(r,\varepsilon) \geq \min\left(\frac{\varepsilon}{2}, (A-1)\frac{\varepsilon^2}{8}\right) > 0.$$

Therefore, ρ is $(UC2)$. Moreover, if we set $\eta_2(r,\varepsilon) = \min\big(\varepsilon/2, (A-1)\varepsilon^2/8\big)$, we conclude that ρ is in fact $(UUC2)$.

$\square$

Remark 6.2. Note that in our proof above, we showed that $\eta_2(r,\varepsilon)$ is in fact a function of ε only. We will make use of this fact throughout this work.

6.5 Applications

Using the new modular uniform convexity property of $\ell_{p(\cdot)}$ proved in the preceding Section, we can prove some interesting modular geometric properties which are not clear to hold in the absence of the Δ_2-condition.

Definition 6.6. Let C be a nonempty ρ-closed convex subset of $\ell_{p(\cdot)}$. C is said to be ρ-proximinal, if for any $x \in \ell_{p(\cdot)}$ such that $d_\rho(x,C) = \inf\{\rho(x-c); c \in C\} < \infty$, the set $P_{\rho,C}(x) = \{c \in C;\ \rho(x-c) = d_\rho(x,C)\}$ is not empty. Moreover if $P_{\rho,C}(x)$ is a singleton for any $x \in \ell_{p(\cdot)}$, we say that C is a Čebyšev subset.

In the next theorem, we investigate the relationship between the uniform convexity of the modular and the *unique best approximant property* which generalizes well known properties of Hilbert and uniformly convex Banach spaces. For other results on best approximation in modular spaces see [13, 18].

Theorem 6.3. *Consider the vector space $\ell_{p(\cdot)}$. Assume the modular ρ is $(UUC2)$. Let $C \subset \ell_{p(\cdot)}$ be nonempty, convex, and ρ-closed. Let $x \in \ell_{p(\cdot)}$ be such that $d_\rho(x,C) < \infty$. There exists then a unique best ρ-approximant of x in C, that is, a unique $y \in C$ such that $\rho(x-y) = d_\rho(f,C)$, i.e., C is a Čebyšev subset.*

Proof. The uniqueness follows from the strict convexity of ρ, since $(UUC2)$ implies (SC). Let us prove the existence of the ρ-approximant. Since C is ρ-closed, we may assume without loss of any generality that $d := d_\rho(x,C) > 0$, i.e., $x \notin C$. Consider a sequence $\{y_n\} \in C$ such that

$$\rho(x-y_n) \leq d\left(1+\frac{1}{n}\right).$$

We claim that $\left\{\frac{1}{2}y_n\right\}$ is ρ-Cauchy. Assume, on the contrary, that it is not. Then there exist an $\varepsilon > 0$ and a subsequence $\{y_{n_k}\}$ of $\{y_n\}$ such that

$$\rho\left(\frac{y_{n_k}-y_{n_p}}{2}\right) \geq \varepsilon,$$

for any $p,k \geq 1$. Since ρ is (UUC2), we have:

$$\rho\left(x-\frac{y_{n_k}+y_{n_p}}{2}\right) \leq \left(1-\delta_2\left(d(k,p),\frac{\varepsilon}{d(k,p)}\right)\right)d(k,p),$$

where $d(k,p) = \left(1+\frac{1}{\min(n_p,n_k)}\right)d$. For $p,k \geq 1$, we have $d(k,p) \leq 2\,d$. Hence,

$$\delta_2\left(d(k,p),\frac{\varepsilon}{d(k,p)}\right) \geq \delta_2\left(d(k,p),\frac{\varepsilon}{2d}\right).$$

Since ρ is (UUC2), there exists $\eta > 0$ such that

$$\delta_2\left(r,\frac{\varepsilon}{2d}\right) \geq \eta$$

for any $r > \frac{d}{3}$. Since $d(k,p) \geq d > \frac{d}{3}$, we get:

$$\rho\left(x-\frac{y_{n_k}+y_{n_p}}{2}\right) \leq (1-\eta)\,d(k,p),$$

for any $k, p \geq 1$. By the convexity of C, $\dfrac{y_{n_k} + y_{n_p}}{2} \in C$. Using the definition of d it follows that

$$d \leq \rho\left(x - \frac{y_{n_k} + y_{n_p}}{2}\right) \leq (1-\eta)\, d(k,p),$$

for any $k, p \geq 1$. If we let k, p go to infinity, we get $d \leq (1-\eta)d$, which is impossible. Hence $\left\{\dfrac{1}{2} y_n\right\}$ is ρ-Cauchy. By the ρ-completeness of $\ell_{p(\cdot)}$, $\left\{\dfrac{1}{2} y_n\right\}$ ρ-converges to some $z \in \ell_{p(\cdot)}$. Fix $m \geq 1$. Since $\left\{\dfrac{y_m + y_n}{2}\right\} \subset C$ and it ρ-converges to $\dfrac{y_m}{2} + z$, given that C is ρ-closed, one concludes $\dfrac{y_m}{2} + z \in C$. Letting $m \to \infty$, we get $2z \in C$. Using the Fatou property, passing to a subsequence if necessary, it follows that

$$\rho(x - 2z) \leq \liminf_{n\to\infty} \rho\left(x - z - \frac{y_n}{2}\right) \leq \liminf_{n\to\infty}\, \liminf_{m\to\infty} \rho\left(x - \frac{y_n + y_m}{2}\right).$$

Since ρ is convex, it is apparent thatx

$$\liminf_{n\to\infty}\, \liminf_{m\to\infty} \rho\left(x - \frac{y_n + y_m}{2}\right) \leq \liminf_{n\to\infty}\, \liminf_{m\to\infty} \frac{\rho(x - y_n) + \rho(x - y_m)}{2} = d.$$

Hence, $\rho(x - 2z) \leq d$. Since $2z \in C$, it is clear that $\rho(x - 2z) = d$. In other words, $y = 2z$ is the ρ-approximant of x in C, as claimed.

$\square$

Theorem 6.3 allows us to establish a relationship between the modular uniform convexity and a property which is a modular equivalent of the Milman-Pettis theorem, which states that the uniform convexity of a Banach space implies its reflexivity. First, we need the following Definition:

Definition 6.7. [9, 12] Let C be a nonempty ρ-closed, ρ-bounded and convex subset of $\ell_{p(\cdot)}$. We will say that C satisfies the property (R), if for any decreasing sequence $\{C_n\}_{n\geq 1}$ of ρ-closed, convex, nonempty subsets of C, we have $\bigcap_{n\geq 1} C_n$ is nonempty.

Theorem 6.4. *Consider the vector space $\ell_{p(\cdot)}$. Assume the modular ρ is $(UUC2)$. Then any ρ-closed, ρ-bounded and convex nonempty subset C of $\ell_{p(\cdot)}$ satisfies the property (R).*

Proof. Let $\{C_n\}$ be a nonincreasing sequence of nonempty, ρ-closed, convex subsets of C. According to Definition 6.7 we need to demonstrate that $\{C_n\}$ has nonempty intersection. Fix any $x \in C$. Since C is ρ-bounded, there exists $M > 0$ such that for any $n \geq 1$, we have $\rho(x-y) < M$ for any $y \in C_n \subset C$. Using Theorem 6.3, there exists a unique $c_n \in C_n$ such that

$$\rho(x - c_n) = d_\rho(x, C_n) = \inf\{\rho(x-c); c \in C_n\},$$

for every $n \geq 1$. It is easy to show that the sequence $\{d_\rho(f, C_n)\}$ is increasing and bounded by M. Hence $\lim_{n\to\infty} d_\rho(x, C_n) = d$ exists. If $d = 0$, then $d_\rho(x, C_n) = 0$, for any $n \geq 1$, i.e., $x \in \bigcap_{n\geq 1} C_n$ since $\{C_n\}$ are ρ-closed. Otherwise, assume $d > 0$. Following the same ideas used in the proof of Theorem 6.3, it can be easily shown that $\left\{\frac{1}{2}c_n\right\}$ is ρ-Cauchy. Therefore it ρ-converges to some $z \in \ell_{p(\cdot)}$. Let us prove that $2z \in C_n$, for any $n \geq 1$. Indeed, we have $\frac{c_k + c_p}{2} \in C_n$, for any $p, k \geq n$. Fix any $k \geq n$. Since $\left\{\frac{c_k + c_p}{2}\right\}$ ρ-converges to $\frac{c_k}{2} + z$ as $p \to \infty$, and C_n is ρ-closed, then $\frac{c_k}{2} + z \in C_n$, for any $k \geq n$. If we let $k \to \infty$, we get $2z \in C_n$, for any $n \geq 1$. Hence $\bigcap_{n\geq 1} C_n \neq \emptyset$, which completes the proof of Theorem 6.4.

□

In fact, the conclusion of Theorem 6.4 may be improved to have an intersection property for any family of subsets.

Theorem 6.5. *Consider the vector space $\ell_{p(\cdot)}$. Assume the modular ρ is $(UUC2)$. Let $\{C_\alpha\}_{\alpha\in\Gamma}$ be a decreasing family of nonempty, convex, ρ-closed subsets of $\ell_{p(\cdot)}$, where $(\Gamma, \prec)$ is upward directed. Assume that there exists $x \in \ell_{p(\cdot)}$ such that $\sup_{\alpha\in\Gamma} d_\rho(x, C_\alpha) < \infty$. Then $\bigcap_{\alpha\in\Gamma} C_\alpha$ is not empty.*

Proof. Set $d = \sup_{\alpha\in\Gamma} d_\rho(x, C_\alpha)$. Without loss of generality, we may assume $d > 0$. For any $n \geq 1$, there exists $\alpha_n \in \Gamma$ such that

$$d\left(1 - \frac{1}{n}\right) < d_\rho(x, C_{\alpha_n}) \leq d.$$

Since $(\Gamma, \prec)$ is upward directed, we may assume $\alpha_n \prec \alpha_{n+1}$. In particular we have $C_{\alpha_{n+1}} \subset C_{\alpha_n}$ for any $n \geq 1$. Theorem 6.4 implies $C_0 = \bigcap_{n\geq 1} C_{\alpha_n} \neq \emptyset$. Clearly C_0 is convex, ρ-closed and

$$d_\rho(x, C_0) = \sup_{n\geq 1} d_\rho(x, C_{\alpha_n}) = d.$$

By virtue of Theorem 6.3, there exists a unique $c_0 \in C_0$ such that $d_\rho(x, C_0) = \rho(x - c_0)$. We claim that $c_0 \in C_\alpha$, for any $\alpha \in \Gamma$. Indeed, fix $\alpha \in \Gamma$. If for some $n \geq 1$ we have $\alpha \prec \alpha_n$, then obviously we have $c_0 \in C_{\alpha_n} \subset C_\alpha$. Therefore, we may assume that $\alpha \not\prec \alpha_n$, for any $n \geq 1$. Since Γ is upward directed, there exists $\beta_n \in \Gamma$ such that $\alpha_n \prec \beta_n$ and $\alpha \prec \beta_n$ for any $n \geq 1$. We can also assume that $\beta_n \prec \beta_{n+1}$ for any $n \geq 1$. Again we have $C_1 = \bigcap\limits_{n \geq 1} C_{\beta_n} \neq \emptyset$. Since $C_{\beta_n} \subset C_{\alpha_n}$, for any $n \geq 1$, we get $C_1 \subset C_0$. Moreover we have

$$d = d_\rho(x, C_0) \leq d_\rho(x, C_1) = \sup_{n \geq 1} d_\rho(x, C_{\beta_n}) \leq d,$$

which implies $d_\rho(x, C_1) = d$. Theorem 6.3 again implies that there exists a unique point $c_1 \in C_1$ such that $d_\rho(x, C_1) = \rho(x - c_1) = d$. Since C_0 is a Čebyšev subset, we get $c_0 = c_1$. In particular, we have $c_0 \in C_{\beta_n}$, for any $n \geq 1$. Since $\alpha \prec \beta_n$, we conclude that $C_{\beta_n} \subset C_\alpha$, for any $n \geq 1$, which implies $c_0 \in C_\alpha$. Since α was taking arbitrary in Γ, we get $c_0 \in \bigcap\limits_{\alpha \in \Gamma} C_\alpha$, i.e., $\bigcap\limits_{\alpha \in \Gamma} C_\alpha$ is not empty.

□

The conclusions of the above theorems will still hold under weaker assumptions.

Proposition 6.2. *Let C be a nonempty ρ-closed ρ-bounded convex subset of $\ell_{p(\cdot)}$.*

(i) Assume that C satisfies the property (R) and let K be a nonempty ρ-closed convex subset of C. Then K is ρ-proximinal in C, i.e. for any $x \in C$, the set $P_{\rho,K}(x) = \{y \in C;\ \rho(x - y) = \inf\limits_{z \in K} \rho(x - z)\}$ is not empty. Moreover if ρ is (SC), then K is a Čebyšev subset, i.e. $P_{\rho,K}(x)$ is a singleton for any $x \in C$.

(ii) Assume that C satisfies the property (R) and that ρ is (SC). Then for any family $\{C_\alpha\}_{\alpha \in \Gamma}$ of ρ-closed, convex, nonempty subsets of C such that $\bigcap\limits_{\alpha \in \Gamma_f} C_\alpha \neq \emptyset$ for any finite subset $\Gamma_f \subset \Gamma$, we have $\bigcap\limits_{\alpha \in \Gamma} C_\alpha$ is nonempty.

(iii) Assume that C satisfies the property (R) and ρ is (UUCED). Let K be a nonempty, ρ-closed, convex subset of C. Then K has a unique ρ-Čebyšev center $x \in K$, i.e.,

$$\sup\{\rho(x - y);\ y \in K\} = \inf_{z \in K} \Big(\sup\{\rho(z - y);\ y \in K\}\Big).$$

Proof. Assume that C satisfies the property (R). Let K be a nonempty, ρ-closed, convex subset of C. For $x \in C$, we have $d_\rho(x,K) = \inf\{\rho(x-y);\ y \in K\} < \infty$, since C is ρ-bounded. Moreover, since

$$P_{\rho,K}(x) = \bigcap_{n\geq 1} B_\rho\Big(x, d_\rho(x,K) + 1/n\Big) \cap K,$$

where $B_\rho(x,r)$ is the ρ-ball centered at x with radius r, the property (R) implies that $P_{\rho,K}(x)$ is not empty. It is clear that if ρ is (SC), then $P_{\rho,K}(x)$ must consist of a single point, which completes the proof of (i).

In order to prove (ii), assume that C satisfies the property (R) and that ρ is (SC). Let $\{C_\alpha\}_{\alpha\in\Gamma}$ be a family of ρ-closed, convex, nonempty subsets of C such that $\bigcap\limits_{\alpha\in\Gamma_f} C_\alpha$ is not empty, for any finite subset $\Gamma_f \subset \Gamma$. Let $x \in C$. Then $\sup\limits_{\alpha\in\Gamma} d_\rho(x,C_\alpha) < \infty$ holds since C is ρ-bounded. For any subset $F \subset \Gamma$, set $d_F = d_\rho(x, \bigcap\limits_{\alpha\in F} C_\alpha)$. Note that if $F_1 \subset F_2 \subset \Gamma$ are finite subsets, then $d_{F_1} \leq d_{F_2}$. Set

$$d_\Gamma = \sup\Big\{d_\rho\Big(x, \bigcap_{\alpha\in J} C_\alpha\Big),\ J \subset \Gamma \ \textit{such that} \ \bigcap_{\alpha\in J} C_\alpha \neq \emptyset\Big\}.$$

For any $n \geq 1$, there exists a subset $F_n \subset \Gamma$ such that $d_\Gamma - 1/n < d_{F_n} \leq d_\Gamma$. Set $F_n^* = F_1 \cup \cdots \cup F_n$, for $n \geq 1$. Then $\Big\{\bigcap\limits_{\alpha\in F_n^*} C_\alpha\Big\}$ is a decreasing sequence of nonempty ρ-closed convex subsets of C. The property (R) satisfied by C implies that $\bigcap\limits_{\alpha\in J} C_\alpha \neq \emptyset$, where $J = \bigcup\limits_{n\geq 1} F_n^* = \bigcup\limits_{n\geq 1} F_n$. Set $K = \bigcap\limits_{\alpha\in J} C_\alpha$. Note that $d_\rho(x,K) = d_\Gamma$, because $d_\Gamma - 1/n < d_{F_n} \leq d_\rho(x,K) \leq d_\Gamma$, for any $n \geq 1$. Because of (i), there exists a unique $y \in K$ for which $\rho(x-y) = d_\rho(x,K) = d_\Gamma$. For fixed $\alpha_0 \in \Gamma$, one has:

$$K \cap C_{\alpha_0} = \bigcap_{\alpha\in J\cup\{\alpha_0\}} C_\alpha \neq \emptyset$$

because of the same argument based on the property (R). Consequently $d_\rho(x,K) \leq d_\rho(x, K\cap C_{\alpha_0}) \leq d_\Gamma$. Hence $d_\rho(x, K\cap C_{\alpha_0}) = d_\rho(x,K) = d_\Gamma$, which implies $y \in K \cap C_{\alpha_0}$. Therefore, we have $y \in \bigcap\limits_{\alpha\in\Gamma} C_\alpha$, which proves that the family $\{C_\alpha\}_{\alpha\in\Gamma}$ has nonempty intersection.

In order to prove (iii), assume that C satisfies the property (R) and that ρ is $(UUCED)$. Let K be a nonempty, ρ-closed, convex subset of C. Set $r_\rho(x,K) = \sup\limits_{y\in K} \rho(x-y)$, for any $x \in K$, and

$$R_\rho(K) = \inf\{r_\rho(x,K); x \in K\}.$$

All the above numbers are finite, since C is ρ-bounded. Note that the set

$$K_n = \{x \in K;\ r_\rho(x,K) \leq R_\rho(K) + 1/n\},$$

is non-empty for any $n \geq 1$ and that

$$K_n = \bigcap_{y \in K} B_\rho\Big(y, R_\rho(K) + 1/n\Big) \cap K,\ n \geq 1,$$

which shows that $\{K_n\}$ is a decreasing sequence of ρ-closed, convex and nonempty subsets of K. The property (R) implies that $\bigcap_{n\geq 1} K_n$ is nonempty. Clearly, any $x \in \bigcap_{n\geq 1} K_n$ will satisfy $r_\rho(x,K) = R_\rho(K)$. We next set out to show that $\bigcap_{n\geq 1} K_n$ consists of a single point. Assume that there exists $z \in K$ such that $z \neq x$ and $r_\rho(z,K) = R_\rho(K)$. Since $\rho(x-z) \leq r_\rho(x,K) = R_\rho(K)$ and $x \neq z$, we conclude that $R_\rho(K) > 0$. Since ρ is $(UUCED)$, there exists $\delta = \delta(x,z,R_\rho(K)) > 0$ such that

$$\begin{cases} \rho(y-x) & \leq & R_\rho(K) \\ \rho(y-z) & \leq & R_\rho(K) \end{cases} \implies \rho\left(y - \frac{x+z}{2}\right) \leq R_\rho(K)(1-\delta),$$

for any $y \in K$. The latter implies

$$R_\rho(K) \leq r_\rho\left(\frac{x+z}{2}, K\right) \leq R_\rho(K)(1-\delta).$$

This contradiction finishes the proof of (iii), which in turn completes the proof of Proposition 6.2.

□

We aim at utilizing the above ideas for proving an analogue to Kirk's fixed point theorem [14], in $\ell_{p(\cdot)}$. Since this classical theorem uses the normal structure property, the following definition is needed:

Definition 6.8. $\ell_{p(\cdot)}$ is said to have the ρ-normal-structure property if for any nonempty, ρ-closed, convex and ρ-bounded subset C of $\ell_{p(\cdot)}$ that contains more than one point, there exists $x \in C$ such that

$$\sup_{y \in C} \rho(x-y) < \delta_\rho(C).$$

Theorem 6.6. *[2] Assume that* $\inf_{n\in\mathbb{N}} p(n) > 1$. *Then* $\ell_{p(\cdot)}$ *has the* ρ*-normal structure property.*

Proof. Since $\inf_{n\in\mathbb{N}} p(n) > 1$, Theorem 6.2 implies that ρ is $(UUC2)$. Let C be a ρ-closed, convex and ρ-bounded subset of $\ell_{p(\cdot)}$ not consisting of a single one point. It follows that $\delta_\rho(C) > 0$. Set $R = \delta_\rho(C)$. Let $x, y \in C$ such that $x \neq y$. Hence $\rho((x-y)/2) = \varepsilon > 0$. For any $c \in C$, we have $\rho(x-c) \leq R$ and $\rho(y-c) \leq R$. Hence

$$\rho\left(\frac{x+y}{2} - c\right) = \rho\left(\frac{(x-c)+(y-c)}{2}\right) \leq R\ \left(1 - \delta_2\left(R, \frac{\varepsilon}{R}\right)\right),$$

for any $c \in C$. Thus,

$$\sup_{c\in C} \rho\left(\frac{x+y}{2} - c\right) \leq R\ \left(1 - \delta_2\left(R, \frac{\varepsilon}{R}\right)\right) < R = \delta_\rho(C).$$

This completes the proof of Theorem 6.6, since C is convex.

□

Definition 6.9. Consider a nonempty set $C \subset \ell_{p(\cdot)}$ and a mapping $T : C \to C$. T is said to be ρ-Lipschitzian if for some constant $K \geq 0$ one has

$$\rho(T(x) - T(y)) \leq K\ \rho(x-y), \quad \text{for any } x, y \in C.$$

In particular T is called ρ-nonexpansive if $K = 1$ and $x \in C$ is called a fixed point of T if $T(x) = x$. The collection of all fixed points of T will be denoted by $Fix(T)$.

Theorem 6.7. *[2] Assume that* $\inf_{n\in\mathbb{N}} p(n) > 1$. *Let C be a nonempty, ρ-closed, convex and ρ-bounded subset of $\ell_{p(\cdot)}$. Let $T : C \to C$ be a ρ-nonexpansive mapping. Then T has a fixed point.*

Proof. Without loss of generality, it can be assumed that C is not a singleton. Consider the family

$$\mathscr{F} = \{K \subset C;\ K \text{ is nonempty, } \rho\text{-closed, convex and } T(K) \subset K\}.$$

The family $\mathscr{F}$ is not empty since $C \in \mathscr{F}$. Furthermore, $\inf_{n\in\mathbb{N}} p(n) > 1$, ρ is $(UUC2)$. Using Proposition 6.2 combined with Zorn's lemma, we conclude that $\mathscr{F}$ has a minimal element K_0. We claim that K_0 consists of exactly one point. Assume not, i.e., assume that K_0 has more than one point. Set $co(T(K_0))$ to be the intersection of all ρ-closed, convex subsets of C that contain $T(K_0)$. Hence $co(T(K_0)) \subset K_0$ since $T(K_0) \subset K_0$. So we have $T\Big(co(T(K_0))\Big) \subset T(K_0) \subset co(T(K_0))$. The minimality of K_0

implies $K_0 = co(T(K_0))$. Next, we use Theorem 6.6 to secure the existence of $x_0 \in K_0$ such that

$$r_0 = \sup_{y \in K_0} \rho(x_0 - y) < \delta_\rho(K_0).$$

Define the subset $K = \{x \in K_0;\ \sup_{y \in K_0} \rho(x-y) \leq r_0\}$. K is not empty, since $x_0 \in K$. Note that $K = \bigcap_{y \in K_0} B_\rho(y, r_0) \cap K_0$, where $B_\rho(y, r_0) = \{z \in \ell_{p(\cdot)};\ \rho(y - z) \leq r_0\}$. Since ρ satisfies the Fatou property and is convex, $B_\rho(y, r_0)$ is ρ-closed and convex. Hence K is a ρ-closed and convex subset of K_0. Let us show that $T(K) \subset K$. Let $x \in K$, then $T(x) \in \bigcap_{y \in K_0} B_\rho(T(y), r_0) \cap K_0$ since T is ρ-nonexpansive. Hence $T(K_0) \subset B_\rho(T(x), r_0)$, which implies $K_0 = co(T(K_0)) \subset B_\rho(T(x), r_0)$, i.e., $T(x) \in \bigcap_{y \in K_0} B_\rho(y, r_0) \cap K_0$. Therefore, $T(K) \subset K$ holds. The minimality of K_0 implies $K = K_0$, i.e., for any $x \in K_0$, we have $\sup_{y \in K_0} \rho(x-y) \leq r_0$. This clearly implies $\rho(x-y) \leq r_0$, for any $x, y \in K_0$. Hence $\delta_\rho(K_0) \leq r_0$. This is our sought contradiction. Therefore, K_0 hast exactly one point. Since $T(K_0) \subset K_0$, we conclude that T has a fixed point in C.

□

The following modular version of Kirk's fixed point theorem is an improvement to Theorem 6.7 since it does not require that $\inf_{n \in \mathbb{N}} p(n) > 1$.

Theorem 6.8. *[3] Assume that $p(\cdot)$ satisfies the condition (AO). Let C be a nonempty, ρ-closed, convex, ρ-bounded subset of $\ell_{p(\cdot)}$, which has the property (R). Let $T : C \to C$ be a ρ-nonexpansive mapping. Then $Fix(T)$ is a nonempty, ρ-closed and convex subset of C.*

Proof. Let $\emptyset \neq C \subset \ell_{p(\cdot)}$ be ρ-closed, convex and ρ-bounded and consider $T : C \to C$ to be a ρ-nonexpansive mapping. Assume that C is not a singleton: it is clear that no generality is lost under this assumption. Consider the family

$$\mathscr{F} = \{K \subset C; K \neq \emptyset,\ K\ \rho\text{-closed, convex and } T(K) \subset K\}.$$

The family $\mathscr{F}$ is not empty, since $C \in \mathscr{F}$. Our assumption on $p(\cdot)$ implies that ρ is $(UUCED)$. Zorn's Lemma in concert with Proposition 6.2 yields the existence of a minimal element of $\mathscr{F}$, which we denote by K_0. We contend that K_0 consists of exactly one point. For otherwise we set $co(T(K_0))$ to be the intersection of all ρ-closed, convex subsets of C that contain $T(K_0)$.

In particular, $co(T(K_0)) \subset K_0$ since $T(K_0) \subset K_0$ and we readily conclude that

$$T\Big(co(T(K_0))\Big) \subset T(K_0) \subset co(T(K_0)).$$

The minimality of K_0 yields $K_0 = co(T(K_0))$. Next, set $x \in K_0$ to be the unique ρ-Čebyšev center of K_0, i.e.,

$$r_\rho(x, K_0) = \sup\{\rho(x-y);\ y \in K_0\} = \inf_{z\in K_0} \Big(\sup\{\rho(z-y);\ y \in K_0\}\Big),$$

which exists according to (iii) of Proposition 6.2. Since T is ρ-nonexpansive and $K_0 \subset B_\rho(x, r_\rho(x, K_0))$, we get $T(K_0) \subset B_\rho(T(x), r_\rho(x, K_0))$. Hence $K_0 = co(T(K_0)) \subset B_\rho(T(x), r_\rho(x, K_0))$, since the ρ-balls are ρ-closed and convex. Thus we have

$$\begin{array}{rcl} r_\rho(T(x), K_0) & \leq & r_\rho(x, K_0) \\ & = & \sup\{\rho(x-y);\ y \in K_0\} \\ & = & \inf\limits_{z\in K_0} \Big(\sup\{\rho(z-y);\ y \in K_0\}\Big), \end{array}$$

i.e., $T(x)$ is also ρ-Čebyšev center of K_0. Therefore we must have $T(x) = x$, which implies $K_0 = \{x\}$; this is a contradiction to our assumption that K_0 contains more than one point. Hence any minimal element of $\mathscr{F}$ consists of exactly one point, which shows that $Fix(T)$ is not empty. Since T is ρ-nonexpansive, $Fix(T)$ is ρ-closed. Let us show that $Fix(T)$ is convex. Let $z_1, z_2 \in Fix(T)$ with $z_1 \neq z_2$. Let $\alpha \in [0,1]$. Then

$$\begin{array}{rcl} \rho(z_i - T(\alpha\, z_1 + (1-\alpha)z_2)) & = & \rho(T(z_i) - T(\alpha\, z_1 + (1-\alpha)z_2)) \\ & \leq & \rho(z_i - (\alpha\, z_1 + (1-\alpha)z_2)), \end{array}$$

for $i = 1,2$. Since ρ is (UUCED), it follows that ρ is (SC). Hence $T(\alpha\, z_1 + (1-\alpha)z_2) = \alpha\, z_1 + (1-\alpha)z_2$, i.e., $\alpha\, z_1 + (1-\alpha)z_2 \in Fix(T)$, which completes the proof of Theorem 6.8.

□

As a consequence of the properties of the fixed point set proved in Theorem 6.8, we present the following common fixed point result.

Theorem 6.9. *[12, 28, 65] Assume that $p(\cdot)$ satisfies the condition (AO). Let C be a nonempty, ρ-closed, convex, ρ-bounded subset of $\ell_{p(\cdot)}$ which satisfies the property (R). Let $\{T_\alpha\}_{\alpha\in\Gamma}$ be a commutative family of ρ-nonexpansive self-mappings defined on C. Then $\bigcap\limits_{\alpha\in\Gamma} Fix(T_\alpha)$ is a nonempty ρ-closed convex subset of C.*

Proof. Let $S, T : C \to C$ be two commutative ρ-nonexpansive mappings. Theorem 6.8 implies that $Fix(T)$ is a nonempty, ρ-closed, convex subset of C. Since S and T commute, it follows that $S(Fix(T)) \subset Fix(T)$. A further application of Theorem 6.8 yields that the restriction of S to $Fix(T)$ has a fixed point, i.e., $Fix(T) \cap Fix(S) \neq \emptyset$. It follows from this argument that for any finite subset Γ_f of Γ, $\bigcap_{\alpha \in \Gamma_f} Fix(T_\alpha)$ is a nonempty, ρ-closed, convex subset of C. Since C satisfies the property (R), we conclude that $\bigcap_{\alpha \in \Gamma} Fix(T_\alpha) \neq \emptyset$. The fact that this intersection is ρ-closed and convex follows easily.

$\square$

Bibliography

[1] A. A. N. Abdou, M. A. Khamsi, *Fixed point theorems in modular vector spaces*, J. Nonlinear Sci. Appl. **10** (2017), 4046-4057.

[2] M. Bachar, M. Bounkhel, M. A. Khamsi, *Uniform Convexity in $\ell_{p(.)}$*, J. Nonlinear Sci. Appl., **10** (2017), 5292–5299.

[3] M. Bachar, M. A. Khamsi, O. Méndez, M. Bounkhel, *A geometric property in $\ell_{p(\cdot)}$ and its applications*, Mathematische Nachrichten, **292:9** (2019), 1931-1940.

[4] B. Beauzamy, *Introduction to Banach Spaces and Their Geometry*, Elsevier, North-Holland, Amsterdam (1985).

[5] J. A. Clarkson, *Uniformly Convex Spaces*, Trans. Amer. Math. Soc, **40** (1936), 396-414.

[6] R. E. DeMarr, *Common fixed-points for commuting contraction mappings*, Pacific. J. Math. **13** (1963), 1139-1141.

[7] L. Diening, P. Harjulehto, P. Hästö, M., Ruẑiĉka, *Lebesgue and Sobolev Spaces with Variable Exponents*, Lecture Note in Mathematics **2017**, Springer, Berlin (2011).

[8] J. Diestel, *Geometry of Banach spaces - Selected Topics*, Springer Lecture Notes in Math. **485**, Springer, Berlin, New York (1957).

[9] M. A. Khamsi, *On metric spaces with uniform normal structure*, Proc. Amer. Math. Soc. **106** (1989), 723-726.

[10] M. A. Khamsi, W. A. Kirk, *An Introduction to Metric Spaces and Fixed Point Theory*, John Wiley, New York (2001).

[11] M. A. Khamsi, W. K. Kozlowski, S. Reich, *Fixed point theory in modular function spaces*, Nonlinear Analysis, **14** (1990), 935-953.

[12] M. A. Khamsi, W. M. Kozlowski, *Fixed Point Theory in Modular Function Spaces*, Birkhauser, New York (2015).

[13] S. J. Kilmer, W. M. Kozlowski, G. Lewicki, *Sigma order continuity and best approximants in L_ρ spaces*, Commentationes Mathematicae Universitatis Carolinae, **32:2** (1991), 241-250.

[14] W. A. Kirk, *A fixed point theorem for mappings which do not increase distances*, Amer. Math. Monthly, **72** (1965), 1004-1006.

[15] V. Klee, *Summability in $\ell(p_{11}, p_{21}, \cdots)$ Spaces*, Studia Math., **25** (1965), 277-280.

[16] O. Kováčik, J. Rákosník, *On spaces $L^{p(x)}$ and $W^{1,p(x)}$*, Czechoslovak Math. J., **41** (1991), 592-618.

[17] W. M. Kozlowski, *Modular Function Spaces*, Series of Monographs and Textbooks in Pure and Applied Mathematics, **122**, Dekker, New York/Basel (1988).

[18] W. M. Kozlowski, G. Lewicki, *Analyticity and polynomial approximation in modular function spaces*, J. Approximation Theory **58**(1) (1989), 15-35.

[19] T. C. Lim, *A fixed point theorem for families of nonexpansive mappings*, Pacific J. Math., **53** (1974), 487-493.

[20] J. Musielak, *Orlicz spaces and modular spaces*, Lecture Notes in Mathematics, Vol.1034, Springer-Verlag, Berlin/Heidelberg/New York/Tokyo (1983).

[21] H. Nakano, *Modulared Semi-ordered Linear Spaces*, Maruzen Co., Tokyo (1950).

[22] H. Nakano, *Topology of linear topological spaces*, Maruzen Co. Ltd., Tokyo (1951).

[23] H. Nakano, *Modulared sequence spaces*, Proc. Japan Acad., **27** (1951), 508-512.

[24] W. Orlicz, *”Uber konjugierte Exponentenfolgen*, Studia Math., **3** (1931), 200-211.

[25] K. Rajagopal, M. Ružička, *On the modeling of electrorheological materials*, Mech. Research Comm., **23** (1996), 401-407.

[26] M. Ružička, *Electrorheological fluids: modeling and mathematical theory*, Lecture Notes in Mathematics **1748**, Springer-Verlag, Berlin (2000).

[27] K. Sundaresan, *Uniform convexity of Banach spaces* $\ell(\{p_i\})$, Studia Mathematica, **39** (1971), 227-231.

[28] D. Waterman, T. Ito, F. Barber, J. Ratti, *Reflexivity and Summability: The Nakano* $\ell(p_i)$ *spaces*, Studia Math., **331** (1969), 141-146.

Chapter 7

Variational inequalities under the global NPC condition

P. Chaipunya

Department of Mathematics & Center of Excellence in Theoretical and Computational Science,
King Mongkut's University of Technology Thonburi,
126 Pracha Uthit Road, Bang Mond, Thung Khru, Bangkok 10140, Thailand
E-mail: parin.cha@mail.kmutt.ac.th

The aim of this Chapter is to provide a concise summary of recent convex-analytical results obtained under the global NPC condition, used in convex optimization problems and also to study their extensions from convex optimization to variational inequalities and further to equilibrium problems. It will be seen that both the variational inequality and the Minty variational inequality of the subdifferential provide sufficient and necessary optimality conditions wherever a manifold structure is required in the analysis. It will also be clear that the induced variational inequalities are never convex in any variable unless the space has zero curvature. To overcome this difficulty, a method will be introduced to solve a class of nonconvex equilibrium problems. This method is applicable to the treatment of variational inequalities and of convex optimization problems, under appropriate conditions on the space.

7.1 Introduction

Optimization in normed linear spaces, especially in Hilbert spaces, is no doubt a very fruitful area. In view of several emergent applications including computational biology, medical DTI/MRI image processing, manifold-valued data processing, and manifold learning, the need to develop the theory of optimization, convex analysis, and optimization algorithms outside a linear space is apparent. The first step of such extension is to consider optimization on Riemannian or Hadamard manifolds. In such manifolds, one exploits the available calculus defined through the smooth structure and takes advantage of several identification processes using diffeomorphisms, charts, connections, and so on. We refer the reader to [26, 1] for a nice survey on this topic.

The starting point of the present investigation is to consider metric spaces with a synthetic curvature property - namely, the global NPC condition. This notion directly generalizes both that of Hilbert space and that of a Hadamard manifold. The main idea in this approach is to extract metrical properties from the analytical and differential geometrical properties of a good model space (often a manifold). The ability to consider and solve optimization problems as well as other variational problems in this general setting will allow the study of more applications and the consideration of spaces without a smooth manifold structure, such as metric trees, fractals, singular surfaces, nonsmooth configuration spaces, nonsmooth statistical manifolds, etc.. In addition, this general setting allows us to neglect the Riemannian structure which seems superfluous when the problem at hand has to be dealt with in a nonsmooth manner. This topic was first independently introduced in [13] and [24], where the proximal operator was first introduced in the globally NPC space. It took quite a long time until unconstrained optimization problems in globally NPC spaces were considered in [4]. In [2, 3], the concept of a dual space to a globally NPC space was discussed for the first time. In [10] and [14] the authors indepently developed some theory of monotone operators in globally NPC spaces based on the use of dual space of [3, 2] and showed the convergence of the corresponding proximal algorithm to a stationary point. The results obtained in [10] and [14] clearly improve the existing theory in Hilbert spaces. However it is still unknown and puzzling how these problems, or even the dual space itself, are related to the classical Hadamard manifold theory developed earlier in [12, 21, 27]. In a recent preprint [8] the authors have developed a new approach to study the monotone counterpart of convex functions, using the concept of tangent cones due to [25]. In the same paper the subdifferential of convex functions with the tangent cone approach is considered and the authors deduce the first-order optimality condition for an unconstrained optimization problem.

It is quite natural to go beyond the above studies by introducing constraints to the problems. Suppose that (X, d) is a globally NPC space. Throughout this Chapter, we will consider the convex optimization problem (P) with a closed convex constraint set $\Omega \subset X$, as follows:

$$\min\ \varphi(x) \quad \textit{such that}\ \ x \in \Omega, \tag{P}$$

where $\varphi : X \to (-\infty, +\infty]$ is convex. It is typical to link the first-order properties like the directional derivative $D\varphi$ or the subdifferential $\partial\varphi$, to the optimalities of φ. As is known in other settings, the first-order optimality condition for such optimization problem is given in the form of a variational inequality. In what follows, the motivation from such optimality condition will be used and extended to a general variational inequality with properties inherited by the subdifferential as a vector field. In so doing it should be noticed that the resulting variational inequality lacks the convexity (or quasi-convexity) expected in linear spaces. In fact, the convexity holds if and only if the space has identically zero curvature. In other words, the curvature impedes a simple generalization. This difficulty can be overcome by moving to a higher level of generality and by regarding the variational inequality as an equilibrium problem to highlight the properties that each variable precisely contributes. In fact, the general equilibrium problem in globally NPC spaces was investigated in [9, 15, 20] under convexity (and generalized convexity) assumptions on the second argument. Due to such convexity assumptions, the results in the above works are not applicable to the situation in point. Instead, we come up with a new type of continuity, tailored for a bifunction, that captures the situation and that is sufficient to ensure the convergence of the proximal algorithm at hand. To conclude, new notions to solve non-convex equilibrium problems are developed, which are applicable to variational inequalities and constrained optimization problems at different level of generalities - ranging from general globally NPC spaces to the level of an Hadamard manifold.

This Chapter is organized as follows. Section 7.2 recalls the necessary concepts needed in the sequel. It includes the definitions of the global NPC condition, of convexity of sets and of functions, the definition of convergence and of tangent cones. Section 7.3 is devoted to observations on the first-order properties of convex functions. The motivation stemming from variational inequalities as well as from the Minty variational inequalities will be apparent. In Section 7.4, the target variational inequality is formulated. Such formulation is strongly inspired by the optimality condition in Section 7.3. In Section 7.4, the equilibrium problems are posed. In Section 7.5, the idea of a resolvent operator for a bifunction is developed and its fundamental properties are proved. This resolvent operator is crucial for the proximal algorithm used to find the solution of the equilibrium problem. Finally, in Section 7.6, the convergence of proximal algorithms defined by iterating the resolvent operators with different parameters is proved. The results for nonconvex equilibrium problems are then applied to variational inequalities and finally to convex optimization problems.

7.2 Global NPC spaces

This Section is devoted to a very brief overview of the global NPC condition and to a discussion of related properties. Since an exhaustive discussion of the geometry of globally NPC spaces is beyond the scope of our analysis, the information given in this Section is the minimum required for a transparent exposition. Most of the results given in this Section are covered in [5] and [6, 7]. The reader is also referred to [17, 18] for a further discussion on the space of directions and tangent cones. It is assumed that $\mathbb{R}$ and $\mathbb{N}$ stand for the sets of all reals and all positive integers, respectively. Moreover, $\mathbb{R}^n$ will always be assumed to be equipped with its Euclidean norm $\|\cdot\|$.

7.2.1 The global NPC condition

Let (X,d) be a metric space. X is said to be geodesic if any two points $x,y \in X$ can be connected by a curve $\sigma : [0,d(x,y)] \to X$ with $\sigma(0) = x$, $\sigma(d(x,y)) = y$, and $d(\sigma(s),\sigma(t)) = |s-t|$ for all $s,t \in [0,d(x,y)]$. Here, σ is called a *minimizing geodesic* joining x to y. If each pair $x,y \in X$ is connected by a unique minimizing geodesic, X is said to be *uniquely geodesic*. In this case, the symbol $[x,y] := \sigma([0,d(x,y)])$, where σ is the unique geodesic connecting x to y, denotes the corresponding *geodesic segment*. Moreover, we adopt the notation $(1-t)x \oplus ty := \sigma(td(x,y))$ for $t \in [0,1]$. If σ is the minimizing geodesic joining x and y, then $|\sigma| := d(x,y)$ is called the *length* of σ.

As is the case in a Riemannian manifold, the metric (induced from the Riemannian structure) properties alone are not sufficient to obtain strong results. Thus this discussion is restricted to the case when the sectional curvature is bounded either from above or from below. In the sequel the *synthetic* curvature condition is considered, which corresponds to global nonpositive sectional curvature in Riemannian manifolds, on geodesic metric spaces. Note that this concept of curvature can be defined merely in terms of the distance function, but this will not be so useful in any delicate analysis. To fully exploit this notion, the definition of a *comparison triangle* has to be recalled. Since the full generality of this idea is not needed, we rather adopt the notions that are more natural than the more general ones.

Definition 7.1. Suppose that (X,d) is a uniquely geodesic metric space. Let $x,y,z \in X$ and consider the *geodesic triangle* $\Delta \equiv \Delta(x,y,z) = [x,y] \cup [y,z] \cup [x,z]$. A triangle $\overline{\Delta} \equiv \Delta(\bar{x},\bar{y},\bar{z})$ in $\mathbb{R}^2$ is a *comparison triangle* of Δ if $d(x,y) = \|\bar{x}-\bar{y}\|$, $d(y,z) = \|\bar{y}-\bar{z}\|$, and $d(x,z) = \|\bar{x}-\bar{z}\|$.

Remark 7.1. Notice that by virtue of the triangle inequality, a comparison triangle always exists and it is unique up to rigid motions, i.e., up to translations and rotations. Also, the triangles above can be degenerated.

Definition 7.2. Suppose that (X,d) is a uniquely geodesic metric space, $x,y,z \in X$, and $\Delta(\bar{x},\bar{y},\bar{z})$ is a comparison triangle of $\Delta(x,y,z)$. Let $u \in [x,y]$. The point $\bar{u} \in [\bar{x},\bar{y}]$ is called the *comparison point* of u if $d(x,u) = \|\bar{x}-\bar{u}\|$ and $d(u,y) = \|\bar{u}-\bar{y}\|$. Comparison points on $[y,z]$ and $[x,z]$ are defined likewise.

Next, the formal definition of a globally NPC space is stated.

Definition 7.3. A metric space (X,d) is said to be *globally nonpositively curved* (or *globally NPC*), if it is uniquely geodesic and if for any given $x,y,z \in X$ and any $u \in \Delta(x,y,z)$, the following inequality is satisfied:

$$d(x,u) \leq \|\bar{x}-\bar{u}\|,$$

where $\bar{u} \in \overline{\Delta}$ is the comparison point of u.

Remark 7.2. A globally NPC space is also called a $\mathrm{CAT}(0)$ *space*, following the terminology of Gromov. Moreover, a complete globally NPC space is known by the terminology *Hadamard space*. In this Chapter, the term *globally NPC* is opted for, as it best depicts the geometric situation of the space.

There are several equivalent statements for (X,d) being globally NPC. For future reference, the following list is given:

Proposition 7.1. *Let (X,d) be a uniquely geodesic metric space. Then, the following statements are equivalent:*

(1) X is globally NPC.

(2) For any $x,y,z \in X$, the following inequality holds true:

$$d^2(x,(1-t)y \oplus tz) \leq (1-t)d^2(x,y) + td^2(x,z) - t(1-t)d^2(y,z), \tag{7.1}$$

for any choice of $t \in [0,1]$.

(3) For any geodesic triangle Δ in X and $u,v \in \Delta$, the following inequality holds true:

$$d(u,v) \leq \|\bar{u}-\bar{v}\|,$$

where $\bar{u}$ and $\bar{v}$ are the comparison points of u and v, respectively.

(4) For any $x,y,u,v \in X$, the following inequality holds:

$$d^2(x,v) + d^2(y,u) \leq d^2(x,u) + d^2(y,v) + 2d(x,y)d(u,v). \tag{7.2}$$

The class of globally NPC spaces includes Hilbert spaces, metric trees and Gromov-Hausdorff limits of Riemannian manifolds of nonpositive sectional curvature. It is useful to note that if (X_i, d_i), $i = 1, 2, \cdots, n$ are globally NPC spaces, then so is their product $\mathbf{X} := \prod_{i=1}^{n} X_i$, when equipped with the distance function $\mathbf{d}$ given by $\mathbf{d}(\mathbf{x}, \mathbf{y}) := (\sum_{i=1}^{n} d_i(x_i, y_i))^{1/2}$, where $\mathbf{x} = (x_1, \cdots, x_n)$ and $\mathbf{y} = (y_1, \cdots, y_n)$ are elements from $\mathbf{X}$. In the sequel, products of globally NPC spaces are always to be understood in this sense.

Now, suppose that (X, d) is a globally NPC space, $p \in X$, and take any two minimizing geodesics σ^1 and σ^2 with common emanating point $\sigma^1(0) = \sigma^2(0) = p$. Each geodesic triangle $\Delta(p, \sigma^1(s), \sigma^2(t))$ associates with it the comparison triangle $\Delta_{s,t}$ whose vertices $\bar{p}$, $\overline{\sigma^1(s)}$, and $\overline{\sigma^2(t)}$ correspond respectively to p, $\sigma^1(s)$, and $\sigma^2(t)$. Denote by $\bar{\angle}_{\bar{p}}(\sigma^1(s), \sigma^2(t))$ the angle at $\bar{p}$ of the triangle $\Delta_{s,t}$. By virtue of the global NPC condition on X, the angle $\bar{\angle}_{\bar{p}}(\sigma^1(s), \sigma^2(t))$ decreases with s and t. Therefore, the limit as s, t tend to 0^+ always exists; it will be denoted by $\alpha_p(\sigma^1, \sigma^2)$. As a function of two minimizing geodesics on X, $\alpha_p(\cdot, \cdot)$ is called the *Alexandrov angle* at p.

In the remainder of this paper, unless stated otherwise, it will be always assumed that (X, d) is a complete globally NPC space.

7.2.2 Convexity

The main ideas presented in this Section revolve around the concepts of convex sets and convex functions. A nonempty set $\Omega \subset X$ is called *convex* if $x, y \in \Omega$ implies $[x, y] \subset \Omega$. Likewise, a function $\varphi : X \to (-\infty, +\infty]$ is a *convex* function if its epigraph $\operatorname{epi}(\varphi) := \{(x, r) \in X \times \mathbb{R};\ \varphi(x) \leq r\}$ is a convex set. A simple calculation reveals that φ is convex if and only if the inequality

$$\varphi((1-\lambda)x \oplus \lambda y) \leq (1-\lambda)\varphi(x) + \lambda\varphi(y) \tag{7.3}$$

holds for any $x, y \in X$ and $\lambda \in [0, 1]$. Viewing a convex function as one that satisfies the above inequality can simplify the ideas. For instance, φ is called *strictly convex* if (7.3) holds with strict inequality for distinct points x and y, and it is called *strongly convex* if there is $\kappa > 0$ for which the inequality

$$\varphi((1-\lambda)x \oplus \lambda y) \leq (1-\lambda)\varphi(x) + \lambda\varphi(y) - \kappa\lambda(1-\lambda)d^2(x, y)$$

holds for all $x, y \in X$ and all $\lambda \in [0, 1]$. If φ is strongly convex and lsc, then it is bounded from below. If, additionally, the space X is complete, then φ has a unique minimizer. It is noticed right away that (7.1) actually states that $d^2(\cdot, y)$ is strongly convex with $\kappa = 1$, for any fixed $y \in X$. Fix a nonempty, closed, convex set $\Omega \subset X$ and $x \in X$. Then $d^2(\cdot, y)$ has a unique minimizer over Ω provided that Ω is complete. This leads to the very definition of the so-called *metric projection*.

Definition 7.4. Suppose that $\Omega \subset X$ is nonempty, closed and convex. The *metric projection* onto Ω is the mapping $P_\Omega : X \to \Omega$ given by

$$P_\Omega(x) := \operatorname{argmin}_{y\in\Omega} d(x,y) \quad ({}^\forall x \in X).$$

If $\varphi : X \to \mathbb{R}(-\infty,+\infty]$ is proper, convex, lsc, then its proximal operator $\operatorname{prox}_\varphi^\lambda : X \to X$ is well-defined for all $\lambda > 0$ and is given by

$$\operatorname{prox}_\varphi^\lambda(x) := \operatorname{argmin}_{y\in X}\left[\varphi(y) + \frac{1}{2\lambda}d^2(y,x)\right] \quad ({}^\forall x \in X).$$

This operator was independently introduced in [13] and [24] to study gradient flows and harmonic maps. It was recently used in [4] to define the proximal algorithm used in finding an unconstrained minimizer of a convex function.

7.2.3 Δ-convergence

Convergence in the metric topology is known to be irrelevant in infinite dimensional spaces. In a CAT(0) space, the concept of Δ-convergence can be defined, which coincides with the weak convergence when the space in question is a Hilbert space. Let $(x^k) \subset X$ be a bounded sequence and define a function $r(\cdot;(x^k)) : X \to [0,\infty)$ by

$$r(x;(x^k)) := \limsup_{k\to\infty} \rho(x,x^k), \quad \forall x \in X.$$

According to [11], the minimizer of this function exists and is unique. Following [16] (see also [23]), a bounded sequence (x^k) is said to be *Δ-convergent* to a point $\bar{x} \in X$ if $\bar{x} = \operatorname{argmin}_{x\in X} r(x;(u^k))$ for any subsequence (u^k) of (x^k). In this case, $\bar{x}$ is called the Δ-limit of (x^k). Recall that a bounded sequence is Δ-convergent to at most one point. A point $z \in X$ is said to be a Δ-accumulation point of a sequence (x^k) in X if (x^k) has a Δ-convergent subsequence whose Δ-limit point is z. Moreover, if (x^k) is a sequence in a closed convex set Ω, then its Δ-accumulation points are within Ω.

It is currently unknown whether Δ-convergence is equivalent to convergence with respect to a topology on X. This question was partially answered in [2], with additional assumptions.

In practice, it is often not simple to derive Δ-convergence directly from the definition. The following notion is an important tool in showing Δ-convergence of a given sequence. It also plays a vital role in the convergence analysis given in the final Section of this Chapter.

Definition 7.5. A sequence (x^k) in X is said to be *Fejér convergent* with respect to a nonempty set $\Omega \subset X$ if for each $x \in \Omega$, it holds that $d(x^{k+1},x) \leq d(x^k,x)$ for all large $k \in \mathbb{N}$.

Proposition 7.2 ([10]). *Suppose that (x^k) is a sequence in X which is Fejér convergent to a nonempty set $\Omega \subset X$. Then, the following are true:*

(1) (x^k) is bounded.

(2) $(d(x,x^k))$ converges for any $x \in \Omega$.

(3) If every Δ-accumulation point lies within Ω, then (x^k) is Δ-convergent to an element in Ω.

7.2.4 Space of directions and tangent cones

The ultimate goal of this Subsection is to introduce the notions of tangent cones and of intrinsic scalar products. These concepts are originally due to [25]. The tangent cone to a globally NPC space (X,d) can be defined in two equivalent ways, namely, as an Euclidean cone over the space of directions or as a certain limit of rescalings of the space X. The focus in this analysis is on the first approach. The metric calculation emanating from the latter approach will be mentioned only in passing.

Let $p \in X$ and G_p be the set of all geodesic emanating from p with nonzero length. Define an equivalence relation $\sim$ on G_p by

$$\sigma^1 \sim \sigma^2 \iff \alpha_p(\sigma^1,\sigma^2) = 0,$$

for $\sigma^1,\sigma^2 \in G_p$. Then the quotient metric space $(\Sigma_p X, \angle_p) := (G_p,\alpha_p)/\sim$ is called the *space of directions of X at p*. We shall write $\uparrow_p^x$ to denote the equivalence class of $\sim$ containing the minimizing geodesic joining p to x, for $x \in X \setminus \{p\}$. Next, the *logarithm map associated to* $\Sigma_p X$, denoted with $\log_{\Sigma_p X} : X \setminus \{p\} \to \Sigma_p X$, is defined by

$$\log_{\Sigma_p X}(x) = \uparrow_p^x \quad ({}^\forall x \in X \setminus \{p\}).$$

We then take the Euclidean cone $\mathrm{Cone}_p(\Sigma pX)$ defined by the quotient space $([0,\infty) \times \Sigma_p X)/\approx$, where $\approx$ is the equivalence class on $[0,\infty) \times \Sigma_p X$ defined as $(t_1,\uparrow^1) \sim (t_2,\uparrow^2)$ if either of the following holds:

(1) $t_1 = t_2 = 0$, or

(2) $t_1 = t_2 > 0$ and $\uparrow^1 = \uparrow^2$.

Set $T_pX := \mathrm{Cone}_p(\Sigma_p X)$ and define a metric d_p on T_pX by

$$d_p([(t_1,\uparrow^1)]_\approx,[(t_2,\uparrow^2)]_\approx) := \sqrt{t_1^2 + t_2^2 - 2t_1t_2\cos\angle_p(\uparrow^1,\uparrow^2)}$$

for each $[(t_1,\uparrow^1)]_\approx,[(t_2,\uparrow^2)]_\approx \in T_pX$. Denote by 0_p the equivalence class

$[(0,\uparrow)]_\approx \in T_pX$. The metric space (T_pX, d_p) is called the *tangent cone of X at p*. It must be noted that Σ_pX is a complete CAT(1) space and that T_pX is a complete globally NPC space. With the notion of a tangent cone, one defines $\log_{T_pX} : X \to T_pX$, the *logarithm map associated to T_pX*, given by

$$\log_{T_pX}(x) := \begin{cases} [(d(p,x), \uparrow_p^x)]_\approx & \text{for } x \neq p, \\ 0_p & \text{for } x = p. \end{cases}$$

It is customary to use the notation $t\log_{T_pX}(x) := [(td(p,x), \uparrow_p^x]_\approx$ for $x \neq p$. From the second definition of the tangent cone T_pX it follows that

$$d_p(\log_{T_pX}(x), \log_{T_pX}(y)) = \lim_{\varepsilon \to 0^+} \varepsilon^{-1} d(\sigma_p^x(\varepsilon), \sigma_p^y(\varepsilon)),$$

where σ_p^x and σ_p^y are respectively the minimizing geodesics joining p to x and y. It is natural to identify Σ_pX as the a subset $\{[(1,\uparrow_p)]_\approx \mid \uparrow_p \in \Sigma_pX\}$ in T_pX. Next, define the *scalar product* $< \cdot, \cdot >$ on T_pX by

$$< [(t_1, \uparrow^1)]_\approx, [(t_2, \uparrow^2)]_\approx > := t_1 t_2 \cos \angle_p(\uparrow^1, \uparrow^2)$$

for $[(t_1, \uparrow^1)]_\approx, [(t_2, \uparrow^2)]_\approx \in T_pX$.

A final remark is in order, namely that when X is a Hadamard manifold, X is locally compact, T_pX is isometric to the Riemannian tangent space Tan_pX and the map $\log_{T_pX}$ is identified via such isometry with the *inverse exponential map* $\exp_p^{-1} : X \to \mathrm{Tan}_pX$, which is a global diffeomorphism by the Cartan-Hadamard theorem. In this case, Tan_pX is isometric to the Euclidean space of the dimension of the manifold X. Moreover, Σ_pX is exactly the unit sphere in the tangent space Tan_pX. This fact yields the decomposition $T_{(p,q)}(X_1 \times X_2) = T_pX_1 \times T_qX_2$ if X_1 and X_2 are two Hadamard manifolds, $p \in X_1$, and $q \in X_2$.

7.3 First-order properties of convex functions

This Section delves into deeper considerations in the study of convex functions. Specifically, subdifferentials, directional derivatives, convex optimization and optimality conditions will be considered. The motivation for the introduction of such notions will be apparent towards the end, as variational inequalities are introduced.

For convenience, denote by $\Gamma_0(X)$ the family of all functions $\varphi : X \to (-\infty, +\infty]$ that are proper, convex, and lsc.

Firstly, the notion of a subdifferential $\partial\varphi$ will introduced and its connection with convex optimization problem emphasized. Note that even though $\partial\varphi$ can be defined directly from φ, another approach to the definition is presented here, aiming to highlight its geometric features. First, we define the notion of a *normal cone* to a convex set.

Definition 7.6. Let $\Omega \subset X$ be a nonempty, convex set and take $p \in \mathrm{cl}\Omega$. The *normal cone to* Ω *at* p is the set

$$N_\Omega(p) := \{\Uparrow_p \in T_pX \mid <\Uparrow_p, \log_{T_pX}(x)> \leq 0 \ \ (^\forall x \in \Omega)\}.$$

A normal cone $N_\Omega(p)$ is always nonempty, since it contains 0_p. Moreover, $N_\Omega(p)$ is trivial (containing only 0_p) if and only if p is an interior point of Ω.

Definition 7.7. Let $\varphi \in \Gamma_0(X)$ and $p \in \mathrm{dom}\varphi$. The *subdifferential of* φ *at* p is defined by

$$\partial\varphi(p) := \{\Uparrow_p \in T_pX \mid (\Uparrow_p, -1) \in N_{\mathrm{epi}\varphi}(p, \varphi(p))\}.$$

Elements in $\partial\varphi(p)$ are called *subgradients of* φ *at* p.

Proposition 7.3. *Let* $\varphi \in \Gamma_0(X)$ *and let* $p \in dom\varphi$. *Then,*

$$\partial\varphi(p) := \{\Uparrow_p \in T_pX \mid \varphi(y) \geq \varphi(p) + <\Uparrow_p, \log_{T_pX}(y)> \ \ (^\forall y \in X)\}.$$

Proof. Observe that

$$\begin{aligned}
\Uparrow_p \in \partial\varphi(p) &\iff (\Uparrow_p, -1) \in N_{\mathrm{epi}\varphi}(p, \varphi(p)) \\
&\iff <(\Uparrow_p, -1), \log_{T_{(p,\varphi(p))}X \times \mathbb{R}}(y, w)> \leq 0 \quad (^\forall (y, w) \in \mathrm{epi}\varphi) \\
&\iff <\Uparrow_p, \log_{T_pX}(y)> + \varphi(p) - w \leq 0 \quad (^\forall y \in \mathrm{dom}\varphi)(^\forall w \geq \varphi(y)) \\
&\iff <\Uparrow_p, \log_{T_pX}(y)> + \varphi(p) \leq \varphi(y) \quad (^\forall y \in \mathrm{dom}\varphi).
\end{aligned}$$

This completes the proof. □

It is apparent that $\partial\varphi(x) = \emptyset$ if $x \notin \mathrm{dom}\varphi$. In fact, $\mathrm{dom}\partial\varphi$ is dense in $\mathrm{dom}\varphi$. Notice that $\partial\varphi$ is tied the to convex optimization problem (P). Specifically, both sufficient and necessary optimality conditions, can be given for (P) in terms of a variational inequality. (7.4).

Theorem 7.1. *Suppose that* $\varphi \in \Gamma_0(X)$ *and that* $\Omega \subset int(dom\varphi)$ *is nonempty, closed, and convex. If there exist* $x^* \in \Omega$ *and* $\Uparrow^* \in \partial\varphi(x^*)$ *for which*

$$<\Uparrow^*, \log_{T_{x^*}X}(x)> \geq 0 \quad (^\forall x \in \Omega), \tag{7.4}$$

then x^* *is a solution of* (P).

Proof. By the definition of $\partial\varphi$ and on account of (7.4), it is immediately clear that

$$\varphi(x) \geq \varphi(x^*) + < \Uparrow^*, \log_{T_pX}(x) > \geq \varphi(x^*)$$

for any $x \in \Omega$. □

In order to develop necessary optimality conditions, the relationship between $D\varphi$ and $\partial\varphi$ has to be established. This will be facilitated by the notion of *directional derivative* and its fundamental properties. For a slightly different point of view on directional derivatives, see [18].

Definition 7.8. Let $\varphi \in \Gamma_0(X)$ be Lipschitz continuous, $p \in \mathrm{int}(\mathrm{dom}\varphi)$, and $(t, \uparrow_p) \in T_pX$. The *derivative of φ in the direction $(t, \uparrow_p)$* is defined by

$$D\varphi_p(t, \uparrow_p) := \lim_{\lambda \to 0^+} \frac{\varphi \circ \sigma(\lambda t) - \varphi(p)}{\lambda},$$

where $\sigma \in \uparrow_p$ has unit speed, provided that the limit exists. Here, the map $D\varphi_p(\cdot)$ is called the *directional derivative of φ at p*.

Proposition 7.4. *Let $\varphi \in \Gamma_0(X)$ be Lipschitz continuous and $p \in$ int(domφ). Then $D\varphi_p$ is well-defined on T_pX, and it is convex and positively homogeneous.*

Proof. The convexity guarantees that the quotient

$$\lambda \mapsto \frac{\varphi \circ \sigma(\lambda t) - \varphi(p)}{\lambda}$$

is well-defined and increasing for sufficiently small $\lambda > 0$. Therefore, the limit exists for any $t \geq 0$ and any unit speed minimizing geodesic σ emanating from p. Now, suppose that σ^1 and σ^2 are two unit speed minimizing geodesics emanating from a common point p, whose end points are x and y, respectively. For any $s, t \geq 0$, one has

$$\begin{aligned}
|s(\varphi \circ \sigma^1)'(0) - t(\varphi \circ \sigma^2)'(0)| &= \left| \lim_{\lambda \to 0^+} \frac{\varphi \circ \sigma^1(\lambda s) - \varphi \circ \sigma^2(\lambda t)}{\lambda} \right| \\
&\leq L \lim_{\lambda \to 0^+} \frac{d(\sigma^1(\lambda s) - \sigma^2(\lambda t))}{\lambda} \\
&= L d_p(s \log_{\Sigma_pX}(x), t \log_{\Sigma_pX}(y)),
\end{aligned}$$

where $L \geq 0$ is the Lipschitzian constant of φ. If σ^1 and σ^2 represent the same class $\uparrow_p \in \Sigma_pX$, then $\log_{\Sigma_pX}(x) = \log_{\Sigma_pX}(y)$ and the above inequality guarantees that $D\varphi_p$ is well-defined. The Lipschitz continuity of $D\varphi_p$ follows from the same inequality. Finally, the positive homogeneity is obvious from the limiting process and the convexity of $D\varphi_p$ is obtained from that of φ. □

Even though the relationship between $\partial\varphi$ and $D\varphi$ is quite clear in the linear setting, the full characterization between the two notions under the general global NPC condition is still unknown. In the following result, it is required that X be a Hadamard manifold in order to invoke the separation theorem on the tangent space. To the author's best knowledge, neither this technique nor the following result, have appeared before in the literature on convex analysis, outside the linear setting. Here, recall that $\text{epi}^{<}\varphi := \{(z,w) \in X \times \mathbb{R} \mid \varphi(z) < w\}$.

Proposition 7.5. *Let X be a Hadamard manifold and $\varphi \in \Gamma_0(X)$ be Lipschitz continuous. For any $p \in int(dom\varphi)$ and $\Uparrow_p \in T_pX$, it holds*

$$D\varphi_p(\Uparrow_p) = \max_{\xi \in \partial\varphi(p)} < \xi, \Uparrow_p > .$$

Proof. The inequality $D\varphi_p(\Uparrow_p) \geq \max_{\xi\in\partial\varphi(p)} < \xi, \Uparrow_p >$ was already proved in [22]. Thus, only the reverse inequality $D\varphi_p(\Uparrow_p) \leq \max_{\xi\in\partial\varphi(p)} < \xi, \Uparrow_p >$ remains to be proved. Without loss of generality, only the case $\Uparrow_p \neq 0$ needs to be considered, otherwise there is nothing to prove. Fix a representative (t,σ) of $\Uparrow_p$ such that the geodesic segment of σ lies in the interior of domφ and define

$$\Omega_1 := \left\{\lambda \log_{\Sigma_{(p,\varphi(p))}X\times\mathbb{R}}(z,w) \in T_{(p,\varphi(p))}X \times \mathbb{R} \mid (z,w) \in \text{epi}^{<}\varphi,\ \lambda > 0\right\}$$

and

$$\Omega_2 := \left\{ \log_{T_{(p,\varphi(p))}X\times\mathbb{R}}(y,v) \in T_{(p,\varphi(p))}X \times \mathbb{R} \;\middle|\; \begin{array}{l} y = \sigma(\lambda t), \\ v = \varphi(p) + \lambda D\varphi_p(\Uparrow_p), \\ 0 \leq \lambda \leq |\sigma|/t \end{array} \right\}.$$

Observe that the convexity of Ω_1 and Ω_2 follows, respectively, from the convexity of $\text{epi}^{<}\varphi$ and from the fact that Ω_2 is a line segment. Moreover, since $\varphi\circ\sigma(\lambda t) \geq \varphi(p) + \lambda D\varphi_p(\Uparrow_p)$ for all $0 \leq \lambda \leq |\sigma|/t$, the sets Ω_1 and Ω_2 are disjoint. By virtue of the Hahn-Banach separation theorem, there exists a nonzero vector $(\upsilon,\mu) \in (T_pX) \times \mathbb{R} = T_{(p,\varphi(p))}X \times \mathbb{R}$ such that

$$< \upsilon, \log_{T_pX} > (\sigma(\lambda t)) + \mu(\varphi(p) + \lambda D\varphi_p(\Uparrow_p)) \leq < \upsilon, \eta \log_{\Sigma_pX}(z) > + \mu w \tag{7.5}$$

for all $0 \leq \lambda \leq |\sigma|/t$, all $(z,w) \in \text{epi}^{<}\varphi$, and all $\eta > 0$. If $\mu < 0$, it follows that $w < \varphi(p) + \lambda D\varphi_p(\Uparrow_p))$, which is a contradiction. On the other hand, letting $\mu = 0$ one concludes that $< \upsilon, \log_{T_pX}(\sigma(\lambda t)) > \leq < \upsilon, \delta >$ for every $\delta \in \Sigma_pX$, since p is in the interior of domφ. Hence $\upsilon = 0$ and $(\upsilon,\mu) = 0$, which is also a contradiction. Therefore, $\mu > 0$. Dividing by μ in (7.5) and letting $\bar{\upsilon} := \upsilon/\mu$ it follows that

$$< \bar{\upsilon}, \log_{T_pX}(\sigma(\lambda t)) > + \varphi(p) + \lambda D\varphi_p(\Uparrow_p) \leq < \bar{\upsilon}, \eta \log_{\Sigma_pX}(z) > + w \tag{7.6}$$

for all $0 \leq \lambda \leq |\sigma|/t$, all $(z,w) \in \mathrm{epi}^{<}\varphi$, and all $\eta > 0$. Choosing $\lambda = 0$ and letting $w \to \varphi(z)^{+}$, one has

$$\varphi(p) \leq < \bar{\upsilon}, \eta \log_{\Sigma_p X}(z) > + \varphi(z) \quad (^{\forall} z \in \mathrm{dom}\varphi).$$

This shows that $-\bar{\upsilon} \in \partial\varphi(p)$. Letting $z = p$, $w \to \varphi(p)^{+}$, and $\lambda = |\sigma|/t$ in (7.6), it is clear that

$$(|\sigma|/t) D\varphi_p(\Uparrow_p) \leq < -\bar{\upsilon}, \log_{T_p X}(\sigma(|\sigma|)) > .$$

A simple rearrangement yields

$$D\varphi_p(\Uparrow_p) \leq < -\bar{\upsilon}, \Uparrow_p > \leq \max_{\xi \in \partial\varphi(p)} < \xi, \Uparrow_p > .$$

Consequently, $D\varphi_p(\Uparrow_p) = \max_{\xi \in \partial\varphi(p)} < \xi, \Uparrow_p >$ where the maximum is attained at $-\bar{\upsilon}$. This completes the proof. □

Proposition 7.5 yields the necessary optimality condition on Hadamard manifolds.

Theorem 7.2. *Suppose that X is a Hadamard manifold, that $\varphi \in \Gamma_0(X)$ is Lipschitz continuous, and assume that $\Omega \subset int(dom\varphi)$ is nonempty, closed, and convex. If $x^* \in \Omega$ is a solution of* (P)*, then there exists $\Uparrow^* \in \partial\varphi(x^*)$ for which* (7.4) *holds.*

Proof. Assume that (7.4) does not hold. Thus, for any $\Uparrow \in \partial\varphi(x^*)$ we can find $x_{\Uparrow} \in \Omega$ such that $< \Uparrow, \log_{T_{x^*}X}(x_{\Uparrow}) > < 0$. By Proposition 7.5, it is clear that $D\varphi_{x^*}(\log_{T_{x^*}X}(x_{\Uparrow})) < 0$. It is immediate from the definition of $D\varphi$ that $\varphi((1-\lambda)x^* \oplus \lambda x_{\Uparrow}) < \varphi(x^*)$, for all sufficiently small $\lambda > 0$. Since Ω is convex, all elements $(1-\lambda)x^* \oplus \lambda x_{\Uparrow}$ lie in Ω. This shows that x^* is not a solution to (P), which is a contradiction. Hence (7.4) must be true. □

Remark 7.3. The above optimality has been studied in the Riemannian manifold setting by exploiting the subdifferential calculus. A second proof is presented, that better evidences the nature of general globally NPC spaces where subdifferential calculus is not yet developed. Moreover, the question of the full generalization to globally NPC space is reduced to the construction of an effective separation theorem.

Next, another optimality condition is presented via a different type of inequality, the *Minty variational inequality*.

Theorem 7.3. *Suppose that X is a Hadamard manifold, $\varphi \in \Gamma_0(X)$ is Lipschitz continuous, and $\Omega \subset int(dom\varphi)$ is nonempty, closed, and convex. If $x^* \in \Omega$ satisfies the inequality*

$$< \eta, \log_{T_y X}(x^*) > \leq 0 \quad (^{\forall} y \in \Omega)(^{\forall} \eta \in \partial\varphi(y)), \tag{7.7}$$

then x^* is a solution of (P).

Proof. Assume that $x^* \in \Omega$ is not a solution of (P). Hence, there exists a point $z \in \Omega$ for which $\varphi(z) < \varphi(x^*)$. Define a function $h : [0,1] \to \mathbb{R}$ by

$$h(t) := \varphi((1-t)z \oplus tx^*) \quad ({}^\forall t \in [0,1]).$$

From the assumption of the theorem and basic facts of convex analysis, it is seen that h is convex and differentiable on a dense subset of $[0,1]$. Applying the mean value theorem, it is clear that there exists $\hat{t} \in (0,1)$ for which

$$\frac{dh}{dt}(\hat{t}) \geq \varphi(x^*) - \varphi(z) > 0.$$

Since $\frac{dh}{dt}(\hat{t}) = D\varphi_{\hat{x}}(\log_{T_{\hat{x}}X}(x^*))$, where $\hat{x} = (1-\hat{t})z \oplus \hat{t}x^*$, it follows from Proposition 7.5 that $< \upsilon, \log_{T_zX}(x^*) >$ is strictly positive for some $\upsilon \in \partial\varphi(z)$. Therefore x^* cannot be a solution to the inequality (7.7). $\square$

It shall be proved in the next Section that the Minty variational inequality (7.7) is also a necessary optimality condition for (P).

7.4 Vector fields and their variational inequalities

In the previous Section we introduced the subdifferential $\partial\varphi$ for each $\varphi \in \Gamma_0$; it was moreover observed there that partial optimality conditions for convex optimization problems in globally NPC spaces can be expressed in terms of $\partial\varphi$. Also, a full optimality characterization in the case of Hadamard manifolds can also be expressed in terms of the subdifferential. In this Section, the variational inequality given by (7.4) is extended to any vector field.

7.4.1 Subdifferential as a vector field

As motivation towards the study of variational inequalities, the following properties of $\partial\varphi$, considered as a vector field, are recalled.

Definition 7.9. A mapping $\Phi : X \multimap TX$ is called a *(set-valued) vector field* if $\Phi(x) \subset T_xX$ for all $x \in X$.

It is now clear that, when viewed as a mapping, $\partial\varphi : X \multimap TX$ has the property of a vector field. Moreover, $\partial\varphi$ is called the subdifferential of φ.

The following property of a vector field plays a vital role in what follows. It is observed that $\partial\varphi$ also behaves as a monotone vector field - a fact to be kept in mind throughout the paper.

Definition 7.10 ([8]). A vector field $\Phi : X \multimap TX$ is said to be *monotone* if the inequality

$$< \log_{T_pX}(q), \eta > \leq - < \log_{T_qX}(p), \nu >$$

holds for all $p, q \in X$, all $\eta \in \Phi(p)$ and $\nu \in \Phi(q)$. Moreover, a nonempty set $\Omega \subset X$, Φ is said to be *maximally monotone* if it is monotone and for any $x \in X$ and $\xi \in T_xX$ satisfying $< \xi, \log_{T_xX}(y) > \leq - < \eta, \log_{T_yX}(x) >$ for all $(y, \eta) \in \mathrm{gr}\Phi$, it holds $\xi \in \Phi(x)$.

A few important properties of the subdifferential are next summarized. We refer the reader to [8] for the proofs of the following statements.

Proposition 7.6 ([8]). *Let $\varphi \in \Gamma_0(X)$ and let $\partial\varphi : X \multimap TX$ be its subdifferential. Then the following properties are satisfied:*

(1) $\partial\varphi$ is maximally monotone.

(2) $0 \in \partial\varphi(x^*)$ if and only if x^* minimizes φ on X.

(3) $\lambda^{-1}\log_{T_{x^*}X}(x) \in \partial\varphi(x^*)$ if and only if $x^* = prox_\varphi^\lambda(x)$.

(4) For any given $\lambda > 0$ and $x \in X$, there exists a unique point $x^* \in X$ such that $\lambda^{-1}\log_{T_{x^*}X}(x) \in \partial\varphi(x^*)$.

7.4.2 Variational inequalities and equilibrium problems

In view of the above properties of $\partial\varphi$ in conjunction with the optimality conditions given in Section 7.3, there naturally arises the need for a general theory of variational inequalities for a monotone vector field. More precisely, for a given monotone vector field $\Phi : X \multimap TX$ and a nonempty closed convex set $\Omega \subset X$, we will consider the following *variational inequality* and denote it by $VI(\Phi, \Omega)$:

$$\text{Find } (x^*, \xi) \in \mathrm{gr}\Phi \text{ such that } < \xi, \log_{T_{x^*}X}(x) > \geq 0 \text{ for all } x \in \Omega. \tag{$VI(\Phi,\Omega)$}$$

In this case, the point x^* is said to be the solution of $VI(\Phi, \Omega)$. It is natural to also extend (7.7) to include a general vector field. The *Minty variational inequality*, referred to as $MVI(\Phi, \Omega)$, is given by:

$$\text{Find } x^* \in \Omega \text{ such that } < \eta, \log_{T_yX}(x^*) > \geq 0 \text{ for all } (y, \eta) \in \mathrm{gr}\Phi.$$

The equality $VI(\Phi, \Omega) = MVI(\Phi, \Omega)$ will be used in case the correspoding sets of solutions coincide.

In the linear setting, the study of variational inequalities relies largely on the linearity, or more precisely, the quasi-convexity, of $y \mapsto < \xi, \log_{T_{x^*}X}(x) >$. However, when X is a globally NPC space, neither linearity nor quasi-convexity is expected . In fact, one has the following characterization from [19], even in the case of a Hadamard manifold.

Theorem 7.4 ([19]). *Let X be an Hadamard manifold and $p \in X$. Then the following statements are equivalent:*

(1) $y \mapsto < \xi, \log_{T_{x^*}X}(x) >$ *is convex for each* $\xi \in T_pX$.

(2) $y \mapsto < \xi, \log_{T_{x^*}X}(x) >$ *is affine for each* $\xi \in T_pX$.

(3) The exponential map $\exp_p : T_pX \to X$ *is a global isometry.*

(4) The curve $t \mapsto \exp_p((1-t)\exp_p^{-1}(q^1) + t\exp_p^{-1}(q^2))$ *is a minimal geodesic.*

(5) X has identically zero sectional curvature (i.e., it is isometric to a Euclidean space).

In view of the preceding Theorem, the classical approach to solving $VI(\Phi,\Omega)$ has to be redesigned in the setting of Hadamard manifolds, or more generally, in the setting of globally NPC spaces. To solve $VI(\Phi,\Omega)$, the problem is written in the simpler, more general form of an equilibrium problem. Recall now that the *equilibrium problem* associated to a given domain $\Omega \subset X$ and to a bifunction $V : \Omega \times \Omega \to \mathbb{R}$, denoted by $EP(V,\Omega)$, is given by:

$$\text{Find } x^* \text{ such that } V(x^*, y) \geq 0 \text{ for all } y \in \Omega. \qquad (EP(V,\Omega))$$

The notion of Minty equilibrium problems, which can be seen as a counterpart of $EP(V,\Omega)$, is also needed. The *Minty equilibrium problem* associated to $\Omega \subset X$ and $V : \Omega \times \Omega \to \mathbb{R}$, denoted by $MEP(V,\Omega)$, is given by:

$$\text{Find } x^* \text{ such that } V(y, x^*) \leq 0 \text{ for all } y \in \Omega. \qquad (MEP(V,\Omega))$$

To simplify the notation, let us write $EP(V,\Omega) = MEP(V,\Omega)$ if the corresponding solution sets of the two problems coincide.

For any given monotone vector field $\Phi : X \multimap TX$ and a nonempty closed convex set $\Omega \subset \mathrm{dom}\Phi$, let $V_{\Phi,\Omega} : \Omega \times \Omega \to \mathbb{R}$ be the bifunction defined by

$$V_{\Phi,\Omega}(x,y) := \sup_{\xi \in \Phi(x)} < \xi, \log_{T_xX}(y) > \quad ({}^\forall x, y \in \Omega). \tag{7.8}$$

If $\Phi : X \multimap TX$ is a vector field and $\Omega \subset X$, then the problems $EP(V_{\Phi,\Omega},\Omega)$ and $MEP(V_{\Phi,\Omega},\Omega)$ are the variational inequality $VI(\Phi,\Omega)$ and the Minty variational inequality $MVI(\Phi,\Omega)$, respectively.

Recall now the following notion of *monotonicity* of a bifunction and the implication concerning the Minty equilibrium problem.

Definition 7.11. A bifunction $V : \Omega \times \Omega \to \mathbb{R}$ is said to be *monotone* if

$$V(x,y) \leq -V(y,x)$$

for every $x, y \in \Omega$.

Proposition 7.7. *Suppose that $\Omega \subset X$ is nonempty and that $V : \Omega \times \Omega \to \mathbb{R}$ is monotone. If $x^* \in \Omega$ is a solution of $EP(V,\Omega)$, then it is also a solution of $MEP(V,\Omega)$.*

Proof. If $x^* \in \Omega$ is a solution of $EP(V,\Omega)$, then for any $y \in \Omega$ it holds that

$$0 \leq V(x^*,y) \leq -V(y,x^*).$$

Thus, x^* solves $MEP(V,\Omega)$, as claimed. □

The next proposition shows that the monotonicity of Φ is inherited by $V_{\Phi,\Omega}$.

Proposition 7.8. *If $\Phi : X \multimap TX$ is a monotone vector field and $\Omega \subset X$ is nonempty, then the bifunction $V_{\Phi,\Omega}$ defined by* (7.8) *is monotone.*

Proof. Since Φ is monotone, one has

$$\begin{aligned} V_{\Phi,\Omega}(x,y) &= \sup_{\xi \in \Phi(x)} < \eta, \log_{T_xX}(y) > \leq \inf_{\eta \in \Phi(y)} - < \chi, \log_{T_yX}(x) > \\ &= - \sup_{\eta \in \Phi(y)} < \eta, \log_{T_yX}(x) > = -V_{\Phi,\Omega}(y,x), \end{aligned}$$

for any $x, y \in \Omega$. This shows the monotonicity of $V_{\Phi,\Omega}$. □

As in the above discussion together with Propositions 7.7 and 7.8, the following optimality condition complements Theorem 7.3.

Theorem 7.5. *Suppose that $\varphi \in \Gamma_0(X)$ and that $\Omega \subset X$ is nonempty, closed, and convex. If x^* is a solution of the variational inequality* (7.4)*, then, x^* solves the Minty variational inequality* (7.7).

Proof. Since $V_{\Phi,\Omega}$ is monotone, a solution of $EP(V_{\Phi,\Omega},\Omega)$ is also a solution of $MEP(V_{\Phi,\Omega},\Omega)$. □

The following result summarizes the relationships between the convex optimization problem, the variational inequality, and the Minty variational inequality.

Corollary 7.1. *Let X be a Hadamard manifold, $\varphi \in \Gamma_0$ be Lipschitz continuous and $\Omega \subset int(dom\varphi)$ be nonempty, closed and convex. Then the following statements are equivalent:*

(1) $x^* \in \Omega$ *is a solution of* (P).

(2) $x^* \in \Omega$ *is a solution of the variational inequality* (7.4).

(3) $x^* \in \Omega$ *is a solution of the Minty variational inequality* (7.7).

This corollary suggests the consideration of a special class of vector field, namely one whose variational inequality generalizes inequality (7.4), and hence the problem (P) within the manifold structure. This will be done by systematically passing some properties of such variational inequality to a more general nonconvex equilibrium problem, as shall be seen in the sequel.

7.5 Resolvent operators

In this Section, the notion of resolvent operator of a given bifunction is introduced. This concept is the pivotal ingredient in the construction of the proximal method. It is assumed in what follows that $\Omega \subset X$ is nonempty, closed and convex.

Definition 7.12. Let $V : \Omega \times \Omega \to \mathbb{R}$ be a given bifunction and $\lambda > 0$. The λ*-resolvent operator of* V is the mapping $R_V^\lambda : X \multimap X$ defined by

$$R_V^\lambda(x) := \{z \in \Omega \mid V(z,y) - \lambda^{-1} < \log_{T_zX}(x), \log_{T_zX}(y) > \geq 0 \ (^\forall y \in \Omega)\},$$

for all $x \in X$. V is said to be *prox-friendly* if $\mathrm{dom} R_V^\lambda \supset \Omega$ for any $\lambda > 0$.

Proposition 7.9. *Suppose that* $V : \Omega \times \Omega \to \mathbb{R}$ *is a monotone, prox-friendly bifunction. Then the following properties hold for all* $\lambda > 0$*:*

(1) R_V^λ *is single-valued.*

(2) R_V^λ *is nonexpansive on* Ω*.*

(3) $x^* \in \Omega$ *solves* $EP(V,\Omega)$ *if and only if* $x^* = R_V^\lambda(x^*)$*.*

Proof. 7.9 Let $x \in \mathrm{dom}(R_V^\lambda)$ and suppose that $z, z' \in R_V^\lambda(x)$. Thus,

$$\begin{cases} V(z,z') \geq \lambda^{-1} < \log_{T_zX}(x), \log_{T_zX}(z') >, \\ V(z',z) \geq \lambda^{-1} < \log_{T_{z'}X}(x), \log_{T_{z'}X}(z) > . \end{cases}$$

Adding up the two preceding inequalities, applying the monotonicity of F

and after some calculations, one concludes

$$\begin{aligned}0 &\geq V(z,z')+V(z',z)\\ &\geq \lambda^{-1} < \log_{T_z X}(x), \log_{T_z X}(z') > + \lambda^{-1} < \log_{T_{z'} X}(x), \log_{T_{z'} X}(z) >\\ &\geq \lambda^{-1}\rho^2(z,z').\end{aligned}$$

This shows $z = z'$.

7.9 Let $x, y \in \Omega$. By the definition of R_V^λ, it follows that

$$\begin{cases} V(R_V^\lambda(x), R_V^\lambda(y)) - \lambda^{-1} < \log_{T_{R_V^\lambda(x)}X}(x), \log_{T_{R_V^\lambda(x)}X}(R_V^\lambda(y)) > \geq 0, \\ V(R_V^\lambda(y), R_V^\lambda(x)) - \lambda^{-1} < \log_{T_{R_V^\lambda(y)}X}(y), \log_{T_{R_V^\lambda(y)}X}(R_V^\lambda(x)) > \geq 0. \end{cases}$$

Adding up the two inequalities above, applying the monotonicity of V, and multiplying both sides by λ, it becomes clear that

$$\begin{aligned} 0 \geq\ & < \log_{T_{R_V^\lambda(x)}X}(x), \log_{T_{R_V^\lambda(x)}X}(R_V^\lambda(y)) > \\ & + < \log_{T_{R_V^\lambda(y)}X}(y), \log_{T_{R_V^\lambda(y)}X}(R_V^\lambda(x)) > \\ \geq\ & \left[\rho^2(R_V^\lambda(x), x) + \rho^2(R_V^\lambda(x), R_V^\lambda(y)) - \rho^2(x, R_V^\lambda)\right] \\ & + \left[\rho^2(R_V^\lambda(y), y) + \rho^2(R_V^\lambda(x), R_V^\lambda(y)) - \rho^2(y, R_V^\lambda(x))\right]. \end{aligned}$$

Rearranging terms in the above inequality and using the global NPC condition, one has

$$\begin{aligned} \rho^2(R_V^\lambda(x), R_V^\lambda(y)) &\leq \frac{1}{2}\Big[\rho^2(x, R_V^\lambda(y)) - \rho^2(y, R_V^\lambda(x)) \\ &\qquad -\rho^2(x, R_V^\lambda(x)) - \rho^2(y, R_V^\lambda(y))\Big] \\ &\leq \rho(x,y)\rho(R_V^\lambda(x), R_V^\lambda(y)), \end{aligned}$$

which shows that R_V^λ is nonexpansive.

7.9 Let $x^* \in \Omega$. Observe that

$$\begin{aligned} x^* = R_V^\lambda(x^*) &\iff V(x^*, y) - \lambda^{-1} < \log_{T_{x^*}X}(x^*), \log_{T_{x^*}X}(y) > \geq 0 \\ &\iff V(x^*, y) \geq 0, \end{aligned}$$

for all $y \in \Omega$. That is, $x^* \in \Omega$ solves $EP(V, \Omega)$ if and only if $x^* = R_V^\lambda(x^*)$. $\square$

The formula for R_V^λ can be replaced with a simpler, more explicit expression in the case of convex optimization (P). Specifically,

Proposition 7.10. *If $\varphi \in \Gamma_0(X)$, $\lambda > 0$, and $x \in X$, then $z = R^{\lambda}_{V_{\partial\varphi,\Omega}}(x)$ if and only if*

$$z = argmin_{y\in\Omega}\left[\varphi(y) + \frac{1}{2\lambda}d^2(y,x)\right] = prox^{\lambda}_{\varphi+\delta_\Omega}(x).$$

Proof. Taking the maximal monotonicity of $\partial\varphi$ one has

$$\begin{aligned}
z = R^{\lambda}_{V_{\partial\varphi,\Omega}}(x) &\iff V_{\partial\varphi,\Omega}(z,y) - \lambda^{-1} < \log_{T_zX}(x), \log_{T_zX}(y) > \geq 0 \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (^\forall y \in \Omega) \\
&\iff \sup_{\upsilon\in\partial\varphi(z)} < \upsilon, \log_{T_zX}(y) > \geq \lambda^{-1} < \log_{T_zX}(x), \log_{T_zX}(y) > \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (^\forall y \in \Omega) \\
&\iff \varphi(y) - \varphi(z) \geq \lambda^{-1} < \log_{T_zX}(x), \log_{T_zX}(y) > \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (^\forall y \in \Omega) \\
&\iff \lambda^{-1}\log_{T_zX}(x) \in \partial(\varphi + \delta_\Omega)(z) \\
&\iff z = \mathrm{prox}^{\lambda}_{\varphi+\delta_\Omega}(x).
\end{aligned}$$

□

7.6 Proximal algorithms

In this Section, the convergence of the proximal methods associated to a monotone bifunction $V : \Omega \times \Omega \to \mathbb{R}$ is proved. Since the aim is to apply our results to the variational inequalities for monotone vector fields that are not convex in the second argument (unless the curvature vanishes), a completely new condition on V is needed. Recall that $\Omega \subset X$ is always assumed to be nonempty, closed and convex. The *proximal algorithm* consists of generating a sequence (x^k) from a given initial $x^0 \in \Omega$ by setting

$$x^{k+1} := R^{\lambda_k}_V(x^k) \quad (k = 0, 1, 2, \cdots), \tag{7.9}$$

where (λ_k) is a sequence of positive reals.

Definition 7.13. A bifunction $V : \Omega \times \Omega \to \mathbb{R}$ is said to be *skewed Δ-upper semicontinuous* (for short, *skewed Δ-usc*) if $-V(y,x^*) \geq \limsup_k V(x^k,y)$ for all $y \in \Omega$, whenever (x^k) is a Δ-convergent sequence in Ω with Δ-limit point $x^* \in \Omega$.

It is now time to state and prove the main convergence result of this Chapter, which concerns the proximal algorithm applied to a general nonconvex equilibrium problem.

Theorem 7.6. *Suppose that V is a prox-friendly monotone bifunction which is skewed Δ-usc and assume that $EP(V,\Omega) = MEP(V,\Omega)$ has a solution. Let (λ_k) be a sequence of positive reals which is bounded away from* 0*. Then the sequence generated by* (7.9) *is Δ-convergent to a solution of $EP(V,\Omega)$ for any initial starting point $x^0 \in K$.*

Proof. Let $x^0 \in \Omega$ be an initial starting point and let $x^* \in \Omega$ be a solution of $EP(V,\Omega)$. Clearly,

$$d(x^*, x^{k+1}) = d(R_V^{\lambda_k}(x^*), R_V^{\lambda_k}(x^k)) \leq d(x^*, x^k),$$

which implies that (x^k) is Fejér convergent with respect to $\mathbf{S}$, where $\mathbf{S}$ is the set of all solutions to $EP(V,\Phi)$. In view of Proposition 7.2, the real sequence $(\rho(x^k,x^*))$ is bounded, and therefore it converges to some $\ell \geq 0$. Since $x^{k+1} = R_V^{\lambda_k}(x^k)$, we have

$$V(x^{k+1}, x^*) \geq \lambda_k^{-1} < \log_{T_{x^{k+1}}X}(x^k), \log_{T_{x^{k+1}}X}(x^*) > .$$

Since x^* is a solution of $EP(V,\Omega)$ and V is monotone, one has $V(x^{k+1},x^*) \leq 0$. In view of the above inequality, it is apparent that $< \log_{T_{x^{k+1}}X}(x^k), \log_{T_{x^{k+1}}X}(x^*) > \leq 0$. On account of the definition of the scalar product and by the law of cosines in the model space, it follows that

$$\begin{aligned}
0 &\geq < \log_{T_{x^{k+1}}X}(x^k), \log_{T_{x^{k+1}}X}(x^*) > \\
&= d(x^{k+1},x^k) d(x^{k+1},x^*) \cos \angle_{x^{k+1}} (\log_{\Sigma_{x^{k+1}}X}(x^k), \log_{\Sigma_{x^{k+1}}X}(x^*)) \\
&\geq d(x^{k+1},x^k) d(x^{k+1},x^*) \cos \overline{\angle}_{\overline{x^{k+1}}} (\overline{x^k}, \overline{x^*}) \\
&= d^2(x^{k+1},x^k) + d^2(x^{k+1},x^*) - d^2(x^k,x^*).
\end{aligned}$$

Taking the limit as $k \to \infty$ in the above inequalites, it is clear that $\lim_k d(x^{k+1},x^k) = 0$.

Suppose that $\hat{x} \in \Omega$ is a Δ-accumulation point of (x^k). There exists thus a subsequence (x^{k_j}) of (x^k) whose Δ-limit point is $\hat{x}$. From the definition of the resolvent operator and by virtue of the Cauchy-Schwarz inequality, the following inequalities hold:

$$\begin{aligned}
V(x^{k_j}, y) &\geq \lambda_{k_j-1}^{-1} < \log_{T_{x^{k_j}}X}(x^{k_j-1}), \log_{T_{x^{k_j}}X}(y) > \\
&\geq -\lambda_{k_j-1}^{-1} d(x^{k_j}, x^{k_j-1}) d(x^{k_j}, y),
\end{aligned}$$

for any $y \in \Omega$. From the boundedness of (x^k) and recalling that (λ_k) is bounded away from 0, one gets

$$V(x^{k_j}, y) \geq -Md(x^{k_j}, x^{k_j - 1})$$

for some constant $M > 0$. Since V is skewed Δ-usc, letting $j \to \infty$ one has the following inequalities

$$-V(y, \hat{x}) \geq \limsup_{j\to\infty} V(x^{k_j}, y) \geq 0.$$

Since $y \in \Omega$ is arbitrary and $EP(V, \Omega) = MEP(V, \Omega)$, it is clear that $\hat{x}$ belongs to the solution set $\mathbf{S}$ and so does every Δ-accumulation point of (x^k). In view of Proposition 7.2, it follows that (x^k) is Δ-convergent to a solution of $EP(V, \Omega)$. □

It is next proved that $V_{\Phi,\Omega}$ with a monotone vector field Φ, is skewed usc in a locally compact space. Note that the assumption of local compactness should not be an issue of concern, for the Δ-convergence (or weak convergence in the linear settings) in an infinite dimensional space, cannot be instantly detected in practical implementations. In fact, it is known that any Hadamard manifold is locally compact and therefore all subsequent results are perfectly useful.

Proposition 7.11. *Let X be locally compact and $\Phi : X \multimap TX$ be a monotone vector field. If $V_{\Phi,\Omega}$ is defined as in* (7.8)*, then it is skwed usc.*

Proof. Suppose that (x^k) is a sequence in Ω that is convergent to a point $x^* \in \Omega$. Let $y \in \Omega$ be arbitrary. Then, by the monotonicity of Φ and hence of $V_{\Phi,\Omega}$, one has

$$\begin{aligned}
\limsup_{k\to\infty} V_{\Phi,\Omega}(x^k, y) &\leq \limsup_{k\to\infty}[-V_{\Phi,\Omega}(y, x^k)] \\
&= \limsup_{k\to\infty}[-\sup_{\eta\in\Phi(y)} < \eta, \log_{T_yX}(x^k) >] \\
&\leq \limsup_{k\to\infty}[- < \eta_0, \log_{T_yX}(x^k) >] \\
&\leq - < \eta_0, \log_{T_yX}(x^*) >,
\end{aligned}$$

for any $\eta_0 \in \Phi(y)$. It follows that

$$\begin{aligned}
\limsup_{k\to\infty} V_{\Phi,\Omega}(x^k, y) &\leq \inf_{\eta\in\Phi(y)}[- < \eta, \log_{T_yX}(x^*) >] \\
&= -\sup_{\eta\in\Phi(y)} < \eta, \log_{T_yX}(x^*) > \\
&= -V(y, x^*).
\end{aligned}$$

Since $y \in \Omega$ is arbitrary, it is clear that V is skewed usc. □

The following is the proof of the convergence of the proximal algorithm applied to the solution of a class of variational inequalities in a globally NPC space, which naturally lacks the affinity/convexity/concavity valid in the linear settings.

Theorem 7.7. *Let X be locally compact and $\Phi : X \multimap TX$ be a monotone vector field, such that $VI(\Phi,\Omega) = MVI(\Phi,\Omega)$ has a solution. Suppose that $V_{\Phi,\Omega}$, defined by* (7.8)*, is prox-friendly and that (λ_k) is a sequence of positive reals which is bounded away from* 0*. Then the sequence generated by* (7.9) *applied to $V_{\Phi,\Omega}$ is convergent to a solution of $VI(\Phi,\Omega)$.*

Proof. By Proposition 7.11, V_Φ is skewed usc. Apply Theorem 7.6 to $V_{\Phi,\Omega}$ to obtain the desired result. □

The final main result in this Chapter concerns the proximal algorithm associated to the solution of (P) in the context of a Hadamard manifold.

Theorem 7.8. *Let X be a Hadamard manifold, $\varphi \in \Gamma_0(X)$, $\Omega \subset int(dom\varphi)$, and suppose that* (P) *has a solution. Let (λ_k) be a sequence of positive reals which is bounded away from* 0*. Then the sequence (x^k) generated by*

$$x^{k+1} := prox^{\lambda_k}_{\varphi+\delta_\Omega}(x^k) \quad (k = 0, 1, 2, \cdots),$$

with any initial start $x^0 \in \Omega$, is convergent to a solution of (P).

Proof. By Propositions 7.6 and 7.10, the resolvent $R^{\lambda_k}_{V_{\partial\varphi,\Omega}} = \mathrm{prox}_{\varphi+\delta_\Omega}$ is well-defined. Moreover, Proposition 7.1 shows that $EP(V_{\partial\varphi,\Omega},\Omega) = MEP(V_{\partial\varphi,\Omega})$. Again, it transpires from Proposition 7.6 that $\partial\varphi$ is monotone and hence it follows from Proposition 7.11 that $V_{\partial\varphi,\Omega}$ is skew usc. By Theorem 7.7, one has that (x^k) converges to a solution of $EP(V_{\partial\varphi,\Omega},\Omega)$, which is no other than the solution of (P). □

Conclusion

We have developed the fundamental theory of convex optimization of variational inequalities and equilibrium problems. Despite the different levels of generality, the foundation of the interplay between convex optimization problems and variational inequalities has been laid. It has been underlined that the variational inequalities generated in such situations are never convex, unless the space has zero curvature. This fact highlights the important difference between the linear setting and the case of spaces with non-zero curvature. Therefore, a new approach has to be found to deal with such difficulty. This Chapter presented one new way that handles the problem in such a way that it can be solved using very classical methods, such as proximal algorithms.

Taking into account the results where a manifold structure is needed, it is apparent that, in fact, only a *powerful* separation theorem is required. Thus, the problem of passing from a Hadamard manifold to, at least, a locally compact globally NPC space, consists of developing the required separation theorem. To the best of the authors' knowledge, the present work is the first in the literature to adopt the separation theorem and prove the characterization of subdifferentials by directional derivatives with *equalities* outside the linear setting.

Bibliography

[1] P.-A. Absil, R. Mahony, R. Sepulchre, *Optimization algorithms on matrix manifolds*, Princeton University Press, Princeton, NJ, 2008. With a foreword by Paul Van Dooren.

[2] B. A. Kakavandi, *Weak topologies in complete* $\mathrm{CAT}(0)$ *metric spaces*, Proc. Amer. Math. Soc., 141(3) (2013), 1029–1039.

[3] B. A. Kakavandi, M. Amini, *Duality and subdifferential for convex functions on complete* $\mathrm{CAT}(0)$ *metric spaces*, Nonlinear Anal., 73(10) (2010), 3450–3455.

[4] M. Bačák, *The proximal point algorithm in metric spaces*, Israel J. Math., 194(2) (2013), 689–701.

[5] M. Bačák, *Convex analysis and optimization in Hadamard spaces*, De Gruyter Series in Nonlinear Analysis and Applications Vol. 22, De Gruyter, Berlin, 2014.

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

[7] D. Burago, Y. Burago, S. Ivanov, *A course in metric geometry*, Graduate Studies in Mathematics Vol. 33, American Mathematical Society, Providence, RI, 2001.

[8] P. Chaipunya, F. Kohsaka, P. Kumam, *Monotone vector fields and generation of nonexpansive semigroups in complete cat(0) spaces*, arXiv:1906.05984, 2019.

[9] P. Chaipunya, P. Kumam, *Nonself KKM maps and corresponding theorems in Hadamard manifolds*, Appl. Gen. Topol., 16(1) (2015), 37–44.

[10] P. Chaipunya, P. Kumam, *On the proximal point method in Hadamard spaces*, Optimization, 66(10) (2017), 1647–1665.

[11] S. Dhompongsa, W. A. Kirk, B. Sims, *Fixed points of uniformly Lipschitzian mappings*, Nonlinear Anal., 65(4) (2006), 762–772.

[12] T. Iwamiya, H. Okochi, *Monotonicity, resolvents and Yosida approximations of operators on Hilbert manifolds*, Nonlinear Anal., 54(2) (2003), 205–214.

[13] J. Jost, *Convex functionals and generalized harmonic maps into spaces of nonpositive curvature*, Comment. Math. Helv., 70(4) (1995), 659–673.

[14] H. Khatibzadeh, S. Ranjbar, *Monotone operators and the proximal point algorithm in complete* $\mathrm{CAT}(0)$ *metric spaces*, J. Aust. Math. Soc., 103(1) (2017), 70–90.

[15] Y. Kimura, Y. Kishi, *Equilibrium problems and their resolvents in Hadamard spaces*, Journal of nonlinear and convex analysis 19(9) (2017), 1503-1513.

[16] W. A. Kirk, B. Panyanak, *A concept of convergence in geodesic spaces*, Nonlinear Anal., 68(12) (2008), 3689–3696.

[17] B. Kleiner, *The local structure of length spaces with curvature bounded above*, Math. Z., 231(3) (1999), 409–456.

[18] B. Kleiner, B. Leeb, *Rigidity of quasi-isometries for symmetric spaces and Euclidean buildings*, Inst. Hautes Études Sci. Publ. Math., (86) (1997), 115–197.

[19] A. Kristály, C. Li, G. López-Acedo, A. Nicolae, *What do 'convexities' imply on Hadamard manifolds?*, J. Optim. Theory Appl., 170(3) (2016), 1068–1074.

[20] P. Kumam, P. Chaipunya, *Equilibrium problems and proximal algorithms in hadamard spaces*, Journal of Nonlinear Analysis and Optimization: Theory & Applications, 8(2) (2017), 155–172.

[21] C. Li, G. López, V. Martín-Márquez, *Monotone vector fields and the proximal point algorithm on Hadamard manifolds*, J. Lond. Math. Soc. (2), 79(3) (2009), 663–683.

[22] C. Li, B. S. Mordukhovich, J. Wang, J.-C. Yao, *Weak sharp minima on Riemannian manifolds*, SIAM J. Optim., 21(4) (2011), 1523–1560.

[23] T. C. Lim, *Remarks on some fixed point theorems*, Proc. Amer. Math. Soc., 60 (1977), 179–182.

[24] U. F. Mayer, *Gradient flows on nonpositively curved metric spaces and harmonic maps*, Comm. Anal. Geom., 6(2) (1998), 199–253.

[25] I. Nikolaev, *The tangent cone of an Aleksandrov space of curvature* $\leq K$, Manuscripta Math., 86(2) (1995), 137–147.

[26] C. Udrişte, *Convex functions and optimization methods on Riemannian manifolds*, Mathematics and its Applications Vol.297, Kluwer Academic Publishers Group, Dordrecht, 1994.

[27] J. H. Wang, G. López, V. Martín-Márquez, C. Li, *Monotone and accretive vector fields on Riemannian manifolds*, J. Optim. Theory Appl., 146(3) (2010), 691–708.

Chapter 8

Motion factorization and bond theory in hyperbolic kinematics

D. F. Scharler
Department of Basic Sciences in Engineering Sciences, University of Innsbruck, Innsbruck, Austria,
E-mail: daniel.scharler@uibk.ac.at

H. P. Schröcker
Department of Basic Sciences in Engineering Sciences, University of Innsbruck, Innsbruck, Austria,
E-mail: hans-peter.schroecker@uibk.ac.at

In this Chapter we look at bond theory and motion factorization, two recently developed theories for the analysis and synthesis of closed loop linkages, from the viewpoint of planar hyperbolic kinematics. We do this in order to better understand phenomena observed in Euclidean kinematics. Bond theory in hyperbolic geometry allows for a real and finite visualization and gives a concrete meaning to an originally abstract and imaginary concept. In hyperbolic geometry the number of real factorizations (even if finite, which is not always the case) can be substantially larger than in Euclidean geometry. The presence of zero divisors in the algebraic description makes factorization algorithms more subtle but also more interesting.

8.1 Introduction

In recent years, the parametrization of Euclidean motion groups via quaternions has been used to gain new insight on flexible structures with revolute and translational joints. The factorization of motion polynomials [8, 12, 13] yields new mechanisms with certain prescribed properties while "bond theory" [7, 9, 14] turned out to be a versatile tool for the analysis of mechanical structures. Here, we develop both theories in planar hyperbolic kinematics mainly for the purpose of illustrating certain phenomena that also occur in Euclidean kinematics but cannot be easily explained or visualized there.

This is especially true for bond theory where “bonds” are defined purely algebraically and their geometric or kinematic meaning is rather unclear. They can be thought of as points in the closure of a configuration variety with “degenerate” kinematic behavior. In hyperbolic kinematics they can be real and allow visualization in the Cayley-Klein model of hyperbolic geometry.

Motion factorization in hyperbolic kinematics is more challenging than in Euclidean kinematics because the underlying algebraic structure of split quaternions is more intricate than that of quaternions or dual quaternions. While algorithms for computing factorizations generically still work for left polynomials over split quaternions, the presence of non-invertible elements and their interesting geometric structure accounts for phenomena that are hidden in Euclidean geometry. Most notably, the number of factorizations can be larger in hyperbolic kinematics, even for generic polynomials.

The main purposes of this text are illustration and visualization. This necessitates abandoning the viewpoint of traditional (axiomatic) hyperbolic geometry in favor of “universal hyperbolic geometry” in the sense of [22, 23, 24]. In contrast to traditional hyperbolic geometry, points inside and outside the “absolute circle” or “null circle” are treated on equal footing. This also fits well with our algebraic approach where the additional inequality constraint of traditional hyperbolic geometry would only complicate matters and obscure results.

We continue this text with a split-quaternion-oriented introduction to universal hyperbolic geometry in Section 8.2, before we present the factorization theory for split quaternion polynomials and its kinematic interpretation in Section 8.3. Section 8.4 is devoted to bond theory. Due to the algebraic equivalence of spherical and hyperbolic kinematics over the complex numbers, the transfer of definitions and results from Euclidean geometry is straightforward, whence the focus will mainly be on examples and their illustration.

8.2 Split quaternions and hyperbolic geometry

In this Section we introduce notions and notation concerning split quaternions, the Section also features an introduction to universal hyperbolic geometry.

8.2.1 Split quaternion basics

The split quaternions $\mathbb{S}$ form an associative real algebra of dimension four. A split quaternion can be written as $h = h_0 + h_1\mathbf{i} + h_2\mathbf{j} + h_3\mathbf{k}$. Multiplication is defined by the generating rules

$$\mathbf{i}^2 = -\mathbf{j}^2 = -\mathbf{k}^2 = -\mathbf{ijk} = -1.$$

From this, the complete multiplication table can be worked out:

	$\mathbf{i}$	$\mathbf{j}$	$\mathbf{k}$
$\mathbf{i}$	-1	$\mathbf{k}$	$-\mathbf{j}$
$\mathbf{j}$	$-\mathbf{k}$	1	$-\mathbf{i}$
$\mathbf{k}$	$\mathbf{j}$	$\mathbf{i}$	1

It is clear then that the split quaternion algebra is not commutative. The split quaternion *conjugate* to $h = h_0 + h_1\mathbf{i} + h_2\mathbf{j} + h_3\mathbf{k}$ is defined as $\overline{h} := h_0 - h_1\mathbf{i} - h_2\mathbf{j} - h_3\mathbf{k}$, the split quaternion *norm* is given by

$$h\overline{h} = \overline{h}h = h_0^2 + h_1^2 - h_2^2 - h_3^2.$$

The split quaternion norm is real but, in contrast to the case of ordinary (Hamiltonian) quaternions (defined by the generating relations $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1$), may attain negative values as well. Every split quaternion with non-zero norm has a multiplicative inverse given by

$$h^{-1} = \frac{\overline{h}}{h\overline{h}}.$$

Split quaternions with zero norm are not invertible.

The *scalar* or *real part* of $h = h_0 + h_1\mathbf{i} + h_2\mathbf{j} + h_3\mathbf{k}$ is defined as $\operatorname{Re} h := \frac{1}{2}(h + \overline{h}) = h_0$, its *vector* or *imaginary part* is $\operatorname{Im} h := \frac{1}{2}(h - \overline{h}) = h_1\mathbf{i} + h_2\mathbf{j} + h_3\mathbf{k}$. Split quaternions with zero scalar part are called *vectorial.*

Split quaternion multiplication gives rise to the *scalar product* and the *cross product.* For $h, k \in \mathbb{S}$, define

$$\langle h, k\rangle := \tfrac{1}{2}(h\overline{k} + k\overline{h}) \quad \text{and} \quad h \times k := \tfrac{1}{2}(hk - kh).$$

Writing $h = h_0 + h_1\mathbf{i} + h_2\mathbf{j} + h_3\mathbf{k}$ and $k = k_0 + k_1\mathbf{i} + k_2\mathbf{j} + k_3\mathbf{k}$ it is clear that

$$\langle h,k\rangle = h_0k_0 + h_1k_1 - h_2k_2 - h_3k_3 \quad \text{and}$$
$$h \times k = (h_3k_2 - h_2k_3)\mathbf{i} + (h_3k_1 - h_1k_3)\mathbf{j} + (h_1k_2 - h_2k_1)\mathbf{k}.$$

Note that $\langle h,h\rangle = h\overline{h}$ and that $h \times k = \operatorname{Im} h \times \operatorname{Im} k$.

8.2.2 Hyperbolic geometry

The points of the hyperbolic plane H^2 are the points of the projective plane over the vector space $\operatorname{Im}\mathbb{S}$ of vectorial split quaternions. The projective point represented by the vector $h \in \operatorname{Im}\mathbb{S}$ will be denoted by $[h]$. The geometric structure of H^2 is given by the split quaternion multiplication rule. In the sequel, the terminology of [22, 23, 24] will be followed. The *absolute circle* or *null circle* $\mathscr{N}$ is the conic defined by the quadratic form $(h,k) \mapsto \langle h,k\rangle$. The null circle is real and regular. Its points are called *null points,* its tangents are *null lines*. The polarity with respect to $\mathscr{N}$ is called *absolute polarity* or *null polarity.*

The *quadrance* of two non-null points $[h]$, $[k] \in H^2 \setminus \mathscr{N}$ is

$$Q([h],[k]) := 1 - \frac{\langle h,k\rangle^2}{\langle h,h\rangle\langle k,k\rangle}. \tag{8.1}$$

The quadrance equals the square of the distance of traditional hyperbolic geometry but is well defined for any two non-null points and may attain negative values.

The line spanned by two different points $[h]$, $[k] \in H^2$ will be represented by the point $[u]^\vee$ of the dual projective plane, given by $u = h \times k$. The line $[u]^\vee$ and the point $[x]$ are incident if and only if $\langle u,x\rangle = 0$. These conventions fit nicely in our split quaternion approach but are slightly (up to certain sign changes) different from more common conventions in projective geometry [2].

The *spread* of two non-null lines $[u]^\vee$, $[v]^\vee$ is defined as

$$S([u]^\vee,[v]^\vee) := 1 - \frac{\langle u,v\rangle^2}{\langle u,u\rangle\langle v,v\rangle}. \tag{8.2}$$

It equals the square of the sine of the angle in traditional hyperbolic geometry. Comparing (8.1) and (8.2), a perfect duality is observed between points and lines in the hyperbolic plane.

The kinematics of hyperbolic geometry can be conveniently introduced via reflection. A *hyperbolic reflection* μ is a homology in the sense of [2, Section 5.7] that preserves $\mathscr{N}$. It has a center $[h]$ and an axis $[h]^\vee$, which are absolutely polar and non-null. Every line through $[h]$ and every point of $[h]^\vee$ is fixed under μ.

Theorem 8.1. *The reflection with center* $[h]$ *and axis* $[h]^\vee$ *is the map*

$$\mu : H^2 \to H^2, \quad [x] \mapsto [hx\overline{h}] \tag{8.3}$$

Note that h in (8.3) *is vectorial and non-null (*$h\overline{h} \neq 0$*).*

Proof. In order to see that (8.3) really describes a reflection, it suffices to show that μ preserves $\mathscr{N}$ and has a line of fixpoints [2, Corollary 5.7.13]. The former follows from

$$\langle hx\overline{h}, hx\overline{h}\rangle = (hx\overline{h})\overline{hx\overline{h}} = (h\overline{h})^2\, x\overline{x} = (h\overline{h})^2\, \langle x, x\rangle.$$

In order to see the latter, consider the line $[h]^\vee$ and any of its points $[x]$. From the incidence condition $\langle h, x\rangle = 0$ it is inferred that $xh = -hx$, whence $\mu([x]) = [h^2 x] = [x]$ because $h^2 = -h\overline{h} \in \mathbb{R}$. □

The composition of two reflections is called a *rotation.* Denote the reflection centers by $[h_1]$ and $[h_2]$ and their axes by $[h_1]^\vee$ and $[h_2]^\vee$, respectively. Then the *rotation center* $[h_1]^\vee \cap [h_2]^\vee$ is a fixed point of the rotation, and the *rotation axis* $[h_1] \vee [h_2]$ is a fixed line. Rotation centers and rotation axes are absolutely polar. Note that a rotation has several decompositions into two reflections but centers and axes are still well-defined. The algebraic description of rotations in terms of split quaternions generalizes (8.3):

Theorem 8.2. *For any split quaternion h with non-zero norm, the map*

$$\rho : H^2 \to H^2, \quad [x] \mapsto [hx\overline{h}] \tag{8.4}$$

is a rotation. The rotation center is $[\operatorname{Im} h]$.

Remark 8.1. In contrast to (8.3), h in (8.4) is *not* required to be vectorial. One consequence of Theorem 8.2 is that reflections should viewed as special rotations or, equivalently, the identity should be considered as a reflection. This is different from the case in Euclidean geometry.

Lemma 8.1. *For every split quaternion h there exist vectorial split quaternions u, w such that* $h = uw$.

Proof. Pick two independent vectors $u = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$ and $v = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$ in the orthogonal space of $\operatorname{Im} h$ in $\operatorname{Im}\mathbb{S}$, whose norm does not vanish. This is possible because the orthogonal space has dimension two (we omit the trivial case $h = 0$) and contains at most two directions of vanishing norm. Set $w := \lambda u + \mu v$ and solve $h = uw$ for λ and μ. With $h = h_0 + h_1\mathbf{i} + h_2\mathbf{j} + h_3\mathbf{k}$ the scalar part of this equation yields

$$\lambda = -\frac{\mu\langle u, v\rangle + h_0}{u\overline{u}}.$$

Plug this in the remaining three equations, which are linear in μ:

$$(u_2v_3 - u_3v_2)\mu + h_1 = (u_1v_3 - u_3v_1)\mu + h_2 = (u_2v_1 - u_1v_2)\mu + h_3 = 0.$$

The vectors u and v are independent. Hence the coefficients of μ do not all vanish and there exists at most one solution. The three equations are independent if and only if

$$\begin{aligned}
(u_2v_3 - u_3v_2)h_2 - (u_1v_3 - u_3v_1)h_1 &= 0,\\
(u_1v_3 - u_3v_1)h_3 - (u_2v_1 - u_1v_2)h_2 &= 0,\\
(u_2v_3 - u_3v_2)h_3 - (u_2v_1 - u_1v_2)h_1 &= 0.
\end{aligned}$$

The left-hand sides of these equations are (up to sign) just the coefficients of $h \times (u \times v)$. Since u and v are in the orthogonal space of $\operatorname{Im} h$, they vanish and a solution for μ does exist. It gives rise to a vectorial split quaternion w that satisfies $h = uw$, as required. □

Proof of Theorem 8.2. By Lemma 8.1, there exist vectorial split quaternions u, w such that $h = uw$. Since $h\overline{h} \neq 0$, the norms of u and w cannot vanish. Thus, the reflections μ_u and μ_w in $[u]$ and $[w]$, respectively, are well-defined and

$$hx\overline{h} = (uw)x\overline{(uw)} = u(wx\overline{w})\overline{u}$$

implies $\rho = \mu_u irc \mu_w$. It can be immediately confirmed that $h = -\langle u, w\rangle + u \times w$, whence the rotation center is indeed $[u \times w] = [\operatorname{Im} h]$. □

8.3 Motion factorization

Motion factorization is an algebraic procedure for representing polynomials with quaternion coefficients as products of linear polynomials. The factorization theory of split quaternions is not difficult in generic cases but, due to non-commutativity and to the existence of non-invertible elements, it requires some care. Some phenomena are not encountered in the more familiar case of polynomial factorization over the real numbers. Most notably, factorizations are, in general, no longer unique.

One motivation for studying quaternion polynomial factorization is kinematics. In the field of kinematics, factorization corresponds to the decomposition of a rational motion into the product of coupled rotations. Non-uniqueness of factorizations allows the construction of mechanisms from them. In Euclidean geometry, the properties of factorizations are interesting enough to merit publication in an engineering context [10, 11, 20]. Here factorizations are presented from the point of view of hyperbolic geometry which is, of course, less relevant in engineering. Nonetheless, factorizations in this case illustrate phenomena that are hidden in Euclidean geometry and thus furthers our understanding of quaternion polynomial factorization theory.

8.3.1 Split quaternion polynomials

A split quaternion polynomial C of degree d in the indeterminate t, is an expression of the form

$$C = \sum_{i=0}^{d} c_i t^i, \tag{8.5}$$

with $c_0, c_1, \cdots, c_d \in \mathbb{S}$ and $c_d \neq 0$. The addition of polynomials is defined in the usual way but multiplication and evaluation at h require some consideration. We will use split quaternion polynomials to describe rational motions in universal hyperbolic geometry, where the indeterminate t serves as a real parameter. This leads to postulate that *t commutes with all coefficients.* The product of C defined in (8.5) and $D = \sum_{i=0}^{e} d_i t^i$ is $CD = \sum_{i=0}^{d+e} e_i t^i$ where

$$e_i = \sum_{j+k=i} c_i d_j \quad \text{for } i = 0, 1, \cdots, \, d+e.$$

While this is just what is required for kinematics, it is just one among several multiplication rules from a mathematical viewpoint [19]. Since the multiplication order matters, the convention is adopted to always write coefficients to the left of the indeterminate. Even if this is often emphasized by speaking of "left polynomials" we will simply refer to the thus defined ring $\mathbb{S}[t]$ as the *ring of split quaternion polynomials.* The *value of C at h* $\in \mathbb{S}$ is defined as

$$C(h) := \sum_{i=0}^{d} c_i h^i.$$

The computation of $C(h)$ requires the expanded form of C, whence evaluation of C at h is not a homomorphism between the rings $\mathbb{S}[t]$ and $\mathbb{S}$. At any rate, evaluation is additive, that is, $C(h) + D(h) = (C + D)(h)$ for any C, $D \in \mathbb{S}[t]$ and $h \in \mathbb{S}$.

The polynomial $\overline{C}$ *conjugate* to C is obtained by conjugating the coefficients of C, namely

$$\overline{C} = \sum_{i=0}^{d} \overline{c_i} t^i.$$

The *norm polynomial* is defined as $N(C) := C\overline{C}$. Its coefficients are real.

Via (8.4), a split quaternion h of non-zero norm represents a rotation. Substitute h in (8.4) by a split quaternion polynomial C to obtain a one-parametric motion of the hyperbolic plane whose trajectories depend rationally on the motion parameter t:

$$[x] \mapsto [x(t)] = [C(t)x\overline{C}(t)]. \tag{8.6}$$

For real zeros of $C(t)$, the corresponding value of $[x(t)]$ can be defined by continuity requirements. This is called a rational motion and will be on the agenda for the remainder of this Chapter. At this point, only a simple corollary to Theorem 8.2 will be considered:

Corollary 8.1. *If $C = c_1t + c_0 \in \mathbb{S}[t]$ is a linear polynomial with $N(C) \neq 0$ and independent coefficients, then the rational motion* (8.6) *is the composition (from the left) of a fixed rotation with all rotations around a fixed center.*

Proof. If c_1 is not invertible, there exist suitable values $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ with $\alpha\delta - \beta\gamma \neq 0$ such that the leading coefficient of $\tilde{C} := (\alpha t + \beta)c_1 + (\gamma t + \delta)c_0$ is invertible. The polynomials C and $\tilde{C}$ are just different parametrizations of the same rational motion. Hence, it may assumed, without loss of generality, that c_1 is invertible and one can write $C = c_1(t + h)$ with $h = c_1^{-1}c_0$. The claim now follows, because $t + h$ is, indeed, a rotation with fixed center $[\mathrm{Im}\, t + h] = [\mathrm{Im}\, h]$ for every value of t. Note that independence of c_0 and c_1 ensures $\mathrm{Im}\, h \neq 0$, so that the center is actually defined. $\square$

8.3.2 Factorization theory

A well-known consequence of the fundamental theorem of algebra is that any polynomial

$$C = \sum_{i=0}^{d} c_i t^i, \quad c_0, c_1, \cdots, c_d \in \mathbb{C}$$

can be written as

$$C = c_d \prod_{i=1}^{d} (t - z_i),$$

with complex numbers $z_1, z_2, \cdots, z_d$. This representation is unique up to re-ordering of the factors. Left polynomials over Hamiltonian quaternions admit similar factorizations but such factorizations are not unique. The corresponding theory has been developed in the mid-twentieth century [6, 18]. To a certain extent, it carries over to dual quaternions [8, 15] and also to split quaternions [1, 13]. The main difference between polynomials and the latter algebra is the existence of non-invertible elements.

Definition 8.1. We say that the polynomial $C \in \mathbb{S}[t]$ *admits a factorization* if there exist split quaternions $c, h_1, h_2, \cdots, h_d$ such that

$$C = c(t-h_1)(t-h_2)\cdots(t-h_d). \tag{8.7}$$

Remark 8.2. If $C \in \mathbb{S}[t]$ is a split quaternion polynomial of degree d with leading coefficient c, the following hold true:

- If c is invertible, the polynomial C admits a factorization if and only if the monic polynomial $c^{-1}C$ admits a factorization.
- Unless $C\overline{C} = 0$, there exists a fractional linear parameter transformation $t \mapsto (\alpha t+\beta)(\gamma t+\delta)^{-1}$ with α, β, γ, $\delta \in \mathbb{R}$ and $\alpha\delta - \beta\gamma \neq 0$, such that the leading coefficient of

  $$D := (\gamma t+\delta)^{\deg C} C((\alpha t+\beta)(\gamma t+\delta)^{-1})$$

  is invertible. Every factorization of D also gives rise to a factorization of C and vice versa.

In what follows, the focus will be on the factorizations of monic polynomials. Due to Remark 8.2 this is only a minor loss of generality. The only missing case is that of polynomials C with vanishing norm polynomial $N(C)$.

Before delving into the intricacies of polynomial factorization over the split quaternions, a kinematic interpretation is presented. Consider a monic polynomial $C \in \mathbb{S}[t]$ with a factorization, as in (8.7) where $c = 1$. We make two assumptions:

1. The norm polynomial $N(C)$ does not vanish
2. C has no real polynomial factor of positive degree.

By Corollary 8.1, the factor $t - h_\ell$ parameterizes a rotation around $[\operatorname{Im} h_\ell]$, for any $t \in \mathbb{R}$. (Note that $[\operatorname{Im} h_\ell]$ is well-defined, because otherwise $t - h_\ell$ would be a real factor of C.) In other words, the factorization of C corresponds to the decomposition of the rational motion (8.6) into a product of hyperbolic rotations. In this sense, (8.7) describes a rational motion of an open kinematic chain of hyperbolic revolute joints.

Assume that a split quaternion polynomial C can be written as $= C'(t-h)$, with $C' \in \mathbb{S}[t]$ and $h \in \mathbb{S}$. Then

$$C\overline{C} = C'(t-h)\overline{C'(t-h)} = C'(t-h)(\overline{t-h})\,\overline{C'} = C'\overline{C'}(t-\overline{h})(t-h).$$

It is clear that a necessary condition for $t-h$ to be a right factor of C is that $M := (t-\overline{h})(t-h)$ be a quadratic factor of the norm polynomial $N(C)$.

Conversely, it is possible to compute a right factor from a quadratic factor of the norm polynomial, at least in generic cases. Its description profits from a relation between right zeros and right factors (Theorem 2 of [13]):

Theorem 8.3. *Given a polynomial $C \in \mathbb{S}[t]$ and a split quaternion $h \in \mathbb{S}$, $C(h) = 0$ holds if and only if there exists $C' \in \mathbb{S}[t]$ such that $C = C'(t-h)$.*

Given $C \in \mathbb{S}[t]$, pick a quadratic factor M of $N(C)$ and use polynomial division to compute the unique polynomials $Q, R \in \mathbb{S}[t]$, with $\deg R \leq 1$ such that $C = QM + R$. This yields

$$C\overline{C} = (QM+R)(\overline{Q}M+\overline{R}) = (Q\overline{Q}M + Q\overline{R} + R\overline{Q})M + R\overline{R}.$$

It follows that there exists $m \in \mathbb{R}$ such that $R\overline{R} = mM$. Generically (but not necessarily), m is invertible and the same is true for the leading coefficient r_1 of the remainder polynomial $R = r_1 t + r_0$. In this case, there exists a unique zero $h := -r_1^{-1} r_0$ of R, whence $R = r_1(t-h)$ and $M = m^{-1} r_1 \overline{r_1}(t-\overline{h})(t-h)$. But then, by Theorem 8.3, $t-h$ is a right factor of R. It is also a right factor of M and hence a right factor of $C = QM + R$ as well.

This observation yields an iterative procedure to compute right factors for generic polynomials $C \in \mathbb{S}[t]$:

1. Pick a quadratic factor M of $C\overline{C}$.

2. Compute the linear remainder polynomial $R = r_1 t + r_0$ when dividing C by M.

3. Compute the (generically) unique zero $h = -r_1^{-1} r_0$ of R.

4. Use polynomial division once more to compute $C' \in \mathbb{S}[t]$, such that $C = C'(t-h)$ and iterate with C' instead of C.

Note that division of polynomials in $\mathbb{S}[t]$ is possible by means of a "left" version of the Euclidean algorithm , as long as the divisor's leading coefficient is invertible. Step 3 in the above procedure is critical, as the zero need not exist and need not be unique. Moreover, at any iteration there is freedom to pick a quadratic factor M of $C\overline{C}$. Depending on the number of real roots of $C\overline{C}$, there are up to $\binom{2\deg C}{2}$ choices. In generic cases, the zero h of M is unique at any iteration and the above procedure gives all possible factorizations. Thus it holds:

Theorem 8.4. *Generically, a polynomial $C \in \mathbb{S}[t]$ whose norm polynomial has no multiple zeros (over $\mathbb{C}$) admits between $d!$ and*

$$\prod_{i=0}^{d-1} \binom{2(d-i)}{2}$$

different factorizations.

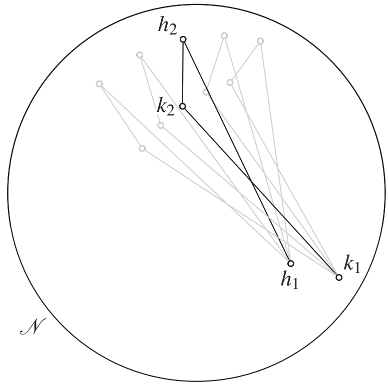


FIGURE 8.1: Hyperbolic four-bar linkage via motion factorization

As already noted in [1], there can be up to six factorizations of generic quadratic polynomials. For polynomials over the quaternions and dual quaternions without multiple factors in the norm polynomial, the number of factorizations is always $d!$.

Returning to the kinematic interpretation, different factorizations correspond to different kinematic chains but, because the respective linear factors have the same products, these factorizations describe the same motion of the distal link. This allows us to connect them and produce a closed loop kinematic structure that is capable of performing the same motion. This is illustrated in Figure 8.1 of a quadratic polynomial $C \in \mathbb{S}[t]$ with two different factorizations

$$C = (t - h_1)(t - h_2) = (t - k_1)(t - k_2).$$

It gives rise to a hyperbolic four-bar linkage with rational coupler motion [16, 21].

A rather strange example is given by a quadratic polynomial $C \in \mathbb{S}[t]$ with six factorizations. It gives rise to a linkage with six "legs", each consisting of a dyad of hyperbolic revolute joints. This can be thought of as a "four-bar linkage" with six legs although that name seems no longer appropriate. Its elementary geometry has been investigated in [16] (Figure 8.2). In particular:

- There exists a conic $\mathscr{S}$ that shares four real tangents with $\mathscr{N}$. These tangents form a complete quadrilateral and the fixed revolute joints, denoted by $H_1, K_1, \cdots, O_1$ in Figure 8.2, are its vertices. In hyperbolic geometry, they are usually called the focal points of $\mathscr{S}$.

- By reflecting the fixed revolute joints in the tangents of $\mathscr{S}$ (or in the

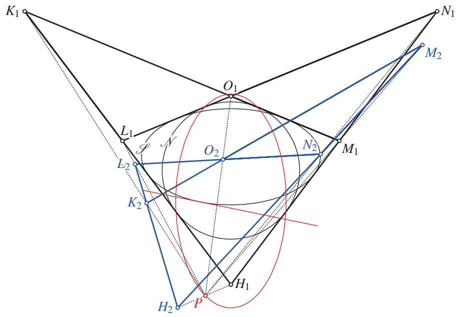


FIGURE 8.2: Hyperbolic "four-bar" linkage with six legs

points of their absolute polar conic), we obtain the possible positions of the moving vertices $H_2, K_2, \cdots, O_2$.

Six-leg four-bars do not exist in traditional hyperbolic geometry because their construction necessarily involves points in the exterior of $\mathscr{N}$. Within traditional hyperbolic geometry, the conic $\mathscr{S}$ is contained in $\mathscr{N}$ and only two focal points are real [21]. This corresponds to quadratic motion polynomials with only two factorizations.

Example 8.1. It is actually quite simple and straightforward to construct quadratic polynomials $C \in \mathbb{S}[t]$ with six factorizations. Start by picking four arbitrary linear polynomials over $\mathbb{R}$, for example

$$m_1 = t+2, \quad m_2 = t+1, \quad m_3 = t-1, \quad m_4 = t-2.$$

For $\ell, r \in \{1,2,3,4\}$ and $\ell < r$ we then define $M_{\ell r} := m_\ell m_r$. For two of these quadratic polynomials with complementary indices, say M_{12} and M_{34}, we compute $h_1, h_2 \in \mathbb{S}$ such that

$$M_{12} = (t-h_1)(t-\overline{h_1}) \quad \text{and} \quad M_{34} = (t-h_2)(t-\overline{h_2}).$$

This amounts to solving a quadratic equation and yields a two-parametric variety of solutions. These degrees of freedom are then used to find a rational example:

$$h_1 = -\frac{3}{2} + \mathbf{i} + \frac{1}{2}\mathbf{j} + \mathbf{k}, \quad h_2 = \frac{1}{2}(3+\mathbf{i}+\mathbf{j}-\mathbf{k}).$$

The quadratic polynomial

$$C := (t-h_1)(t-h_2) = t^2 - \frac{1}{2}(3\mathbf{i}+2\mathbf{j}+\mathbf{k})t - \frac{1}{2}(6-3\mathbf{i}-2\mathbf{j}-5\mathbf{k})$$

has the norm polynomial $N(C) = m_1 m_2 m_3 m_4$ and, consequently, allows for six factorizations:

$$\begin{aligned}
C &= (t - \tfrac{1}{10}\mathbf{i} + \tfrac{1}{10}\mathbf{j} + \tfrac{1}{2}\mathbf{k} - \tfrac{3}{2})(t - \tfrac{7}{5}\mathbf{i} - \tfrac{11}{10}\mathbf{j} - \mathbf{k} + \tfrac{3}{2}) \\
&= (t + \mathbf{i} + \tfrac{3}{2}\mathbf{j} + \mathbf{k} - \tfrac{1}{2})(t - \tfrac{5}{2}\mathbf{i} - \tfrac{5}{2}\mathbf{j} - \tfrac{3}{2}\mathbf{k} + \tfrac{1}{2}) \\
&= (t - \mathbf{i} + \mathbf{j} - \mathbf{k})(t - \tfrac{1}{2}\mathbf{i} - 2\mathbf{j} + \tfrac{1}{2}\mathbf{k}) \\
&= (t + \tfrac{7}{2}\mathbf{i} + 4\mathbf{j} + \tfrac{1}{2}\mathbf{k})(t - 5\mathbf{i} - 5\mathbf{j} - \mathbf{k}) \\
&= (t - \tfrac{1}{2}\mathbf{i} + \tfrac{1}{2}\mathbf{j} - \tfrac{3}{2}\mathbf{k} + \tfrac{1}{2})(t - \mathbf{i} - \tfrac{3}{2}\mathbf{j} + \mathbf{k} - \tfrac{1}{2}) \\
&= (t - \mathbf{i} - \tfrac{1}{2}\mathbf{j} - \mathbf{k} + \tfrac{3}{2})(t - \tfrac{1}{2}\mathbf{i} - \tfrac{1}{2}\mathbf{j} + \tfrac{1}{2}\mathbf{k} - \tfrac{3}{2}).
\end{aligned}$$

The factorization of split-quaternion polynomials in non-generic cases goes beyond the scope of this discussion. We confine ourselves to one example to demonstrate that split quaternion polynomials can have infinitely many factorizations. Needless to say that this is only possible under very special circumstances.

Example 8.2. The polynomial $C = (1+\mathbf{j})t^2 + \mathbf{i} - \mathbf{k} \in \mathbb{S}[t]$ allows the factorizations

$$C = (1+\mathbf{j})(t - h_0 - h_1\mathbf{i} - h_2\mathbf{j} - h_3\mathbf{k})(t - k_0 - k_1\mathbf{i} - k_2\mathbf{j} - k_3\mathbf{k}),$$

where

$$h_0 = -\frac{2h_2k_0 - w}{2k_0}, \quad h_1 = \frac{2h_3k_0 + 1}{2k_0}, \quad k_2 = \frac{-2k_0^2 - w}{2k_0}, \quad k_3 = \frac{2k_0k_1 + 1}{2k_0}$$

and $w = \sqrt{-1 - 4k_0k_1}$. In order to ensure that w is real the inequality $1 + 4k_0k_1 \le 0$ has to be satisfied. Otherwise, the values of h_2, h_3, k_0 and k_1 can be chosen arbitrarily. This example violates our general assumption that the norm polynomial of C be non-zero.

Example 8.3. The polynomial $C = (t+1)(t+\mathbf{j}) \in \mathbb{S}[t]$ allows the factorizations

$$\begin{aligned}
C &= (t + 1 + \lambda(\mathbf{i}+\mathbf{k}))(t + \mathbf{j} - \lambda(\mathbf{i}+\mathbf{k})) \\
&= (t + \mathbf{j} - \lambda(\mathbf{i}-\mathbf{k}))(t + 1 + \lambda(\mathbf{i}-\mathbf{k})),
\end{aligned}$$

where $\lambda \in \mathbb{R}$. This example violates the assumption that C has no real factors.

For more factorization results on split quaternion polynomials, we refer the reader to [1], where zeros of quadratic polynomials are considered. Because of Theorem 8.3, this is actually the topic of this Section in disguise. Some more factorization results for split quaternion polynomials of arbitrary degree can be found in [13].

8.4 Bond theory

Bond theory was introduced in [9] as a tool for the analysis and classification of movable closed kinematic chains with revolute joints [7, 14]. It was soon extended to other mechanical structures [3, 4, 5, 17].

The definition of a bond is rather abstract and algebraic. One has to construct a suitable closure of a coordinate description of the underlying motion group (a kinematic space) and then intersect the linkage's complete configuration variety, with the set of newly added point set. The thus obtained "bond points" of the configuration variety have been shown to bear a lot of information on the underlying mechanism. They exhibit a degenerate kinematic behavior. Therefore, they are not amenable to lucid visualization and evade geometric intuition, at least in Euclidean geometry. Here, we are going to study bonds in planar hyperbolic geometry along the lines of the original paper [9]. In a certain way, this is a rather trivial task: The original theory applies to SO and is algebraically equivalent to the planar hyperbolic motion group over the complex numbers. Since bond points in SO are complex anyway, the complete theory carries over to the hyperbolic setting without any substantial changes and new proofs are not required. Moreover, [9] studies closed loop linkages with only revolute joints and one degree of freedom. In our planar hyperbolic setting, the only linkages falling into this category are four-bar linkages. In spite of all these simplifications, the hyperbolic case is still capable of providing some new insight. In particular, bond points need not be complex, whence they can easily be visualized and their kinematic degeneracy become obvious.

In order not to make the further development of planar hyperbolic bond theory an unnecessarily badly motivated exercise, we briefly summarize important properties of bonds for closed linkages with revolute joints and a one-parametric mobility, in Euclidean three-space [7, 9, 14]. We denote the linkage's joints by j_1, j_2, $\cdots$ j_n, adopt the convention $j_k = j_\ell$ if $n \equiv k \mod \ell$ and make the natural assumption that j_k and j_{k+1} are always neighbors. In order to eliminate trivial cases, we assume that every joint actually moves during the motion of the linkage.

- A bond is a pair of conjugate complex points in the linkage's compactified configuration curve.

- For a bond, the notion of "connecting" two joints with a certain multiplicity can be defined.

- The pattern of all connections with bonds is subject to combinatorial constraints: At least one and at most four connections emanate from every joint and no two neighboring joints are connected. Further restrictions are known, but are more complicated to formulate.

- The existence of connections between joints has implications on the geometry of revolute axes. If, for example, the joints j_k and j_{k+2} are connected, then there exists a revolute joint j_0 such that the four-bar linkage $(j_0, j_k, j_{k+1}, j_{k+2})$ is movable — even if the original joint axes are skew. The resulting spatial four-bar linkage is known under the name of *Bennett linkage* and its axis geometry is quite special. Existence of a joint connecting j_k and j_{k+3} also has implications that were used to classify linkages with configuration curves of maximal genus in [7].

- From the connection combinatorics, the degree (in the dual quaternion model of spatial kinematics) of relative motions between any joint pair can be directly read off.

Most of the properties of bonds stated above are not relevant for planar hyperbolic kinematics, but can easily be transferred to the kinematics of Minkowski's three space, where planar hyperbolic kinematics is embedded as kinematics of the Minkowski unit sphere.

8.4.1 Definition of hyperbolic bonds

Consider a four-bar linkage in the hyperbolic plane. In some initial configuration, its revolute joints are located at points $[h_1]$, $[h_2]$, $[h_3]$, $[h_4]$ and, assuming none of these points is ideal, it can be assumed that

$$h_\ell \overline{h_\ell} = \pm 1. \tag{8.8}$$

The sign depends on whether $[h_1]$ is inside $(+)$ or outside $(-)$ the null circle. The rotation about the ℓ^{th} joint is parameterized by $t_\ell - h_\ell$ where $t_\ell \in \mathbb{R} \cup \{\infty\}$. The four-bar's *closure equation* is

$$E(t_1,t_2,t_3,t_4) := (t_1 - h_1)(t_2 - h_2)(t_3 - h_3)(t_4 - h_4) \in \mathbb{R} \setminus \{0\}. \tag{8.9}$$

Define the *configuration curve* as

$$K := \{(t_1,t_2,t_3,t_4) \in (\mathbb{R} \cup \{\infty\})^4 \mid E(t_1,t_2,t_3,t_4) \in \mathbb{R} \setminus \{0\}\},$$

and denote its Zariski closure by $K_{\mathbb{C}}$. The set of *bond points* is[1]

$$B := \{(t_1,t_2,t_3,t_4) \in K_{\mathbb{C}} \mid (t_1 - h_1)(t_2 - h_2)(t_3 - h_3)(t_4 - h_4) = 0\}. \tag{8.10}$$

[1]For reasons of simplicity, it is tacitly assumed that none of the bond points is a singularity of $K_{\mathbb{C}}$. If this were the case, one has to replace $K_{\mathbb{C}}$ by a normalization.

Since $K_{\mathbb{C}}$ has dimension one, B is finite and non-empty. A bond point might be thought of as an element of the Zariski closed configuration curve, for which the closure equation (8.9) is actually violated.

8.4.2 Properties and examples

Properties of bonds in Euclidean or spherical geometry [9] effortlessly carry over to our setting, because spherical and universal hyperbolic geometry over the complex numbers are equivalent. In this Section some of these properties are reviewed and examples are presented and discussed. In doing so, significant benefit will be derived from the fact that bond points in hyperbolic geometry can be real, and thus allow visualization.

Example 8.4. The bond points of the four-bar linkage (h_1,h_2,h_3,h_4) with

$$h_1=\mathbf{i},\quad h_2=\tfrac{1}{3}w(\mathbf{i}+2\mathbf{j}),\quad h_3=-\tfrac{1}{3}(\mathbf{i}+3\mathbf{j}+\mathbf{k}),\quad h_4=\tfrac{1}{29}(20\mathbf{i}+35\mathbf{j}+4\mathbf{k}),$$

where $w=\sqrt{3}$ are:

$$\begin{gathered}
(-\tfrac{110}{23}+\tfrac{73}{23}w,1,\tfrac{41}{3}-\tfrac{22}{3}w,1),\quad (-\tfrac{110}{23}-\tfrac{73}{23}w,-1,\tfrac{41}{3}+\tfrac{22}{3}w,1),\\
(-30-17w,1,-\tfrac{43}{69}+\tfrac{16}{69}w,-1),\quad (-30+17w,-1,-\tfrac{43}{69}-\tfrac{16}{69}w,-1),\\
(\mathrm{i},(\tfrac{8}{15}+\tfrac{1}{15}\mathrm{i})w,1,\tfrac{35}{29}-\tfrac{16}{29}\mathrm{i}),\quad (-\mathrm{i},(\tfrac{8}{15}-\tfrac{1}{15}\mathrm{i})w,1,\tfrac{35}{29}+\tfrac{16}{29}\mathrm{i}),\\
(\mathrm{i},(-\tfrac{2}{3}-\tfrac{1}{3}\mathrm{i})w,-1,-\tfrac{152}{145}-\tfrac{11}{145}\mathrm{i}),\quad (-\mathrm{i},(-\tfrac{2}{3}+\tfrac{1}{3}\mathrm{i})w,-1,-\tfrac{152}{145}+\tfrac{11}{145}\mathrm{i}).
\end{gathered}\tag{8.11}$$

Observe that two elements of $\{1,-1,\mathrm{i},-\mathrm{i}\}$ occur in every bond quadruple. This is no coincidence.

Theorem 8.5. *If $b=(t_1,t_2,t_3,t_4)$ is a bond point, then there are at least two coordinates ℓ, m such that t_ℓ, $t_m\in\mathbb{I}:=\{1,-1,\mathrm{i},-\mathrm{i}\}$.*

Proof. For any $\ell\in\{1,2,3,4\}$ one has

$$(t_\ell-h_\ell)(t_\ell-\overline{h}_\ell)=t_\ell^2-(h_\ell+\overline{h}_\ell)t_\ell+h_\ell\overline{h}_\ell=t^2\pm1,$$

by virtue of the normalization condition (8.8) and because h_ℓ is assumed to be vectorial. Moreover, for any bond point , the equation $E(t_1,t_2,t_3,t_4)=0$ is fulfilled. This implies

$$0=E(t_1,t_2,t_3,t_4)\overline{E(t_1,t_2,t_3,t_4)}=(t_1^2\pm1)(t_2^2\pm1)(t_3^2\pm1)(t_4^2\pm1).$$

Hence, there is at least one coordinate ℓ such that $t_\ell\in\mathbb{I}$. If $\ell=1$, multiplying the bond equation $E(t_1,t_2,t_3,t_4)=0$ from the right times $(t_4-\overline{h_4})(t_3-\overline{h_3})(t_2-\overline{h_2})$ one obtains

$$(t_1-h_1)(t_2^2\pm1)(t_3^2\pm1)(t_4^2\pm1)=0.\tag{8.12}$$

Because $t_1 - h_1 \neq 0$, there exists a second index m with $t_m \in \mathbb{I}$. If $\ell = 2$, the same argument applies after multiplication times $t_1 - \overline{h_1}$ from the left and by $(t_4 - \overline{h_4})(t_3 - \overline{h_3})$. The cases $\ell = 3$ and $\ell = 4$ are similar. □

Generically, precisely two coordinates t_ℓ and t_m of a bond point b are in the set $\mathbb{I}$. In this case it is said that b "connects" the joints of index ℓ and m. Moreover, the proof of Theorem 8.5 shows that generic bonds that connect two joints come in pairs corresponding to the two solutions of (8.12). Paired bonds connect the same joints. In (8.11), paired bond points are written in the same respective lines.

Example 8.5. We consider another example of a four-bar linkage. In the notation of Section 8.3, start with the quadratic motion polynomial $C = (t - h_1)(t - h_2)$, where

$$h_1 = \mathbf{i} \quad \text{and} \quad h_2 = \mathbf{i} + \mathbf{j} + 2\mathbf{k}.$$

The second factorization $C = (t - k_1)(t - k_2)$ with

$$k_1 = 11\mathbf{i} + 5\mathbf{j} + 10\mathbf{k} \quad \text{and} \quad k_2 = -9\mathbf{i} - 4\mathbf{j} - 8\mathbf{k}$$

implies

$$(t - h_1)(t - h_2)(t - h_3)(t - h_4) \in \mathbb{R}[t], \tag{8.13}$$

where $h_3 = \overline{k_2}$ and $h_4 = \overline{k_1}$. Expression (8.13) can be interpreted as the closure equation of a four-bar linkage , which is fulfilled for any $t \in \mathbb{R}$ (and also for $t = \infty$). In order to compute the bonds, we first compute the zero set $\{\pm 2, \pm \mathrm{i}\}$ of (8.13). With

$$n_\ell := \frac{1}{|h_\ell \overline{h_\ell}|^{1/2}} \quad \text{for } \ell \in \{1,2,3,4\},$$

that is, $n_1 = n_3 = 1$ and $n_2 = n_4 = \frac{1}{2}$, the corresponding bond points are

$$\pm 2(n_1, n_2, n_3, n_4) = (\pm 2, \pm 1, \pm 2, \pm 1), \quad \pm \mathrm{i}(n_1, n_2, n_3, n_4) = (\pm \mathrm{i}, \pm \tfrac{1}{2}\mathrm{i}, \pm \mathrm{i}, \pm \tfrac{1}{2}\mathrm{i}).$$

This is, however, only half of the story. The reason is that the four-bar linkage's configuration curve has a second component, not parameterized by C, that also has to be taken into account. The closure equation of this component is given by

$$(t - \ell_1)(t - \ell_2)(t - \ell_3)(t - \ell_4) \in \mathbb{R}[t],$$

where

$$\ell_1 = h_1, \quad \ell_2 = \tfrac{2}{3} h_2, \quad \ell_3 = -h_3, \quad \ell_4 = -\tfrac{2}{3} h_4,$$

so that, in fact, $[h_1] = [\ell_1]$, $[h_2] = [\ell_2]$, $[h_3] = [\ell_3]$, $[h_4] = [\ell_4]$ in the two respective configurations corresponding to $t = \infty$. The bonds of the second motion component are

$$\pm(\tfrac{4}{3}, 1, \tfrac{4}{3}, 1) \quad \pm(\mathrm{i}, \tfrac{3}{4}\mathrm{i}, \mathrm{i}, \tfrac{3}{4}\mathrm{i}).$$

As expected, precisely two entries are always in the set $\{\pm 1, \pm \mathrm{i}\}$.

A more profound theory of bonds is rather involved. One has to take into account connection multiplicities, non-generic bonds need to be properly paired and the notion of connecting joints needs to be defined. Moreover, the computation of bonds also has to take into account the parameter value $t_n = \infty$. All of this is important for a consistent mathematical theory. Here it suffices to mention that these concepts can be derived from the vanishing orders at bond points of the *coupling maps* $f_{m,n}$, which are defined on the configuration curve's normalization and map a regular point (t_1, t_2, t_3, t_4) to

$$f_{m,n} := (t_{m+1} - h_{m+1})(t_{m+2} - h_{m+2}) \cdots (t_{n-1} - h_{n-1})(t_n - h_n). \qquad (8.14)$$

Here, indices are reduced modulo four and the value of $f_{m,n}$ at singular points is defined by continuity.

Remark 8.1. It is easy to see that some of the maps $f_{i,j}$ indeed vanish at generic bond points. Consider for example a bond point $b := (t_1, t_2, t_3, t_4)$ with $t_2 = t_4 \in \{\pm 1, \pm \mathrm{i}\}$ and t_1, $t_3 \notin \{\pm 1, \pm \mathrm{i}\}$. By (8.10) it is clear that

$$(t_1 - h_1)(t_2 - h_2)(t_3 - h_3)(t_4 - h_4) = 0.$$

Multiplying by $t_1 - \overline{h_1}$ from the left it follows tat

$$(t_1^2 - h_1\overline{h_1})(t_2 - h_2)(t_3 - h_3)(t_4 - h_4) = 0.$$

Since $t_1^2 - h_1\overline{h_1}$ is a non-zero scalar, $f_{2,4}$ vanishes at b. The same is true for $f_{4,2} = (t_4 - h_4)(t_1 - h_1)(t_2 - h_2)$.

Example 8.6. Example 8.4 is revisit here to consider the coupling maps $f_{m,m+2}$ at the bond $b = (t_1, t_2, t_3, t_4)$, with

$$t_1 = -30 - 17w, \quad t_2 = 1, \quad t_3 = -\tfrac{43}{69} + \tfrac{16}{69}w, \quad t_4 = -1$$

and $w = \sqrt{3}$. A straightforward computation confirms that

$$f_{2,4}(t_1, t_2, t_3, t_4) = (t_2 - h_2)(t_3 - h_3)(t_4 - h_4) = 0 \quad \text{and}$$
$$f_{4,2}(t_1, t_2, t_3, t_4) = (t_4 - h_4)(t_1 - h_1)(t_2 - h_2) = 0$$

but $f_{1,3}(t_1, t_2, t_3, t_4) \neq 0$ and $f_{3,1}(t_1, t_2, t_3, t_4) \neq 0$.

Example 8.7. A simple but, as far as bonds are concerned, rather atypical four-bar linkage, is obtained from the motion polynomial $C = (t-h_1)(t-h_2) = (t-k_1)(t-k_2)$, where

$$h_1 = \mathbf{i}, \quad h_2 = \mathbf{i}+\mathbf{j}+\mathbf{k}, \quad k_1 = -\mathbf{i}-\mathbf{j}-\mathbf{k}, \quad k_2 = 3\mathbf{i}+2\mathbf{j}+2\mathbf{k}.$$

Its configuration curve has precisely three irreducible components, which are parameterized by

$$(t,t,t,t), \ (\infty,t,\infty,-t) \text{ and } \ (t,0,-t,0) \quad t \in \mathbb{R} \cup \{\infty\}. \tag{8.15}$$

Setting $h_3 := \overline{k_2}$ and $h_4 := \overline{k_1}$, this can be verified by the following computation:

$$\begin{gathered} (t-h_1)(t-h_2)(t-h_3)(t-h_4) = t^4 - 1, \\ (t-h_2)(-t-h_4) = -t^2+1, \quad (t-h_1)h_2(-t-h_3)h_4 = -t^2-1. \end{gathered} \tag{8.16}$$

Because $h_1\overline{h_1} = k_2\overline{k_2} = 1$ and $h_2\overline{h_2} = k_1\overline{k_1} = -1$, the bond points are obtained from the zeros of the polynomials in (8.16):

$$\begin{gathered} b_\ell := (\ell,\ell,\ell,\ell) \text{ for } \ell \in \{\pm 1, \pm \mathrm{i}\}, \\ b'_\ell := (\infty,\ell,\infty,-\ell) \text{ for } \ell \in \{\pm 1\}, \quad b''_\ell := (\ell,0,-\ell,0) \text{ for } \ell \in \{\pm \mathrm{i}\}. \end{gathered}$$

The bonds b_ℓ are not generic. However, among all coupling functions, only $f_{2,4}$ and $f_{4,2}$ vanish at b_1 and b_{-1}, while only $f_{1,3}$ and $f_{3,1}$ vanish at b_{i} and $b_{-\mathrm{i}}$. In the language of bond theory in the sense of the original paper [9], one would say that "$b_{\pm 1}$ connect joints of index 2 and 4" while "$b_{\pm \mathrm{i}}$ connect joints of index 1 and 3."

Bonds in Euclidean or spherical geometry are always complex. As demonstrated by several of the above examples, this is not the case in hyperbolic geometry. This Section is concluded by visualizing the bonds of Example 8.7. There is a simple geometric explanation for the three components of the configuration curve given in (8.15). Observe that

$$Q(h_1,h_2) = Q(h_2,k_2) = Q(k_2,k_1) = Q(k_1,h_1) = 2.$$

Hence, all four sides of the linkage are equal. The two motion components in (8.15), where two joint parameters are fixed, correspond to configurations where two joints coincide. The motion itself is the rotation about this common joint. Using the data of Example 8.7, these are visualized in Figure 8.3.

In order to better understand the kinematic behavior at bond points, we study point trajectories and their values (or limit) at bonds. In the bottom right drawing, several trajectories of the first motion component are visualized. Four special points can be identified, namely $[\mathbf{i}-\mathbf{j}+\mathbf{k}]$, $[\mathbf{i}+\mathbf{j}-\mathbf{k}]$,

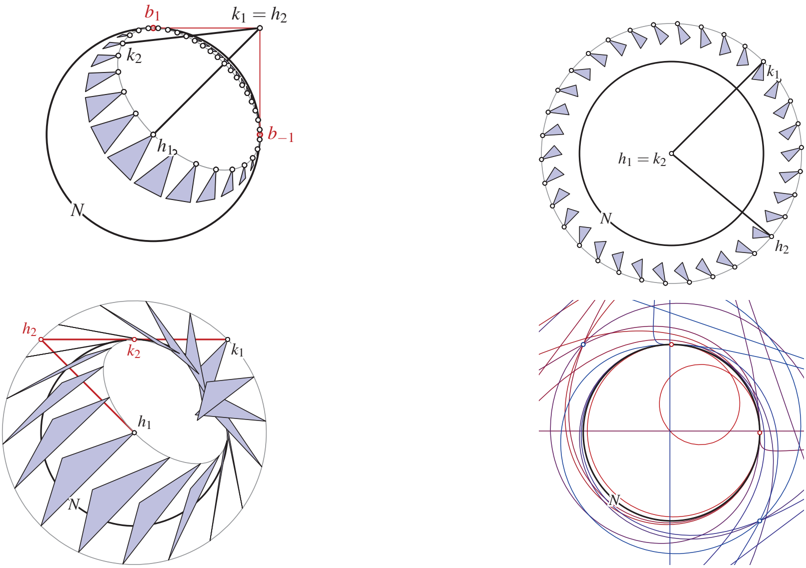


FIGURE 8.3: Motion modes of a hyperbolic four-bar linkage

$[\mathbf{i}+\mathbf{j}]$, $[\mathbf{i}+\mathbf{k}]$, that are apparently incident with many trajectories. Indeed, these points are closely related to the undefined kinematic behavior of the four-bar linkage at bonds. A straightforward computation yields the parametric trajectory of the point $[x] := [x_1\mathbf{i}+x_2\mathbf{j}+x_3\mathbf{k}]$ as $[y] = [Cx\overline{C}]$, where $y = y_1\mathbf{i}+y_2\mathbf{j}+y_3\mathbf{k}$ and

$$
\begin{aligned}
y_1 &= (t^2+1)(x_1t^2-(2x_2-2x_3)t+3x_1-2x_2-2x_3),\\
y_2 &= (t+1)(x_2t^3-(2x_1+x_2-4x_3)t^2+(8x_1-5x_2-6x_3)t-2x_1+x_2+2x_3),\\
y_3 &= (t-1)(x_3t^3+(2x_1-4x_2+x_3)t^2+(8x_1-6x_2-5x_3)t+2x_1-2x_2-x_3).
\end{aligned}
\tag{8.17}
$$

Evaluation at $t = 1$ (one of the real bond points), generically yields $[y](1) = [8(x_1-x_2)(\mathbf{i}+\mathbf{j})]$. The case $x_1 = x_2$ requires additional considerations. Here, (8.17) simplifies to

$$
\begin{aligned}
y_1 &= (t^2+1)(x_1t^2-(2x_1-2x_3)t+x_1-2x_3),\\
y_2 &= (t+1)(x_1t^3-(3x_1-4x_3)t^2+(3x_1-6x_3)t-x_1+2x_3),\\
y_3 &= (t-1)(x_3t^3+(-2x_1+x_3)t^2+(2x_1-5x_3)t+x_3)
\end{aligned}
$$

and evaluation at $t = 1$ yields an undefined point. However, in the limit, one has:

$$\lim_{t \to 1} \frac{y_2}{y_1} = 1, \quad \lim_{t \to 1} \frac{y_3}{y_1} = -1.$$

This suggests the consideration of $[\mathbf{i}+\mathbf{j}-\mathbf{k}]$ as limiting point of all points on the line given by $x_1 = x_2$ for $t \to 1$ and nicely fits with the intuition provided by the bottom left drawing in Figure 8.3. Observe that two adjacent edges of the four-bar linkage are aligned and null at bond points .

8.5 Summary

We have presented and reviewed some results on the factorization of split quaternion polynomials and extended the theory of bonds to planar hyperbolic kinematics. For algebraic and geometric reasons, both theories exhibit interesting properties that differ from the (real) Euclidean or spherical case and that make them worthwhile studying in this non-Euclidean context. Most notably, the number of factorizations in hyperbolic geometry can be significantly larger, even in generic cases. A quadratic polynomial over Hamiltonian quaternions allows, at most, two different factorizations, while six factorizations are possible for split quaternion polynomials (Example 8.1, Figure 8.2). An interesting property of bond theory in hyperbolic kinematics is the possibility of real bond points. For the first time, this article features an investigation and also a visualization of the kinematic behavior of a mechanism at such a point.

Bibliography

[1] M. Abrate, *Quadratic Formulas for Generalized Quaternions*, J. Algebra Appl. **8:3** (2009), 289-306.

[2] E. Casas-Alvero, *Analytic Projective Geometry*, European Mathematical Society (2014).

[3] M. Gallet, G. Nawratil, J. Schicho, *Bond theory for pentapods and hexapods*, J. Geometry **106:2** (2015), 211-228.

[4] M. Gallet, G. Nawratil, J. Schicho, *Liaison Linkages*, J. Symbolic Comput. **79:1** (2017), 65-98.

[5] M. Gallet, G. Nawratil, J. Schicho, J. M. Selig, *Mobile Icosapods*, Advances in Appl. Math. **88** (2017), 1-25.

[6] B. Gordon, T. S. Motzkin, *On the Zeros of Polynomials Over Division Rings*, Trans. Amer. Math. Soc. **116** (1965), 218-226.

[7] G. Hegedüs, Z. Li, J. Schicho, H.P. Schröcker, *The Theory of Bonds II: Closed 6R Linkages With Maximal Genus*, J. Symbolic Comput. **68:2** (2015), 167-180. DOI 10.1016/j.jsc.2014.09.035.

[8] G. Hegedüs, J. Schicho, H.P. Schröcker, *Factorization of Rational Curves in the Study Quadric and Revolute Linkages*, Mech. Mach. Theory **69:1** (2013), 142-152. DOI 10.1016/j.mechmachtheory.2013.05.010.

[9] G. Hegedüs, Z. Li, J. Schicho, H.P. Schröcker, *The Theory of Bonds: A New Method for the Analysis of Linkages*, Mech. Mach. Theory **70** (2013), 407-424. DOI 10.1016/j.mechmachtheory.2013.08.004.

[10] G. Hegedüs, J. Schicho, H.P. Schröcker, *A non-unique factorization theorem and its engineering applications*, In: A. Cerić, S. Lukušić (eds.) Proceedings of Future Trends in Civil Engineering (2014), 123-140. University of Zagreb, Faculty of Civil Engineering.

[11] G. Hegedüs, J. Schicho, H.P. Schröcker, *Four-Pose Synthesis of AngleSymmetric 6R Linkages*, ASME J. Mechanisms Robotics **7:4** (2015). DOI 10.1115/1.4029186.

[12] Z. Li, T.D. Rad, J. Schicho, H.P. Schröcker, *Factorization of Rational Motions: A Survey with Examples and Applications*, In Proceedings of the 14th IFToMM World Congress (2015). DOI 10.6567/IFToMM.14TH.WC.OS2.013.

[13] Z. Li, D.F. Scharler, H.P. Schröcker, *Factorization Results for Left Polynomials in Some Associative Real Algebras: State of the Art, Applications and Open Questions*, J. Comput. Appl. Math. **349**, 508-522. DOI 10.1016/j.cam.2018.09.045.

[14] Z. Li, J. Schicho, H.P. Schröcker, *A Survey on the Theory of Bonds*, IMA J. Math. Control Inform. **35:1** (2018), 279-295. DOI 10.1093/imamci/dnw048.

[15] Z. Li, J. Schicho, H. P. Schröcker, *Factorization of Motion Polynomials*, Accepted for publication in J. Symbolic Comp. (2018). DOI 10.1016/j.jsc.2018.02.005.

[16] Z. Li, J. Schicho, H. P. Schröcker, *The Geometry of Quadratic Quaternion Polynomials in Euclidean and Non-Euclidean Planes*, In Proceedings of the 18th International Conference on Geometry and Graphics (2019), 298-309. Springer International Publishing, Cham.

[17] G. Nawratil, J. Schicho, *Duporcq Pentapods*, ASME J. Mechanism Robotics **9:1** (2017), 011001 (7 pages).

[18] I. Niven, *Equations in Quaternions*, Amer. Math. Monthly **48:10** (1941), 654-661.

[19] O. Ore, *Theory of Non-Commutative Polynomials*, Annh. of Math. (2) **34:3** (1933), 480-508.

[20] T. D. Rad, H.P. Schröcker, *Optimal Synthesis of Overconstrained 6R Linkages by Curve Evolution*, In Computational Kinematics. Springer (2017). DOI 10.1007/978-3-319-60867-9.

[21] A. U. Schoger, *Koppelkurven und Mittelpunktskegelschnitte in der hyperbolischen Ebene,* J. Geom. **57:1-2** (1996), 160-176.

[22] N. Wildberger, *Universal Hyperbolic Geometry II: A Pictorial Overview*, KoG **14** (2010), 3-24.

[23] N. Wildberger, *Universal Hyperbolic Geometry III: First Steps in Projective Triangle Geometry*, KoG **15** (2011), 25-49.

[24] N. Wildberger, *Universal Hyperbolic Geometry I: Trigonometry*, Geom. Dedicata **163** (2013), 215-274. DOI 10.1007/s10711-012-9746-9.

Chapter 9

Sydpoints and parabolas in relativistic linear algebra and universal hyperbolic geometry

A. H. Alkhaldi
Department of Mathematics, King Khalid University, Abha, Saudi Arabia,
E-mail: ahalkhaldi@kku.edu.sa

N. J. Wildberger
School of Mathematics and Statistics, University of New South Wales Sydney Australia,
E-mail: n.wildberger@unsw.edu.au

We give an introduction to a completely algebraic form of hyperbolic geometry, which extends to general fields and connects naturally to relativistic physics. It also allows remarkable new constructions, which typically have meaning also outside the usual light-cone. In particular, we describe the role of *sydpoints* and of *twin circles*, which allow us to extend hyperbolic triangle geometry to non-classical triangles. We also connect these ideas with the modern theory of the *parabola* in hyperbolic geometry.

9.1 Introduction

In this Chapter we present an overview of an exciting new algebraic approach to hyperbolic geometry, which extends the subject beyond the light cone of relativistic physics, and allows it to be built from arbitrary fields, not necessarily of characteristic two. In particular we look at the new theory of sydpoints and their connections with the hyperbolic parabola.

Universal Hyperbolic Geometry (UHG), first set out in Wildberger [44], is the projective study of a distinguished conic, called the *absolute* in classical geometry, and the *null conic* in UHG. Typically, the null conic can be taken to be the unit circle, from which the interior of the circle is then essentially the Beltrami Klein projective model of classical hyperbolic geometry, but with the concepts of distance and angle replaced by algebraic equivalents called *quadrance* and *spread*. These terms come originally from rational trigonometry, introduced a few years earlier in [42].

From this algebraic point of view, points both inside and outside of the null circle are part of the geometry, which is no longer homogeneous and supports both a Riemannian geometry inside the disk and a Lorentzian geometry outside the disk. We connect naturally to relativistic geometry by projectively viewing a three-dimensional vector space with a Minkowskian inner product on it, this is also in the direction of Ungar's work [38] and [39]. The algebraic aspect allows theorems to apply uniformly, which in physics language unifies the geometry inside and outside the light cone of special relativity.

Since its inception the subject has developed to include a rich hyperbolic triangle geometry [46] and to provide many new insights into conics. UHG provides not only new and wider understanding of existing subjects; it also brings entirely new concepts and results into life. In particular we describe in this Chapter how the classical notion of midpoints of a side, which typically come in pairs when they exist, can be extended to incorporate *sydpoints*, which connect points both interior and exterior to the null conic.

This Chapter is intended to be mostly an expository introduction to this rich new form of metrical geometry. We first review classical hyperbolic geometry, then introduce universal hyperbolic geometry, (UHG) first pictorially via projective geometry then analytically using projective linear algebra. We then look in more detail at how the idea of sydpoints extends the notion of midpoints and at how sydpoints relate to the hyperbolic parabola.

9.2 Classical hyperbolic geometry

In an effort to comprehend Euclid's axiomatic basis for geometry, Lobachevsky, Bolyai and Gauss discovered the concept of hyperbolic geometry by the first half of the 19th century. Hyperbolic geometry is a non-Euclidean geometry which goes beyond the parallel postulate of Euclid [8], [18], [19]. As Milnor states in [26], Lobachevsky was the first mathematician to publish on hyperbolic geometry in 1830; in particular, he showed the existence of a natural unit distance in this new geometry. In 1832, another work on non-Euclidean geometry was published independently by Bolyai, while Gauss revealed that he had studied the subject some years earlier. The theory was more fully developed by Beltrami, Klein and Poincaré, who introduced models which made this geometry more explicit ([34], [35]).

In the 20-th century, Einstein and Minkowski realized that an understanding of physical time and space necessitated the study of more general geometries than both that of Riemann and non-Euclidean geometry. With the work of Thurston and others, it has been recognized that the negatively-curved geometries, of which hyperbolic geometry is a prototype, are in some sense generic forms of geometry, and crucial for low-dimensional topology, see for example [10].

Non-Euclidean geometries have also applications to different fields such as complex variables, analytic number theory, harmonic analysis, Lie theory, infinite discrete groups (following the more recent work of Gromov) and even optics. Although Euclidean geometry has a single standard model, various models may describe hyperbolic geometry, most notably the Beltrami-Poincaré disk and the upper half plane model, the Beltrami-Klein projective model, and the hyperboloid model. These models differ from each other in certain aspects and some properties are more evident in one rather than the other, but they were all essentially discovered by Eugenio Beltrami in the second half of the 19th century. The Beltrami-Poincaré models naturally connect to complex analysis, while the Beltrami-Klein model connects naturally to the projective geometry and to more general Cayley-Klein geometries.

9.2.1 The Beltrami-Poincaré model

In this model the underlying space is the interior of the unit circle $\mathscr{C}$: $x^2+y^2=1$ in the complex plane $\mathbb{C}$, that is

$$\zeta = \{(x+iy : x^2+y^2 < 1\}.$$

The points on the unit circle $\mathscr{C}$ itself are assumed to be points ”at infinity” rather than being part of the hyperbolic plane: these points are also referred to as *ideal points*, *omega points*, *vanishing points or null points*–whatever their name, they still play a key role in the theory. Lines of ζ are represented by either circular arcs, which are parts of circles orthogonal to $\mathscr{C}$, or Euclidean lines which are diameters of $\mathscr{C}$. The hyperbolic distance between any two given hyperbolic points a and b in this model is defined in terms of the complex structure by

$$d(a,b) = \tanh^{-1}\left(\left|\frac{b-a}{1-\overline{a}b}\right|\right).$$

Measuring angles between hyperbolic lines in this model is the same as measuring Euclidean angles, making it a conformal model of the hyperbolic plane [33]. So in order for this theory to work we need an underlying theory of "real numbers", and "transcendental functions", depending of course on "infinite processes". With universal hyperbolic geometry, this restriction is avoided, opening up the subject to more general fields, including finite fields.

9.2.2 The Beltrami-Klein model

The underlying space in the Beltram-Klein projective model is the same as that of the previous model: the open disk ζ in $\mathbb{C}$. However, lines in the Beltrami-Klein model are actual Euclidean lines instead of circular arcs. The hyperbolic distance between two points a_1 and a_2 is given by

$$d(a_1, a_2) = \frac{1}{2}\log|R(a_1, a_2 : a_3, a_4)|,$$

where $R(a_1, a_2 : a_3, a_4)$ is the cross ratio of the points a_1, a_2 and the intersection points a_3, a_4 of the line a_1a_2 and the null circle $\mathscr{C}$. This model is not conformal, so the hyperbolic angles are not the same as the Euclidean angles [13].

This model may also be viewed as an example of a Cayley-Klein geometry, whose underlying space consists of the entire projective plane, with a distinguished conic, usually called the *absolute*, which plays the same role as the unit circle does here.

9.2.3 The Beltrami-Poincaré upper half plane model

This model can be obtained from the Beltrami-Poincaré disk model by a Cayley transformation. The underlying space of this model is the upper half plane $\mathscr{H} = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ in the complex plane $\mathbb{C}$, and notions such as that of point and angle will remain the same as the corresponding ones in the complex plane. For instance, the angle between two curves in $\mathscr{H}$ is the angle in $\mathbb{C}$ between the tangent lines to the two curves. Lines are either the intersection of $\mathscr{H}$ with an Euclidean line in $\mathbb{C}$ which is orthogonal to the $x-$axis; or the intersection of $\mathscr{H}$ with an Euclidean circle whose center is on the $x-$axis, see [4]. Given any two hyperbolic points a and b, the hyperbolic distance between them may be defined as

$$d(a, b) = \cosh^{-1}\left(1 + \left|\frac{b-a}{1-\overline{a}b}\right|\right).$$

9.2.4 The hyperboloid model

There is another model coming from a three-dimensional vector space with quadratic form

$$Q((x,y,z)) \equiv x^2 + y^2 - z^2, \tag{9.1}$$

which also goes back to Beltrami and that in some sense is more fundamental than the previous models. Here the hyperbolic space is represented by the upper sheet of the hyperboloid of two sheets with equation $x^2 + y^2 - z^2 = -1$, which turns out to be a Riemannian manifold with the metric inherited from $\mathbb{R}^3$. Lines or geodesics are given by intersections of the hyperboloid with planes through the origin. This model is quite similar in spirit to the usual spherical model of elliptic geometry, where antipodal points on a sphere are identified, see [4] and [30].

The hyperbolic model is closer to the 20-th century relativistic geometry of Einstein and Minkowski, and brings up the question of how the relativistic space outside the light cone, which the physicists sometimes refer to as de Sitter geometry, figures in hyperbolic geometry .

There are well-known projections from the hyperboloid model to the Beltrami-Klein and Beltrami-Poincaré disk models, see for example [33].

9.3 Universal hyperbolic geometry

Universal hyperbolic geometry (UHG) is a new model of hyperbolic geometry introduced and developed by Wildberger in [43], [44], [45] and [46]. In this new model, the Beltrami-Klein model has been extended to the entire projective plane, so instead of working in the interior of a disk, we are allowed to consider exterior points (including points at infinity), and also points on the boundary of the disk—which are called *null points.* These play a very important role, at least as important as the role played by interior and exterior points.

The lines in this geometry are complete projective lines rather than straight line segments. All measurements and theorems ultimately are projective and work more generally with an underlying projective plane together with a distinguished conic in it.

The symmetry between points and lines of the projective plane is inherited by this hyperbolic geometry; the main measurements of quadrance and spread which replace the hyperbolic distance and the hyperbolic angle respectively become completely dual in nature. The introduction of these metrical concepts grants new perspectives to this geometry because they give a purely *algebraic* approach to Cayley-Klein geometries, emphasizing a projective metrical formulation *without transcendental functions,* valid both inside and outside the usual null circle (or absolute), and working over a general field, generally not of characteristic two. Because of the projective nature of the measurements, this geometry extends to the case of a general conic in the projective plane.

From the point of view of the hyperboloid model, we are looking at *all* one-dimensional and two-dimensional subspaces of the three-dimensional vector space with quadratic form (9.1); these form the points and lines of the geometry, and the quadrance and spread measurements are intimately and naturally linked to the associated bilinear form. From the point of view of projective linear algebra, it is straightforward to extend this to more general quadratic forms.

Since projective geometry is key to understanding this model of hyperbolic geometry, we review some related facts in the next Section.

9.3.1 Projective geometry

Projective geometry emerged as the result of the attempt to properly present $3D$ figures in the plane; Renaissance artists were concerned with giving their drawings a more realistic resemblance to actual scenes [48]. In the 17-th century, the work of artists had been mathematically expressed by Desargues but unfortunately his work was largely ignored for about two hundred years, perhaps due to the wide interest generated by the analytic geometry of Descartes and Fermat, which relates algebra to classical geometry. Desargues is considered to be the founder of the subject (although of course Pappus' theorem had been discovered much earlier); in 1636 Desargues published a paper on perspective, which was the first account of projective geometry as an independent discipline. Pascal also contributed to the subject; of particular importance is his celebrated theorem that extends Pappus' theorem to conics [35].

Projective geometry, also known as the geometry of the straightedge, because in it there is no need for a compass, is the study of properties which are invariant under projective transformations, such as incidence of points, concurrency of lines, and the cross ratio. On the other hand, some familiar and fundamental concepts from classical geometry are not preserved in projective geometry; for instance, the usual notion of parallelism has no meaning, since any two lines always meet at a point. Moreover, length between points and angles between lines are variable quantities under projection, and so they lose their applicability in this kind of geometry.

Although quite different from the usual classical geometry, projective

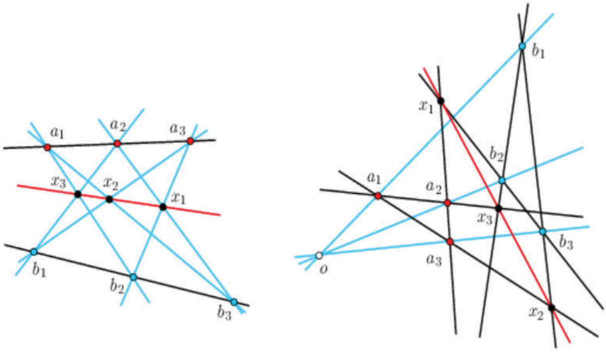


FIGURE 9.1: Theorems of Pappus and Desargues

geometry has many advantages when dealing with the study and classification of curves; for example the exceptions to Bezout's theorem can been removed; this becomes the foundational subject for modern algebraic geometry [35]. One of Wildberger's main points is that projective geometry is also the natural framework in which to develop hyperbolic geometry, see [43], [44], [45] and [46].

So far we have dealt almost exclusively with situations in which only points and lines were involved. Large parts of classical Euclidean geometry deal also with constructions involving circles. While circles are not intrinsically a concept of projective geometry, Steiner showed that *conic sections* have a natural place in this framework.

9.3.2 The projective plane

The projective plane can be obtained from the usual affine plane by adjoining to it a new additional line called the *line at infinity*. This can be done by introducing new points called *points at infinity*. For every set of parallel lines, a new point can be added that represents the intersection of these parallel lines. In this geometry, thus, every two parallel lines meet at a point.

More precisely, the projective plane can be introduced using linear algebra as follows; consider a fixed affine plane in the three-dimensional space and a fixed point O (corresponding to the origin), not on this plane. Then each point P on the plane is represented by a line which passes through O and through P. The remaining lines through O which are parallel to the affine plane represent the points at infinity and the plane containing them is the line at infinity.

In other words, the projective plane is identified with the set of all lines in space that pass through a fixed point, or equivalently as the one-dimensional subspaces of a vector space. Lines are then the two-dimensional subspaces of that vector space, which typically meet the fixed plane in a line. If our fixed plane is the plane $z = 1$ in $\mathbb{R}^3$, then we can specify a projective point by its homogeneous coordinates $[x : y : z]$, with the convention that this is the same as $[\lambda x : \lambda y : \lambda z]$, where $\lambda \neq 0$. A point such as $[x : y : 1]$ represents an actual point on the affine plane, and a point at infinity has the form $[x : y : 0]$. Dually, a line is determined by the equation of a plane, such as $lx + my + nz = 0$ and so it is represented by the proportion $\langle l : m : n \rangle$, which again is identical to $\langle \lambda l : \lambda m : \lambda n \rangle$, for any $\lambda \neq 0$. The line at infinity is $\langle 0 : 0 : 1 \rangle$, while any other line is represented on the fixed affine plane as a usual line, see [20].

Given any two points a and b in the projective plane, there is then exactly one line ab passing through them both, and given any two lines L and M there is exactly one point LM which lies on both of them.

The symmetries of the projective plane in this model are just the linear transformations of the vector space, which naturally preserve both one- and two-dimensional subspaces, hence points and lines. So the projective group here is the projective general linear group, since two linear transformations which are multiples of each other yield the same map on subspaces. At this point, the projective plane has no intrinsic metrical structure. However, there is an important quantity that can certainly be measured!

9.3.3 The cross ratio

The cross ratio is a fundamental measurement in projective geometry; it has multiple applications in different areas and it possesses beautiful algebraic properties. As mentioned earlier, the usual length and angle are not invariant under projections, but the cross ratio compensates for this deficiency, and plays a key role in UHG.

This important notion concerns four collinear points a, b, c and d on a line L, in any order. Suppose we choose affine coordinates on L so that the coordinates of a, b, c and d are respectively x, y, z and w. Then the **cross-ratio** is defined to be the extended-number (i.,e., possibly ∞) given by the ratio of ratios:

$$R(a, b : c, d) \equiv \left(\frac{a - c}{b - c} \right) / \left(\frac{a - d}{b - d} \right).$$

This is independent of the choice of affine coordinates on L.

The cross-ratio is also projectively invariant, meaning that if a_1, b_1, c_1 and d_1 are also collinear points on a line L_1 which are perspective to a, b, c and d from some point p, as in Figure 9.2, then $(a, b : c, d) = (a_1, b_1 : c_1, d_1)$.

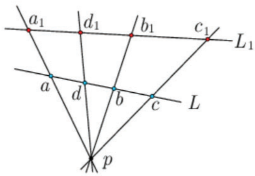


FIGURE 9.2: Projective invariance of the cross-ratio: $(a,b:c,d) = (a_1,b_1:c_1,d_1)$

Dually, the cross ratio of four concurrent (projective) lines can be defined in a similar way. An important fact is that if a line A meets the concurrent lines A_1,A_2,A_3,A_4 at four distinct points a_1,a_2,a_3,a_4, then $R(A_1,A_2;A_3,A_4) = R(a_1,a_2;a_3,a_4)$. In addition, if four collinear points a_1,a_2,a_3,a_4 are projective with four collinear points a_1',a_2',a_3',a_4', then $R(a_1,a_2;a_3,a_4) = R(a_1',a_2';a_3',a_4')$.

Four collinear points a_1,a_2,a_3,a_4 are said to be a **harmonic range** when $R(a_1,a_2;a_3,a_4) = -1$. In such a case, a_1 and a_2 are referred to as **harmonic conjugates** with respect to the pair a_3,a_4 and vice versa.

9.3.4 UHG from a synthetic projective view

Universal Hyperbolic Geometry can be approached from either a synthetic, projective-geometry-like point of view, or from an analytic, linear-algebra-like point of view; both are useful and they shed light on each other. In this Section we present a synthetic introduction that is useful for dynamic geometry packages such as GSP, C.a.R., Cabri, GeoGebra and Cinderella. We describe the situation generally in the projective plane over a field, which in our diagrams will be the field of rational numbers, with a distinguished conic, called the **null circle**, but elsewhere also the *absolute*. In the included pictures, this will be the standard unit circle, always depicted in blue. The points on the unit circle are called **null points**. We may consider this to be the affine circle $X^2+Y^2=1$ or the projective version $x^2+y^2-z^2=0$.

Wildberger's set-up of universal hyperbolic geometry is more general, being set in a Cayley-Klein geometry both inside and outside the null circle (absolute), *over a general field*. One of the challenges is to extend basic notions, such as midpoints, angle bisectors/bilines, perpendicular bisectors/midlines etc., to rather general triangles. This is the component of our contribution that we want to describe in this Chapter. The notions of midpoints and bilines can be extended, but one must be prepared to undertake the study of *sydpoints* and *silines* as well.

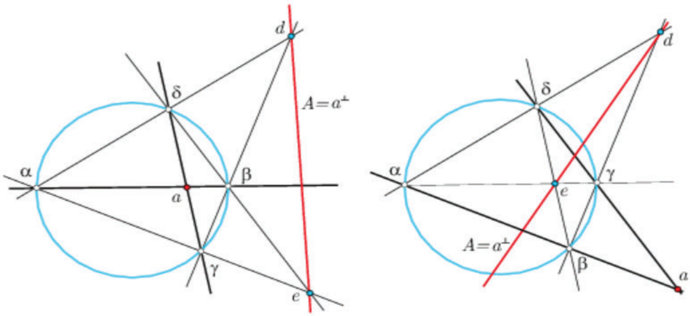


FIGURE 9.3: Duality and pole-polar pairs

Figure 9.3 shows a construction for the Apollonian dual of a point a; this is the line $A = a^{\perp}$ formed by the other two diagonals of any null quadrangle $\alpha\beta\gamma\delta$ for which a is a diagonal point. Note that the construction works both when a is inside as well as when a is outside, the null circle. To construct the dual of a line L, take the intersection of the duals of any two points on it.

This duality between points and lines induced by the null circle allows a notion of perpendicularity: two points a and b are **perpendicular**, written $a \perp b$, precisely when b lies on the dual $a^{\perp}$ of a, or, equivalently, if a lies on the dual $b^{\perp}$ of b. Similarly two lines L and M are said to be **perpendicular** ($L \perp M$) if and only if L passes through the dual of M (equivalently, if M passes through the dual of L). In Figure 9.4, several lines perpendicular to the line A are depicted: notice that they pass through the dual point a.

The basic isometries in the above described geometry are reflections in points (or reflections in lines—these two notions turn out to be the same). If m is not a null point, the reflection r_m in m interchanges the two null points on any line through m, should there be such. In Figure 9.5 for example, r_m interchanges x and w, and also y and z. It is then a remarkable and fundamental fact that r_m extends to a projective transformation: to find the image of a point a, construct any line through a which meets the null circle at two points, say x and y, then find the images of x and y under r_m, namely w and z, and then define $r_m(a) = b \equiv (am)(wz)$ as shown. Perpendicularity of both points and lines is preserved by r_m.

The notion of reflection allows us to define midpoints without metrical measurements: if $r_m(a) = b$ then we may say that m is a **midpoint** of the

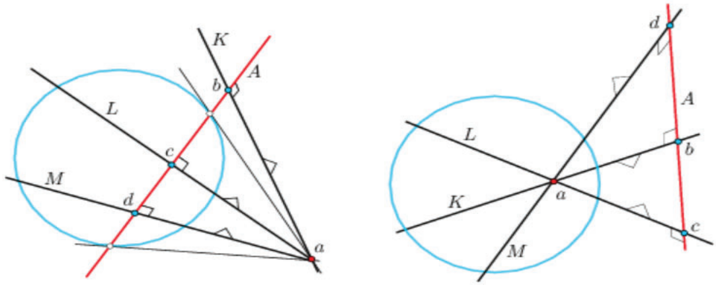


FIGURE 9.4: Perpendicular points and lines

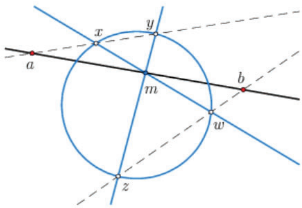


FIGURE 9.5: Reflection r_m in m sends a to b

side $\overline{ab}$. To construct the midpoints of a side $\overline{ab}$, when they exist (this is essentially a quadratic condition), we essentially invert the above construction.

Following [44] a side $\overline{a_1a_2} = \{a_1, a_2\}$ is said to be a set of two points, and a vertex $\overline{L_1L_2} = \{L_1, L_2\}$ is said to be a set of two lines. Figure 9.7 shows two situations where midpoints m and n of the side $\overline{ab}$ can be constructed, at least approximately over the rational numbers, which is the orientation of the Geometer's Sketchpad and other dynamic geometry packages. In the left diagram, we take the dual c of the line ab, and if the lines ac and bc meet the null circle, we take the other two diagonal points of this null quadrangle. This is also the case in Figure 9.5. In the right diagram, the lines ac and bc do not meet the null circle, but the dual lines A and B of a and b, which necessarily pass through c, do meet the null circle in a quadrangle, whose other diagonal points are the required midpoints m and n.

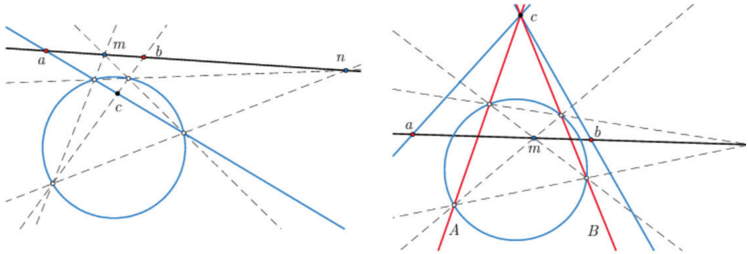


FIGURE 9.6: Constructing midpoints m and n of the side $\overline{ab}$

9.3.5 Quadrance and spread in UHG

Now we show how the main metrical notions of quadrance and spread, as introduced by Wildberger, can be framed purely projectively. In the Beltrami-Klein model, the hyperbolic distance between two points a and b may be defined by means of a logarithmic function, together with the absolute value of the cross ratio of the points a, b, and the intersection points of the line ab and the null circle. However in universal hyperbolic geometry, the **quadrance** between two points a and b may be defined simply as the cross ratio

$$q(a,b) = R(a,d:b,c),$$

where $c \equiv (ab)\,a^{\perp}$ and $d \equiv (ab)\,b^{\perp}$.

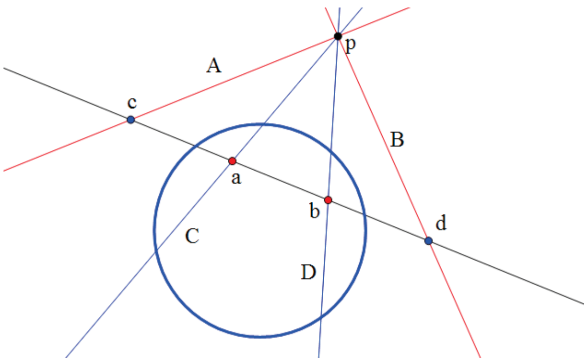


FIGURE 9.7: Defining quadrance and spread via cross ratios

This hyperbolic quadrance is superior to the usual hyperbolic distance in several ways. First of all for two general points a,b the existence of the conjugate points c,d is always guaranteed, and so the quadrance extends to general points. In addition, there is no ambiguity in the ordering of the four points when applying the cross-ratio, while this is not true in the Beltrami Klein model. Furthermore, hyperbolic quadrance does not involve any transcendental function, such as the logarithm or an inverse circular function.

Such an algebraic approach allows the theory to extend to finite fields. In addition, the purely projective nature of the cross ratio guarantees that the same results will hold when applying a projective transformation to replace the chosen distinguished conic with a more general one.

Over the rational numbers, quadrance takes negative values when a and b are either interior or exterior to the null circle, while it is positive when one of the points a or b is interior and the other one is exterior.

Dually, the **spread** between the lines A and B is defined to be the cross-ratio

$$S(A,B) \equiv R(A,D:B,C),$$

where C and D are the dual lines of the conjugate points c and d. Notice that Figure **??** shows important points and lines used in the definitions of quadrance and spread. It follows from basic properties of the cross-ratio, that $q(a,b) = S(A,B)$. This means that the fundamental projective duality between points and lines extends to the metric notions of quadrance and spread. Note also that it is not necessary for the line ab to meet the null circle; in fact, these metric notions are valid for all points and lines, except when null points or null lines are involved, when the cross-ratio becomes infinite.

As shown in [44], it is possible to translate the theorems in UHG to formulas of classical hyperbolic trigonometry in the special case of interior points and lines using the following (necessarily approximate, transcendental) relations;

$$q(a,b) = -\sinh^2(d(a,b)) \qquad \text{and} \qquad S(A,B) = \sin^2(\theta(A,B)).$$

Hence, all formulas in this Chapter can be interpreted and reformulated in case the points and lines are interior to the null circle, yielding relations between classical distances and angles via these relations, if such are desired.

While we could proceed with this synthetic point of view, we prefer to work in an analytic environment using projective linear algebra.

9.3.6 Circles

A circle $\mathscr{C}$ in this setting may be defined synthetically or algebraically. Suppose that c and p are points; then the locus of the reflections $r_x(p)$ as x runs along the dual line of c, is the **circle** with center c, through p. This definition immediately yields a correspondence between a circle and a line. Algebraically a circle is an equation of the form $q(x,c) = k$, where c is the centre and k is the quadrance.

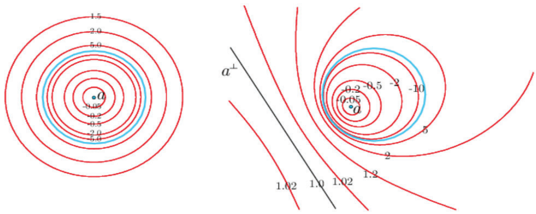


FIGURE 9.8: Circles centered at a, for an interior point

In Figure 9.8 examples of hyperbolic circles centered at a are presented, for two different choices of a inside the null circle, also showing the values of the quadrance k. These curves appear in our diagrams always as conics: as ellipses, parabolas or hyperbolas.

Also in Figure 9.9, examples of hyperbolic circles centered at a are given, for a choice of a outside the null circle. In classical hyperbolic geometry, The circles that result here are usually called curves of constant width (at least for those inside the null circle). Notice that all such circles are tangent to the null circle at the two points where the dual line $a^\perp$ of the center meets the null circle.

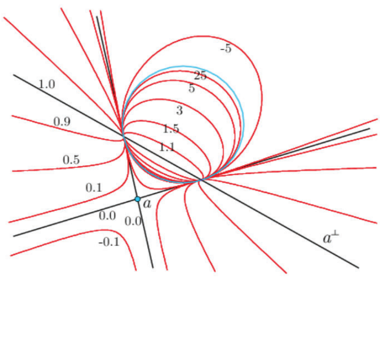


FIGURE 9.9: Circles centered at a, for an exterior point

9.3.7 Relativistic geometry

In 1905 Einstein wrote his fundamental paper on Special Relativity (SR) and a few years later Minkowski introduced a geometric framework for it, which proved vital for Einstein's introduction of General Relativity in 1918. Minkowski's idea was to introduce a four dimensional space time, consisting of vectors (x,y,z,t), in which x,y and z represent spacial coordinates and t is time. In this space-time, particles now have trajectories, which describe their entire histories, and observers moving in inertial frames (at a constant velocity with respect to the fixed stars) can now compare their measurements of events, which are points in this space-time.

The Lorentz transformations that underlie the symmetry of SR turn out to be isometries of this four dimensional space, under the indefinite inner product

$$(x_1,y_1,z_1,t_1)\cdot(x_2,y_2,z_2,t_2) = x_1x_2+y_1y_2+z_1z_2-t_1t_2$$

or the associated quadratic form

$$Q(x,y,z,t) = x^2+y^2+z^2-t^2.$$

The physical implication is that neither the traditional three-dimensional distance (or quadrance) or the time are actually physically meaningful quantities: only the Einstein interval represented by the quadratic form Q is observable.

In this geometry the null cone consisting of vectors v for which $Q(v) = 0$, plays the role of the path of photons traveling at the speed of light. If a particle's world-line passes through the origin $(0,0,0,0)$, then Einstein's fundamental principle according to which nothing can exceed the speed of light, implies that the world-line must then lie entirely in the interior of this light-cone given by $x^2+y^2+z^2=t^2$, if units are chosen appropriately. If a particle's world-line were to traverse outside this light cone, a speed greater than that of light would be implied, which would yield curious relations with respect to ”time travel”. Physicists are still divided on whether or not such a phenomenon might exist. Nevertheless, the geometry outside his/her light cone is important to an observer, because the future light cones of other observers or events will inevitably meet.

In SR, the nature of the geometry inside the light cone is different from that outside the light cone; the latter is often associated with the work of de Sitter. It was early on realized that the geometry of SR inside the light cone, at least in the simpler case of a three dimensional space-time (x,y,t), was very closely connected to hyperbolic geometry, since Beltrami's hyperboloid model could be viewed as a Riemannian submanifold inside the interior of the upper portion of the light cone: actually, as an orbit under the isometry group of the geometry.

Another way to say this is that classical hyperbolic geometry corresponds to a projective view of the interior of the light cone, as such one-dimensional subspaces (lines through the origin) meet such a submanifold at unique points. Furthermore the hyperbolic plane then naturally has the light cone, or its projective analog the circle, as a limiting object.

This picture can be made explicit in the three-dimensional space-time model by viewing one-dimensional subspaces via their meets with the viewing plane $t = 1$. In this case, the light cone becomes the null circle $x^2 + y^2 = 1$ and two-dimensional subspaces represent geodesics , meeting the viewing plane in lines, and so we recover the Beltrami-Klein projective model.

In UHG, we push this physical motivation further and consider also the outside of the light cone, which is represented projectively by points outside the null circle. It becomes then natural to frame things algebraically in the language of projective linear algebra, incorporating the bilinear form of Einstein.

9.3.8 Metrical projective linear algebra

While the synthetic framework is attractive, for explicit computations and formulas it is useful to work with analytic geometry in the context of (projective) linear algebra. We will now proceed to explain this, starting with some notation and basic results in the affine setting, although the projective setting is the main interest [6].

The three-dimensional vector space $V = \mathbb{F}^3$ over a field $\mathbb{F}$, not of characteristic two, consists of row vectors $v = (x, y, z)$, or equivalently, of 1×3 matrices $\begin{pmatrix} x & y & z \end{pmatrix}$. A metric structure is determined by a *symmetric bilinear form*

$$v \cdot u \equiv vCu^T \tag{9.2}$$

where C is an invertible, symmetric 3×3 matrix. The elements of the dual vector space V^* may be viewed as column vectors $f = (l, m, n)^T$, or equivalently as 3×1 matrices.

Two vectors v, u are said to be**perpendicular** if $v \cdot u = 0$. The **(affine) quadrance** of a vector v is the number $Q_v \equiv v \cdot v$. A vector v is **null** precisely when $Q_v = 0$.

A variant of the following also appears in [25].

Theorem 9.1. *(Parallel vectors) If the vectors v and u are parallel, then*

$$Q_v Q_u = (v \cdot u)^2 . \tag{9.3}$$

Conversely if (9.3) holds, then either v and u are parallel, or the bilinear form restricted to the span of v and u is degenerate.

The previous result motivates the following measure of the non-parallelism of two vectors. The **(affine) spread** between two non-null vectors v and u is the number

$$s(v,u) \equiv 1 - \frac{(v \cdot u)^2}{Q_v Q_u}.$$

The spread is unchanged if either v or u are multiplied by a non-zero number.

One-dimensional and two-dimensional subspaces of $V = \mathbb{F}^3$ may be viewed as the basic objects forming the projective plane, with metric notions coming from the affine notions of quadrance and spread in the associated vector space, but we prefer to give independent definitions, so that logically, neither the affine nor the projective setting has priority. In general, our notation in the projective setting is *opposite* to that in the affine setting, in the sense that the roles of small and capital letters are reversed.

In *projective linear algebra,* when multiplying a vector or matrix by a scalar, the same object is obtained, meaning that vectors and matrices are defined only up to non-zero scalar multiples. Here, in order to differentiate between ordinary linear algebra and *projective linear algebra*, *round brackets will be used* for the usual vectors and matrices, and square brackets will be used to indicate vectors and matrices in the projective setting. Notice that the operations of addition and subtraction of projective vectors or matrices are undefined, while the operations of multiplication, transposes and inverses are well-defined. Pleasantly, computing inverses in the projective setting is very easy since common denominators can be scaled away. In particular. over the rational numbers, *integer arithmetic* is usually enough to deal with projective linear algebra.

A (**projective) point** is a proportion $a = [x : y : z]$ in square brackets, or equivalently a projective row vector $a = \begin{bmatrix} x & y & z \end{bmatrix}$, where the square brackets in the latter equality are interpreted projectively: when multiplying by a non-zero number this is unchanged. A (**projective) line** is a proportion $L = \langle l : m : n \rangle$ in pointed brackets, or equivalently, a projective column vector

$$L = \begin{bmatrix} l \\ m \\ n \end{bmatrix}.$$

When the context is clear, projective points and projective lines are simply referred to as **points** and **lines**. The **incidence** between the point $a = [x : y : z]$ and the line $L = \langle l : m : n \rangle$ is given by the relation

$$aL = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} l \\ m \\ n \end{bmatrix} = lx + my + nz = 0.$$

In such a case we say that a **lies on** L, or that L **passes through** a.

The **join** a_1a_2 of two distinct points $a_1 \equiv [x_1 : y_1 : z_1]$ and $a_2 \equiv [x_2 : y_2 : z_2]$ is the line

$$a_1a_2 \equiv [x_1 : y_1 : z_1] \times [x_2 : y_2 : z_2] \equiv \langle y_1z_2 - y_2z_1 : z_1x_2 - z_2x_1 : x_1y_2 - x_2y_1 \rangle . \tag{9.4}$$

This is the unique line passing through a_1 and a_2. The **meet** L_1L_2 of distinct lines $L_1 \equiv \langle l_1 : m_1 : n_1 \rangle$ and $L_2 \equiv \langle l_2 : m_2 : n_2 \rangle$ is the point

$$L_1L_2 \equiv \langle l_1 : m_1 : n_1 \rangle \times \langle l_2 : m_2 : n_2 \rangle \equiv [m_1n_2 - m_2n_1 : n_1l_2 - n_2l_1 : l_1m_2 - l_2m_1] . \tag{9.5}$$

This is the unique point lying on both L_1 and L_2.

Three points a_1, a_2, a_3 are said to be **collinear** if they lie on a common line L; in this case we will also write $L = [[a_1a_2a_3]]$. Similarly three lines L_1, L_2, L_3 are said to be **concurrent** if they pass through a common point a; in this case we will also write $a = [[L_1L_2L_3]]$.

9.3.9 Projective quadrance and spread

If C is a symmetric, invertible 3×3 matrix, with entries in $\mathbb{F}$ and D is its adjugate matrix (the inverse, up to a multiple), then we denote by $\mathbf{C}$ and $\mathbf{D}$ the corresponding projective matrices, each defined up to a non-zero multiple. This pair of projective matrices determines a metric structure on projective points and lines, as follows.

The (projective) points a_1 and a_2 are called **perpendicular** if $a_1\mathbf{C}a_2^T = 0$, written $a_1 \perp a_2$. This is a well-defined, symmetric relation. Similarly, the (projective) lines L_1 and L_2 are said to be **perpendicular** if $L_1^T\mathbf{D}L_2 = 0$, written $L_1 \perp L_2$. The point a and the line L are said to be **dual** if

$$L = a^{\perp} \equiv \mathbf{C}a^T \qquad \text{or equivalently} \qquad a = L^{\perp} \equiv L^T\mathbf{D}. \tag{9.6}$$

Two points are perpendicular if *one is incident with the dual of the other,* and similarly for two lines. So $a_1 \perp a_2$ precisely when $a_1^{\perp} \perp a_2^{\perp}$, because of the projective relation

$$\left(\mathbf{C}a_1^T\right)^T \mathbf{D}\left(\mathbf{C}a_2^T\right) = \left(a_1\mathbf{C}^T\right)\mathbf{D}\left(\mathbf{C}a_2^T\right) = a_1\left(\mathbf{CD}\right)\left(\mathbf{C}a_2^T\right) = a_1\mathbf{C}a_2^T.$$

A point a is said to be **null** if it is perpendicular to itself, that is, when $a\mathbf{C}a^T = 0$, and a line L is called**null** when it is perpendicular to itself, that is, when $L^T\mathbf{D}L = 0$. The null points determine the **null conic**, sometimes also called the *absolute*.

Hyperbolic and *elliptic geometries* arise, respectively, from the special cases

$$C = J \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = D \qquad \text{and} \qquad C = I \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = D. \tag{9.7}$$

In the hyperbolic case, which is the main object of interest for us in this discussion and which forms the basis for almost all examples, the point $a = [x : y : z]$ is null precisely when

$$x^2 + y^2 - z^2 = 0$$

and dually, the line $L = (l : m : n)$ is null precisely when

$$l^2 + m^2 - n^2 = 0.$$

This is the reason why the null circle can be pictured in affine coordinates $X \equiv x/z$ and $Y \equiv y/z$ as the (blue) circle $X^2 + Y^2 = 1$. Note that in the elliptic case, the null circle over the rational numbers, has no points lying on it. This is why visualizing hyperbolic geometry is often easier than visualizing elliptic geometry!

In the general setting, the bilinear forms determined by C and D can be used to define the metric structure in the associated projective plane. The dual notions of (**projective**) **quadrance** $q(a_1, a_2)$ between two points a_1 and a_2, and (**projective**) **spread** $S(L_1, L_2)$ between two lines L_1 and L_2, are defined by

$$q(a_1, a_2) \equiv 1 - \frac{\left(a_1 \mathbf{C} a_2^T\right)^2}{\left(a_1 \mathbf{C} a_1^T\right)\left(a_2 \mathbf{C} a_2^T\right)} \qquad \text{and} \qquad S(L_1, L_2) \equiv 1 - \frac{\left(L_1^T \mathbf{D} L_2\right)^2}{\left(L_1^T \mathbf{D} L_1\right)\left(L_2^T \mathbf{D} L_2\right)}. \tag{9.8}$$

While the numerators and denominators of these expressions depend on choices of representative vectors and matrices for $a_1, a_2, \mathbf{C}, L_1, L_2$ and $\mathbf{D}$, the *quotients are independent of scaling*, so the overall expressions are indeed well-defined projectively. If $a_1 = [v_1]$, $a_2 = [v_2]$, and $L_1 = [f_1]$, $L_2 = [f_2]$, then we may write

$$q(a_1, a_2) = 1 - \frac{(v_1 \cdot v_2)^2}{(v_1 \cdot v_1)(v_2 \cdot v_2)} \qquad \text{and} \qquad S(L_1, L_2) = 1 - \frac{(f_1 \odot f_2)^2}{(f_1 \odot f_1)(f_2 \odot f_2)},$$

where we use (9.2) and introduce the dual bilinear form on column vectors by $f_1 \odot f_2 \equiv f_1^T D f_2$.

Clearly $q(a, a) = 0$ and $S(L, L) = 0$ for any point a and any line L, while $q(a_1, a_2) = 1$ precisely when $a_1 \perp a_2$, and dually, $S(L_1, L_2) = 1$ precisely when $L_1 \perp L_2$. Then using (9.6) we see that

$$S\left(a_1^{\perp}, a_2^{\perp}\right) = q(a_1, a_2).$$

In [43], using suitable cross ratios, Wildberger showed that these metrical notions agree with the projective formulation.

Example 9.1. In the hyperbolic case, the quadrance between points and the spread between lines are given by essentially similar formulas:

$$q\left([x_1 : y_1 : z_1], [x_2 : y_2 : z_2]\right) = 1 - \frac{\left(x_1x_2 + y_1y_2 - z_1z_2\right)^2}{\left(x_1^2 + y_1^2 - z_1^2\right)\left(x_2^2 + y_2^2 - z_2^2\right)} \quad (9.9)$$

$$S\left(\langle l_1 : m_1 : n_1\rangle, \langle l_2 : m_2 : m_2\rangle\right) = 1 - \frac{\left(l_1l_2 + m_1m_2 - n_1n_2\right)^2}{\left(l_1^2 + m_1^2 - n_1^2\right)\left(l_2^2 + m_2^2 - n_2^2\right)}. \quad \diamond$$

9.3.10 Hyperbolic trigonometry in UHG

The following formula, introduced in [41], is given in a more general setting in [42].

Theorem 9.2. *(Projective Triple quad formula) Suppose that a_1, a_2, a_3 are collinear points, with quadrances $q_1 \equiv q(a_2, a_3)$, $q_2 \equiv q(a_1, a_3)$ and $q_3 \equiv q(a_1, a_2)$. Then*

$$(q_1 + q_2 + q_3)^2 = 2\left(q_1^2 + q_2^2 + q_3^2\right) + 4q_1q_2q_3. \quad (9.10)$$

We present a few useful consequences of the Triple quad formula. If one of the quadrances is $q_3 = 1$, then $q_1 + q_2 = 1$; this is a consequence of the identity

$$(q_1 + q_2 + 1)^2 - 2q_1^2 - 2q_2^2 - 2 - 4q_1q_2 = -\left(q_1 + q_2 - 1\right)^2.$$

Also, if two of the quadrances are equal, say $q_1 = q_2 = r$, then $q_3 = 0$ or $q_3 = 4r(1-r)$; this is a consequence of the identity

$$(2r + q_3)^2 - 4r^2 - 2q_3^2 - 4r^2q_3 = -q_3\left(q_3 - 4r + 4r^2\right).$$

We are now in a position to give a more complete list of the main trigonometric laws in this setting. The laws are taken from [44]. These completely algebraic rules incorporate and extend the more familiar, transcendental ones found in the Klein model and in the Poincaré model. A **triangle** $\overline{a_1a_2a_3} = \{a_1, a_2, a_3\}$ is a set of three non-collinear points. A **trilateral** $\overline{L_1L_2L_3} = \{L_1, L_2, L_3\}$ is a set of three lines that are not concurrent. Every triangle $\overline{a_1a_2a_3}$ has three sides, namely $\overline{a_1a_2}$, $\overline{a_2a_3}$ and $\overline{a_1a_3}$; similarly, any trilateral $\overline{L_1L_2L_3}$ has three vertices, namely $\overline{L_1L_2}$, $\overline{L_2L_3}$ and $\overline{L_1L_3}$.

We will use the convention that $q_1 \equiv q(a_2, a_3)$, $q_2 \equiv q(a_1, a_3)$ and $q_3 \equiv q(a_1, a_2)$, and that $S_1 \equiv S(L_2, L_3)$, $S_2 \equiv S(L_1, L_3)$ and $S_3 \equiv S(L_1, L_2)$, along with the notation shown in the Figure.

The following are the main trigonometric laws of UHG.

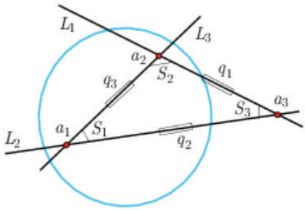


FIGURE 9.10: Quadrance and spreads in a hyperbolic triangle

Theorem 9.3. *(Triple quad formula) If a_1, a_2 and a_3 are collinear points, then*

$$(q_1+q_2+q_3)^2 = 2\left(q_1^2+q_2^2+q_3^2\right)+4q_1q_2q_3.$$

Theorem 9.4. *(Triple spread formula) If L_1, L_2 and L_3 are concurrent lines, then*

$$(S_1+S_2+S_3)^2 = 2\left(S_1^2+S_2^2+S_3^2\right)+4S_1S_2S_3.$$

Theorem 9.5. *(Pythagoras) If L_1 and L_2 are perpendicular lines, then*

$$q_3 = q_1+q_2-q_1q_2.$$

Theorem 9.6. *(Pythagoras' dual) If a_1 and a_2 are perpendicular points, then*

$$S_3 = S_1+S_2-S_1S_2.$$

Theorem 9.7. *(Spread law) For a triangle $\overline{a_1a_2a_3}$ with quadrances q_1, q_2, q_2 and spreads S_1, S_2, S_3, one has*

$$\frac{S_1}{q_1} = \frac{S_2}{q_2} = \frac{S_3}{q_3}.$$

Theorem 9.8. *(Cross law) For a triangle $\overline{a_1a_2a_3}$ with quadrances q_1, q_2, q_2 and spreads S_1, S_2, S_3, it holds that*

$$(q_1q_2S_3-(q_1+q_2+q_3)+2)^2 = 4(1-q_1)(1-q_2)(1-q_3).$$

Theorem 9.9. *(Cross dual law) For a triangle $\overline{a_1a_2a_3}$ with quadrances q_1, q_2, q_2 and spreads S_1, S_2, S_3, one has*

$$(S_1S_2q_3-(S_1+S_2+S_3)+2)^2 = 4(1-S_1)(1-S_2)(1-S_3).$$

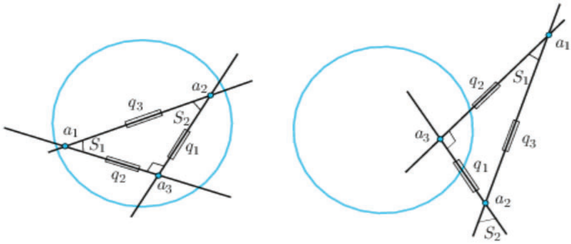


FIGURE 9.11: Right triangles in UHG

Theorem 9.10. *(Napier's Rules) Suppose that $\overline{a_1a_2a_3}$ is a right triangle with $S_3 = 1$. Then, any two of the five quantities S_1, S_2, q_1, q_2 and q_3 determine the other three, using only the three basic equations from Thales' theorem and Pythagoras' theorem:*

$$S_1 = \frac{q_1}{q_3},\ S_2 = \frac{q_2}{q_3} \quad \text{and} \quad q_3 = q_1 + q_2 - q_1 q_2.$$

Another important theorem is:

Theorem 9.11. *(Menelaus') Suppose that $\overline{a_1a_2a_3}$ is an arbitrary triangle and let b_1, b_2 and b_3 be three points on a_2a_3, a_1a_3 and a_1a_2, respectively. If b_1, b_2 and b_3 are collinear, then*

$$\frac{q(a_1,b_3)}{q(b_3,a_2)}\frac{q(a_2,b_1)}{q(b_1,a_3)}\frac{q(a_3,b_2)}{q(b_2,a_1)} = 1.$$

9.4 Sydpoints, triangle geometry and twin circles

In [46] it was shown that if each of the three sides of a triangle (in UHG) has midpoints m, then these six points lie three at a time on four circumlines C, whose duals are the four *circumcenters* c. These are the centers of the four *circumcircles* which pass through the three points of the triangle. In the geometry under consideration, a circle $\mathscr{C}$ may be defined as an equation of the form $q(c,x) = k$, for a fixed point c called the **center**, and a fixed number k called the **quadrance** of the circle, and a point a is said to **lie on** the circle, if $q(c,a) = k$. Since in this case the circle is also determined by c and a, we may refer to it as $\mathscr{C}_c^{(a)}$. The bracket is a reminder of the non-uniqueness of the point a.

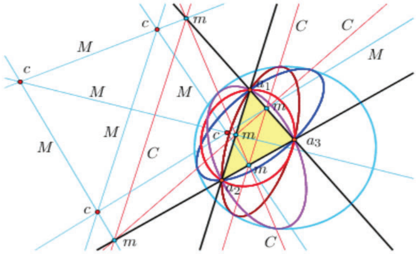


FIGURE 9.12: Midpoints, midlines, circumlines, circumcenters and circumcircles

This four circumcircles are shown for a classical triangle inside the null circle in Figure 9.12. The duals of the midpoints m are the *midlines* M, traditionally called the perpendicular bisectors of the sides. While the red circumcircle is a classical circle in the Cayley-Beltrami-Klein model of hyperbolic geometry, the other three are usually described as *curves of constant width*, but in UHG they are *all just circles*. This is the start of the circumcenter hierarchy as studied in [46].

Remarkably, much of these observations extend also to triangles with points both interior and exterior to the null circle; in the process of this study we also find new phenomena related to circumcircles, that suggest a reconsideration of the classical case above.

Next, the new notion of the sydpoints of a side in hyperbolic geometry is explained, and we delve into how it allows us to widen the study of hyperbolic triangle geometry. This is parallel to, but with features different from, the Euclidean case laid out in [23] and [24], and in a related but different direction from Ungar in [40].

We have seen that a midpoint of a side $\overline{ab}$ is a point m on ab satisfying

$$q(a,m) = q(b,m).$$

The new concept is the following: a **sydpoint** of $\overline{ab}$ is a point s on ab satisfying

$$q(a,s) = -q(b,s).$$

Example 9.2. In the usual hyperbolic case, suppose that $a = [x:0:1]$ and $b = [y:0:1]$. It can be verified by direct computation that

$$q(a,b) = -\frac{(x-y)^2}{(1-x^2)(1-y^2)},$$

Thus, midpoints $m = [w : 0 : 1]$ and sydpoints $s = [z : 0 : 1]$ of $\overline{ab}$ exist precisely when $\left(x^2 - 1\right)\left(y^2 - 1\right) = r^2$ and $\left(x^2 - 1\right)\left(y^2 - 1\right) = -t^2$ respectively, in which cases

$$w = \frac{xy + 1 \pm r}{x + y} \qquad \text{and} \qquad z = \frac{(1 - xy)(x + y) \pm t(x - y)}{x^2 + y^2 - 2}.$$

It is thus clear that algebraically sydpoints are somewhat more complicated than midpoints in general. $\diamond$

As shown in [44], the existence of midpoints is equivalent to $1 - q(a, b)$ being a square in the underlying field. It turns out that the existence of sydpoints is equivalent to $q(a, b) - 1$ being a square. Asis the case with midpoints, if sydpoints exist, there are generally two of them.

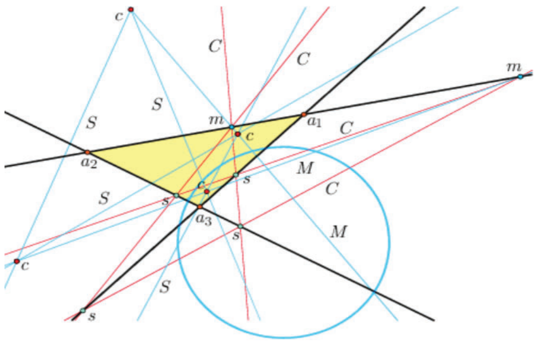


FIGURE 9.13: A non-classical triangle with both midpoints and sydpoints

Hence, sydpoints allow us to extend much of triangle geometry to non-classical triangles, with points inside and outside of the null conic. In Figure 9.13, the non-classical triangle $\overline{a_1a_2a_3}$ has one side $\overline{a_1a_2}$ with midpoints m whose duals are *midlines* M and two sides $\overline{a_1a_3}$ and $\overline{a_2a_3}$ with sydpoints s whose duals are *sydlines* S. Somewhat remarkably, the six midpoints and sydpoints lie three at a time on four *circumlines* C, whose duals are the four *circumcenters* c. The connection between these new circumcenters and the idea of circumcircles is particularly interesting, since, in this case, it is impossible to find *any* circles which pass through all three points of the triangle $\overline{a_1a_2a_3}$.

In UHG, circles can often be paired: two circles are defined to be **twins** if they share the same center and their quadrances add up to 2. The circumcenters c are the centers of *twin circumcircles* passing through all the three points of the triangle. This notion extends our understanding even in the classical case. The four pairs of twin circumcircles give eight *generalized circumcircles*, even for the classical case.

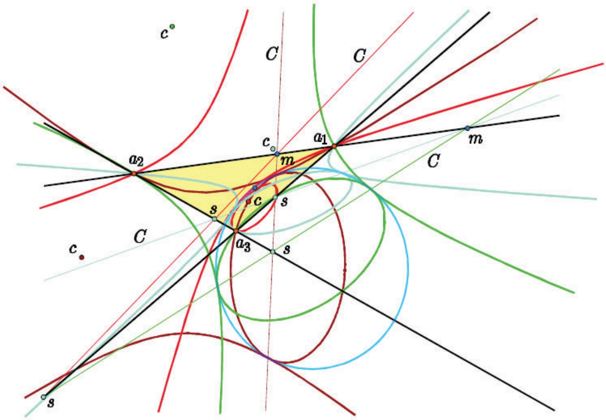


FIGURE 9.14: Four twin circumcircles of a non-classical triangle

In Figure 9.14, we see the twin circumcircles of the triangle of the previous Figure; some of these appear in this model as hyperbolas tangent to the null circle—these are invisible in classical hyperbolic geometry, but have a natural interpretation in terms of hyperboloids of one sheet in three-dimensional space (de Sitter space).

9.4.1 The construction of Sydpoints

The following theorem is helpful in constructing sydpoints using a dynamic geometry package.

Theorem 9.12. *(Sydpoints null points) Suppose that the non-null side $\overline{ab}$ has sydpoints s and r, and that $\overline{ac}$ has midpoints m and n, where $c = (ab)^{\perp}$. Then $x \equiv (mr)(bc) = (ns)(bc)$ and $y \equiv (ms)(bc) = (nr)(bc)$, are null points.*

We use this theorem to give practical constructions of sydpoints using Geometer's Sketchpad, C.a.R., Cabri, GeoGebra or Cinderella. To construct the sydpoints r and s of $\overline{ab}$ as in Figure 9.15, first construct the dual $c = (ab)^{\perp}$, then the midpoints m and n of $\overline{ac}$, and then use the null points x and y lying on bc as shown.

The required sydpoints are $s = (nx)(ab) = (my)(ab)$ and $r = (ny)(ab) = (mx)(ab)$. Similarly, given the sydpoints r and s of $\overline{ab}$, a and b can be constructed as the sydpoints of $\overline{rs}$ using the null points w and z lying on rc, and

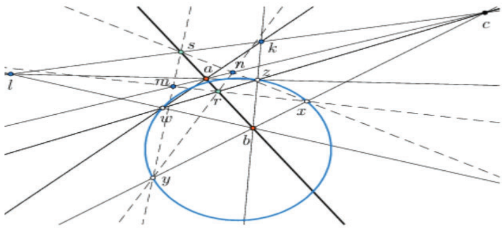


FIGURE 9.15: Construction of sydpoints of $\overline{ab}$

the midpoints k and l of $\overline{cs}$. The required points are $a = (lz)(rs) = (kw)(rs)$ and $b = (lw)(rs) = (kz)(rs)$. So the construction of sydpoints can be reduced, at least in this kind of situation, to the computation of midpoints.

Another useful construction is to find, given the point b and one of the sydpoints s, the other point a and the other sydpoint r as in Figure 9.16. First construct the dual $c = (bs)^{\perp}$, then find the midpoints k and l of $\overline{cs}$. Use the null points u,t lying on bk and the null points v,w lying on bl to construct $r = (cuv)(bs)$ and $a = (lu)(bs) = (kv)(bs)$.

However, by symmetry there is a second solution, namely $\overline{r} = (cwt)(bs)$ and $\overline{a} = (lt)(bs) = (kw)(bs)$. Thus, s and r can be thought of as being the sydpoints of the side $\overline{ab}$, and s and $\overline{r}$ as the sydpoints of the side $\overline{\overline{a}b}$. Notice also that b is a midpoint of the side $r\overline{r}$ and similarly, s is a midpoint of the side $a\overline{a}$. In fact

$$q(b,r) = q(b,\overline{r}) = q(s,a) = q(s,\overline{a}).$$

9.4.2 Twin circles

The connection between sydpoints and twin circles is described by the following:

Theorem 9.13. *(Sydpoint twin circle) If s is a sydpoint of $\overline{ab}$, and c lies on $S \equiv s^{\perp}$, then the circles $\mathscr{C}_c^{(a)}$ and $\mathscr{C}_c^{(b)}$ are twins. Conversely if $\mathscr{C}_c^{(a)}$ and $\mathscr{C}_c^{(b)}$ are twins, then $s \equiv c^{\perp}(ab)$ is a sydpoint of $\overline{ab}$.*

We note that the theorem has another possible, quite interesting, interpretation: *the locus of a point c such that $q(a,c) + q(b,c) = 2$, is a line.*

The Sydpoint twin circle theorem assists us in the construction of twin

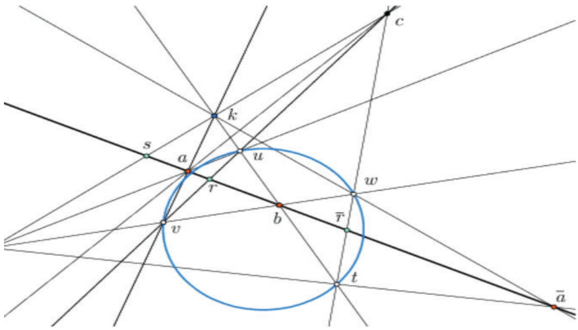


FIGURE 9.16: Constructing r and a (or $\overline{r}$ and $\overline{a}$) from s and b

circles; we generally expect this task to reduce to finding midpoints, but there are also some simpler scenarios. Suppose we are given a circle $\mathscr{C}$ (in brown) with center c as in Figure 9.17. Choose an arbitrary point a on the circle $\mathscr{C}$ and construct $C \equiv c^{\perp}$, then let s be the meet of ac and C, and t the meet of $A \equiv a^{\perp}$ and C.

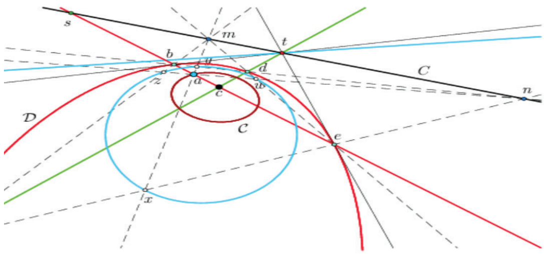


FIGURE 9.17: Constructing the twin circle $\mathscr{D}$ of $\mathscr{C}$

Now we can apply the construction of Figure 9.16; suppose that the side $\overline{st}$ has midpoints m and n, and that x and y are null points on am, and z and w are null points on an. Then $b \equiv (mz)(ac) = (ny)(ac)$ and $e \equiv (mw)(ac) = (nx)(ac)$ lie on the twin circle $\mathscr{D}$ to $\mathscr{C}$. Symmetry implies that we could also use $d \equiv (mw)(ct) = (ny)(ct)$ and $f \equiv (mz)(ct) = (nx)(ct)$.

Figure 9.18 shows another example of construction of the twin $\mathscr{D}$ of a given circle $\mathscr{C}$ (in brown) with center c. In this case c is outside the null circle, so its dual line C passes through null points x and y (approximately—

remember that a dynamic geometry package usually only deals with decimal approximations, so the number-theoretical subtlety is diminished). Choose a point a on $\mathscr{C}$ with dual line $A = a^{\perp}$. Then the twin circle $\mathscr{D}$ (in red) is the locus of the point $b = (ax)A$, or of the point $d = (ay)A$ as a moves along $\mathscr{C}$.

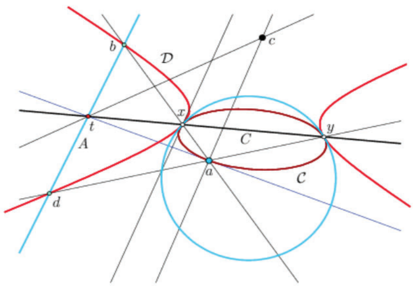


FIGURE 9.18: Another construction of a twin circle

n

The relation $q(a,c) + q(b,c) = 2$ then follows by applying either the Nil Cross law ([43, Thm 80]) or the Null subtended quadrance theorem ([43, Thm 90]), to the triangle $\overline{abc}$. Similarly, given the red circle $\mathscr{D}$, its twin circle $\mathscr{C}$ (in brown) is the locus of the point $a = (bx)b^{\perp}$, as the point b moves on $\mathscr{D}$.

9.4.3 The parabola in hyperbolic geometry

In Euclidean geometry, the parabola plays several distinguished roles. It is the graph resulting from a quadratic function $f(x) = a + bx + cx^2$, and it is also the second degree Taylor expansion of a general function. It is also a conic section in the spirit of Apollonius, obtained by slicing a cone with a plane which is parallel to one of its generatrices. In affine geometry the parabola is the distinguished conic which is tangent to the line at infinity. In everyday life, the parabola occurs in reflecting mirrors and automobile head lamps, in satellite dishes and radio telescopes, and in the trajectories of comets.

The ancient Greeks also studied the familiar metric formulation of a parabola: it is the locus of a point which remains equidistant from a fixed point F, called the *focus*, and a fixed line f (that does not contain F), called the *directrix*. Such a conic $\mathscr{P}$ has a line of symmetry: the *axis* a through F perpendicular to f. It also has a distinguished point V called the *vertex*, which is the only point of the parabola lying on the axis a, aside from the point at infinity. The vertex V is the midpoint between the focus F and the **base point** $B \equiv af$.

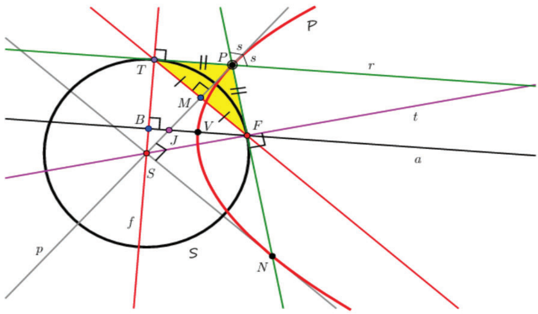


FIGURE 9.19: The Euclidean parabola

For such a classical parabola $\mathscr{P}$ hundreds of facts are known, see for example [1], [8], [12], [15], [22], [31] and [32]; quite a few of them going back to Archimedes and Apollonius. Of particular importance are theorems that relate to an arbitrary point P on the conic and its tangent line p. In particular, the construction of the tangent line p itself is important: there are two common ways of doing this. One is to take the foot T of the altitude from P to the directrix f, and connect P to the midpoint M of $\overline{TF}$; so that $p = PM$. Another is to take the perpendicular line t to PF through F, and find its meet S with the directrix; this gives $p = PS$. The point S is equidistant from T and F, and the circle $\mathscr{S}$ with center S through F is tangent to both the lines PF and PT.

A related and useful fact is that a chord PN is a focal chord—meaning that it passes through F—precisely when the meet of the two tangents at P and N lies on the directrix f, and in this case the two tangents are perpendicular. These facts are illustrated in Figure 9.19. Another result, which appears often in calculus, is that if P and Q are arbitrary points on the parabola with Z the meet of their tangents p and q, and T, U and W are respectively the feet of the altitudes from P, Q and Z to the directrix, then W is the midpoint of $\overline{TU}$.

So when investigating hyperbolic geometry, some natural questions arise, such as what the analog of a parabola in this context is, what properties of the Euclidean case carry over to this setting, and what additional properties might the hyperbolic parabola have that do not hold in the Euclidean case. These issues have of course been studied by quite a few authors, such as [5], [36], [37], [21] and [29].

There is a very natural analog of the idea of a parabola in this hyperbolic setting, and many, but certainly not all, properties of the Euclidean parabola hold or have reasonable analogues for it. But there are also many interesting aspects which have no Euclidean counterpart, such as the existence of a *dual* or *twin parabola*, and an intimate connection with the theory of sydpoints from the previous Chapter. In [2] and [3] we introduced and studied the rich canonical structure on a hyperbolic parabola, with lovely collinearities and concurrences.

In this Section definitions and some basic results for a *parabola* in *universal hyperbolic geometry* are introduced. As already discussed in [21], the definition is not entirely obvious: there are several different possible ways of generalizing the Euclidean theory. Recall that if a is a point and L is a line, then the quadrance $q(a,L)$ is defined to be the quadrance between a and the foot t of the altitude line from a to L. We next present the definition of the parabola.

Suppose that f_1 and f_2 are two non-perpendicular points such that $f_1 f_2$ is a non-null line. The **parabola** $\mathscr{P}_0$ with **foci** f_1 and f_2 is the locus of a point p_0 satisfying

$$q(f_1,p_0)+q(p_0,f_2)=1. \tag{9.11}$$

The lines $F_1 \equiv f_1^{\perp}$ and $F_2 \equiv f_2^{\perp}$ are the **directrices** of the parabola $\mathscr{P}_0$.

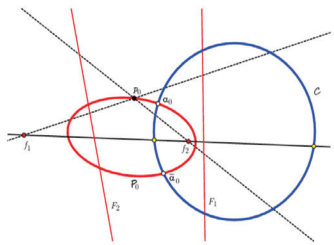


FIGURE 9.20: A parabola $\mathscr{P}_0$ with foci f_1 and f_2

So, at this point, there is no clear justification for the above definition. The following connects our theory with the more traditional approach in [14] and [28].

Theorem 9.14. *(Parabola focus directrix) A point p_0 satisfies (9.11) precisely when either of the following holds:*

$$q(f_1,p_0)=q(p_0,F_2) \qquad \text{or} \qquad q(f_2,p_0)=q(p_0,F_1).$$

In this way we recover the ancient Greek metric definition of the parabola, but it should be noted that there are *two* foci-directrix pairs: (f_1, F_2) and (f_2, F_1). This is a main feature of the hyperbolic theory of the parabola: a fundamental symmetry between the two foci-directrix pairs. In Figure 9.20 a parabola $\mathscr{P}_0$ is displayed in red, with foci f_1 and f_2, and directrices F_1 and F_2, also in red.

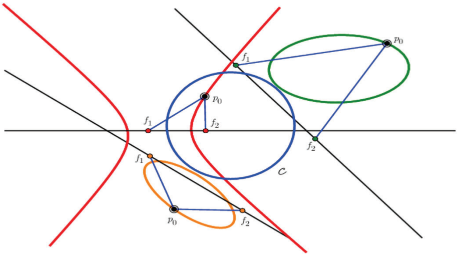


FIGURE 9.21: Various examples of parabolas

Figure 9.21 displays some different examples of parabolas over the rational numbers, at least approximately. When the foci f_1 and f_2 are both interior points of the null circle $\mathscr{C}$, there is no interior point p satisfying the condition $q(p, f_1) + q(p, f_2) = 1$, since the quadrance between any two interior points is always negative and the quadrance between an interior and an exterior points is always greater than or equal 1. Thus, the parabola with both foci inside the null circle, is an empty conic.

In the next theorem we find the equation of a parabola with given foci.

Theorem 9.15. *(Parabola conic) The parabola $\mathscr{P}_0$ with foci f_1 and f_2 is a conic.*

We now define some basic points and lines associated to a parabola $\mathscr{P}_0$ with foci f_1 and f_2, and directrices $F_1 \equiv f_1^{\perp}$ and $F_2 \equiv f_2^{\perp}$, as in Figure 9.22. The **axis** of the parabola $\mathscr{P}_0$ is the line $A \equiv f_1 f_2$. The **axis point** of $\mathscr{P}_0$ is the dual point $a \equiv A^{\perp}$. By assumption the axis A is a non-null line, so that a does not lie on A.

If the axis A has null points, such points are called the **axis null points** of $\mathscr{P}_0$ and are denoted by η_1 and η_2, in no particular order. In the diagrams in this Chapter, the axis point and line will generally be displayed in black, while the axis null points will be colored in yellow.

Theorem 9.16. *(Axis symmetry) The axis $A = f_1 f_2$ of a hyperbolic parabola $\mathscr{P}_0$ is a line of symmetry and its dual point a is a center.*

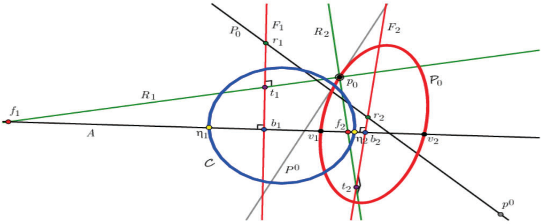


FIGURE 9.22: Dual and tangent lines, twin point and focal lines

9.4.4 Dual conics and the connection with sydpoints

The theory of the hyperbolic parabola connects strongly with the *notion of sydpoints*, as first described in [47].

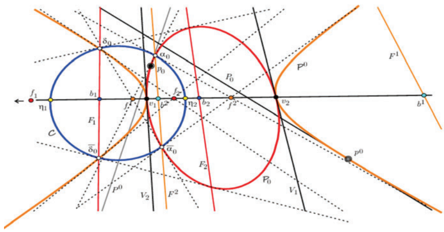


FIGURE 9.23: The parabola $\mathscr{P}_0$ and its twin $\mathscr{P}^0$

The reason is that the sydpoints f^1 and f^2 of the side $\overline{f_1 f_2}$, should they exist (and our assumptions on our field will guarantee that they do), are naturally determined by the geometry of the parabola $\mathscr{P}_0$. They become the foci for the *twin parabola* $\mathscr{P}^0$ (in orange in our diagrams), which turns out to be the dual of the conic $\mathscr{P}_0$ with respect to the null circle $\mathscr{C}$. This means that the dual of the tangent to a point p_0 on $\mathscr{P}_0$ is the point p^0 on the twin $\mathscr{P}^0$. The sydpoint symmetry between the sides $\overline{f_1 f_2}$ and $\overline{f^1 f^2}$ is key to understanding many aspects of these conics.
Figure 9.23 shows the parabola $\mathscr{P}_0$ with foci f_1, f_2 and a point p_0 on it, as well as the twin parabola $\mathscr{P}^0$ with foci f^1, f^2 and the twin point p^0 on it, which is the dual of the tangent P^0 to $\mathscr{P}_0$ at p_0. Reciprocally the dual of p_0 is the tangent to $\mathscr{P}^0$ at p^0. Note that the tangents to both the parabola $\mathscr{P}_0$ and the null circle $\mathscr{C}$ at their common meets, namely the null points α_0 and $\overline{\alpha_0}$, pass through the foci of the twin parabola $\mathscr{P}^0$! Dually, note that the tangents to both the parabola $\mathscr{P}^0$ and the null circle $\mathscr{C}$ at their common

meets, namely the null points δ_0 and $\overline{\delta_0}$, pass through the foci of $\mathscr{P}_0$! This Figure also shows the twin directrices F^1 and F^2, and the twin base points b^1 and b^2.

Bibliography

[1] A. V. Akopyan, A. A. Zaslavsky, *Geometry of conics*, American Mathematical Soc. 2007.

[2] A. Alkhaldi, N. J. Wildberger, *The Parabola in Universal Hyperbolic Geometry I*, KoG, **17** (2013), 14-41.

[3] A. Alkhaldi, N. J. Wildberger, *The Parabola in Universal Hyperbolic Geometry I*, J. Geometry and Graphics, **20** (2016), 1-11.

[4] J. W. Anderson, *Hyperbolic Geometry*, 2nd ed., Springer, 2005.

[5] O. Avcıoğlu, O. Bizim, *The Hyperbolic Conics in the Hyperbolic Geometry*, Adv. Studies in Contemporary Math., **16** (2008), 31-45.

[6] R. Baer, *Linear Algebra and Projective Geometry,* Academic Press, New York, 1952.

[7] M. Berger, *Geometry I*, Springer-Verlag, Berlin, 1987.

[8] M. Berger, *Geometry II*, Springer-Verlag, Berlin, 1987.

[9] M. Berger, *Geometry Revealed: A Jacob's Ladder to Modern Higher Geometry*, Springer, 2010.

[10] F. Bonahon, *Low-dimensional geometry: from Euclidean surfaces to hyperbolic knots*, American Mathematical Soc., 2009.

[11] O. Bottema, *On the medians of a triangle in hyperbolic geometry*, Can. J. Math., **10** (1958), 502-506.

[12] S. C. Choi, *The Universal Parabola*, Honours Thesis, School of Mathematics and Statistics, UNSW, 2011.

[13] H. S. M. Coxeter, *Non-Euclidean Geometry*, 6th ed., Mathematical Association of America, Washington D. C., 1998.

[14] G. Csima, J. Szirmai, *Isoptic curves of the conic sections in the hyperbolic and elliptic plane*, Stud. Univ. Žilina, Math.Ser. 24, **1** (2010), 15-22.

[15] T. de Alwis, *Normal Lines Drawn to a Parabola and Geometric Constructions*, Proceedings of the Third Asian Technology Conference in Mathematics, University of Tsukuba, Japan, 1998.

[16] K. Fladt, *Die Zentralprojektion einer Ebene auf eine andere in der hyperbolischen Geometrie*, Acta Math. Acad. Sci. Hungar, **14** (1963), 403–416.

[17] W. Graustein, *Introduction to Higher Geometry*, MacMillan, New York, 1935.

[18] M. J. Greenberg, *Euclidean and Non-Euclidean Geometries: Development and History*, 4th ed., W. H. Freeman and Co., San Francisco, 2007.

[19] R. Hartshorne, *Geometry: Euclid and Beyond*, Springer, New York, 2000.

[20] R. Hartshorne, *Foundations of Projective Geometry*, 2nd ed., Ishi Press, 2009.

[21] M. Henle, *Will the Real Non-Euclidean Parabola Please Stand up?*, Math. Magazine, **71:5** (1998), 369-376.

[22] K. Kendig, *Conics*, The Math. Assoc. of America, 2005.

[23] C. Kimberling, *Triangle Centers and Central Triangles*, **129**, Congressus Numerantium, Utilitas Mathematica Publishing, Winnepeg, MA, 1998.

[24] C. Kimberling, *Encyclopedia of Triangle Centers*, http://faculty.evansville.edu/ck6/encyclopedia/ETC.html.

[25] N. Le, N. J. Wildberger, *Universal Affine Triangle Geometry and Four-fold Incenter Symmetry*, KoG, **16** (2012), 63-80.

[26] J. Milnor, *Hyperbolic Geometry—The First 150 Years*, Bulletin of the AMS, **6** (1982), 9-24.

[27] E. Molnár, *A hiperbolikus geometria kúpszeleteinek aszimptota tulajdonságai, az aszimptoták és fókuszok kapcsolata szintetikus tárgyalásban*, Mat.Lapok, **22** (1971), 77–91.

[28] E. Molnár, *Kegelschnitte auf der metrischen Ebene*, Acta Mathematica Academiae Scientiarum Hungaricae, Tomus **21:3-4** (1978), 317–343.

[29] R. Perline, *Non-Euclidean Flashlights*, The American Mathematical Monthly, **103:5** (1996), 377-385.

[30] W. F. Reynolds, *Hyperbolic Geometry on a Hyperboloid*, Amer. Math. Monthly, **5** (1993), 442–455.

[31] G. Salmon, *A treatise on conic sections*, New York: Chelsea, 1960.

[32] J. H. Shackleton, *Elementary analytical conics*, 2nd ed., Oxford University Press, 1950.

[33] J. Stillwell, *Geometry of Surfaces*, Springer-Verlag, 1992.

[34] J. Stillwell, *Sources of Hyperbolic Geometry*, American Mathematical Soc., Providence RI, 1996.

[35] J. Stillwell, *Mathematics and its history*, 2nd ed., Springer-Verlag, New York, 2002.

[36] E. Story, *On the Non-Euclidean Geometry*, American Journal of Math., **5:1**, (1882), 180-211.

[37] E. Story, *On Non-Euclidean Properties of Conics*, American Journal of Math. (1883), 358-381.

[38] A. A. Ungar, *Hyperbolic barycentric coordinates*, Aust. J. Math. Anal. Appl., **6:1**, (2009), 1-35.

[39] A. A. Ungar, *Barycentric calculus in Euclidean and Hyperbolic Geometry; A comparative introduction*, World Scientific Publishing, Singapore, 2010.

[40] A. A. Ungar, *Hyperbolic Triangle Centers: The Special Relativistic Approach*, FTP 166, Springer, Dordrecht, 2010.

[41] N. J. Wildberger, *Divine Proportions: Rational Trigonometry to Universal Geometry*, Wild Egg Books, http://wildegg.com, Sydney, 2005.

[42] N. J. Wildberger, *A Rational Approach to Trigonometry*, Math Horizons, (2007), 16–20.

[43] N. J. Wildberger, *Affine and projective metrical geometry*, arXiv: math/0612499v1, (2006), to appear in J. of Geometry.

[44] N. J. Wildberger, *Universal Hyperbolic Geometry I: Trigonometry*, Geometriae Dedicata, (2013), 215-274.

[45] N. J. Wildberger, *Universal Hyperbolic Geometry II: A pictorial overview*, KoG, **14** (2010), 3-24.

[46] N. J. Wildberger, *Universal Hyperbolic Geometry III: First steps in projective triangle geometry*, KoG, **15** (2011), 25-49.

[47] N. J. Wildberger, A. Alkhaldi, *Universal Hyperbolic Geometry IV: Sydpoints and Twin Circumcircles*, KoG, **16** (2012), 43-62.

[48] C. R. Wylie, *Introduction to Projective Geometry*, McGraw-Hill, 1970.

Chapter 10

Metric fixed point theory in weighted graphs

M. R. Alfuraidan
Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia,
E-mail: monther@kfupm.edu.sa

In this Chapter, some known results about the fixed point problem on weighted graphs are discussed. We start by introducing the basic definitions; then we present the concept of monotone mapping with respect to the graph structure. The Chapter is concluded with a discussion of the metric fixed point theory for such mappings. The approximation of fixed points of monotone mappings will also be analyzed. It is worth mentioning that approximate fixed points are useful when computational issues are involved. In particular, most of the approximate fixed points treated in this Chapter are generated by an algorithm that is amenable to computational implementation.

10.1 Introduction

According to F. Browder, considered one of the forefathers of nonlinear functional analysis, ”The theory of fixed points is one of the most powerful tools of modern mathematics”. The origin of fixed point theory has its roots in the works of Poincaré, Lefschetz-Hopf, and Leray-Schauder. This theory is applied to a variety of areas of current interest in analysis. Topological considerations play a crucial role in the study of fixed points, it is worth remarking the relationship of fixed point theory with degree theory. Many mathematical problems involving existence translate into a question involving fixed point theory.

As its name suggests, metric fixed point theory is mainly concerned with problems that involve metric properties. The line dividing metric fixed point theory and its counterpart, topological fixed point theory is, as often the case, fuzzy. Successive approximations are at the core of the metric theory and are used to establish the existence and uniqueness of solutions. These successive approximation ideas were used by Cauchy, Fredholm, Liouville, Lipschitz, Peano and Picard. The Polish mathematician Banach is credited for organizing the ideas of successive approximations into an abstract framework, suitable for broad applications which go beyond the initial use in elementary differential and integral equations.

The fundamental fixed point theorem of Banach is the foundation of metric fixed point theory for contraction mappings, that is, for Lipschitz maps with a Lipschitz constant strictly between zero and one, on a complete metric space. Lipschitz maps with Lipschitz constant equal to one are called nonexpansive mappings. The fixed point problem for nonexpansive mappings differs sharply from that for contraction mappings, in the sense that additional structure of the domain set is needed to ensure the existence of fixed points.

Following the idea by Ran and Reurings [50] on the extension of the Banach contraction principle [12] to metric spaces endowed with a partial order, this new direction of research has attracted a great deal of attention in recent years. The idea in [50] was motivated by the investigation of some special matrix equations. It was Jachymski [33] who gave a more general formulation of these results by considering digraphs instead of a partial order. Since then, many publications appeared in this new direction and thus, a bridge was established between graph theory and metric fixed point theory, [1, 2, 3, 8, 6]. The approach in [50] and [33] is to define a digraph on a metric space and, based on the properties of the metric space, to prove some fixed point results. Our approach (see [5]) is different. We consider a weighted digraph and introduce some necessary topological structures on the set of vertices. These properties are analogous to the ones used in metric spaces, though we show that they are not equivalent. These considerations allow us to prove some fixed point results, which are more general than the ones found in the literature.

The focus in this Chapter is on the fixed points of different types of mappings defined on weighted graphs. It is the topological properties of these weighted graphs that constitute the main tools used in our approach.

A weighted digraph is a digraph with a numeric label associated to each edge. Such labels can be integers, rational numbers, or real numbers and might represent a concept such as distance, connection costs, or affinity. For example, when using a graph to represent roads between cities, if the problem under consideration is to find the fastest way to travel cross-country, it would not be appropriate for all edges to be equal, since some intercity distances will likely be much larger than others. It is thus natural to consider graphs whose edges are not weighted equally.

Because of its applications to industrial fields such as image processing engineering, computer science, economics, ladder networks, control theory, stochastic filtering and telecommunication, among others, this connection between graph theory and fixed point theory is presently attracting a great deal of scientific attention (see [5]).

10.2 Weighted graphs

A graph G consists of two sets, denoted by $V(G) \neq \emptyset$ and $E(G)$ (Or V and E if no confusion arises from this notation). Each element of V is called a vertex. The elements of E, called edges, are unordered pairs of vertices. A directed graph (also called digraph) is obtained by replacing the set E with a set D of ordered pairs of vertices. Replacing set E with a multiset, a so-called multigraph is obtained and allowing edges to connect a vertex to itself (i.e., by allowing loops), a pseudo-graph is obtained. Digraphs can have two edges with the same endpoints, provided they have opposite directions. The underlying graph $\widetilde{G}$ of a digraph is constructed by ignoring all directions and replacing any resulting multiple edges with single edges. All digraphs appearing in this Chapter are assumed to be reflexive i.e., it is assumed that each vertex has a loop.

The digraph G is said to be transitive if $(x,z) \in E(G)$ whenever $(x,y) \in E(G)$ and $(y,z) \in E(G)$, for any $x,y,z \in V(G)$. In other words, G is transitive if for any two vertices x and y that are connected by a directed finite path, it holds that $(x,y) \in E(G)$.

A weighted graph is a graph in which each edge is given a numerical weight. A weighted graph is therefore a special type of labeled graph, in which the labels are real positive numbers. Throughout this work, we consider weighted digraphs such that the weight of each edge is given by a distance function between the vertices.

10.2.1 Topological aspects of weighted graphs

The stage is now set for the introduction of the topological concepts of monotone mappings with respect to the graph structure.

Definition 10.1. Let G be a digraph. A sequence $(x_n) \in V(G)$ is said to be

(a) G-increasing, if $(x_n, x_{n+1}) \in E(G)$, for all $n \in \mathbb{N}$;

(b) G-decreasing, if $(x_{n+1}, x_n) \in E(G)$, for all $n \in \mathbb{N}$;

(c) G-monotone, if (x_n) is either G-increasing or G-decreasing.

In order to define the concept of G-compactness, some kind of sequential convergence in $V(G)$ is needed. For example, if τ is a topology on $V(G)$ one might consider τ-convergent sequences, but there are sequential convergences that are not associated to a topology. Still, the notation τ-convergence will continue to be used, even if τ is not a topology.

Definition 10.2. Let G be a digraph and τ be as described above. A nonempty subset C of $V(G)$ is said to be G_τ-compact if any G-increasing (resp. G-decreasing) sequence $(x_n) \in C$ has a subsequence $(x_{\phi(n)})$, that is τ-convergent to x in C and that satisfies the condition that $(x_{\phi(n)}, x) \in E(G)$ (resp. $(x, x_{\phi(n)}) \in E(G)$), for every $n \in \mathbb{N}$. In particular, if G is transitive it will hold that $(x_n, x) \in E(G)$, (resp. $(x, x_n) \in E(G)$) for every $n \in \mathbb{N}$.

Using the standard definition of τ-compactness in a metric space, a nonempty subset C of $V(G)$ is said to be τ-compact if and only if any sequence in C has a subsequence which τ-converges to a point in C. Note that if G is transitive and the G-intervals are τ-closed, then any τ-compact subset C of $V(G)$ is G_τ-compact.

Example 10.1. Consider the family of intervals $(I_s)_{s\in[0,+\infty)}$, in $\mathbb{R}^2$ defined by

$$I_s = \{(x,y)\ ;\ x = s \textit{ and } 0 \leq y \leq [s]+1\}.$$

On $\mathbb{R}^2$ define the digraph G by $\Big((x,y),(a,b)\Big) \in E(G)$ if and only if (x,y) and (a,b) belong to some I_s, for $s \in [0,+\infty)$, and $y \leq b$. It is clear that if $\Big((x_n,y_n)\Big)$ is a G-monotone sequence, then there exists $s_0 \in [0,+\infty)$ such that $(x_n,y_n) \in I_{s_0}$, for all $n \in \mathbb{N}$. If τ is the Euclidean topology, then $\mathbb{R}^2$ is G_τ-compact and any G-monotone sequence is bounded.

This example suggests the following definition.

Definition 10.3. Let G be a weighted digraph. Let d be a metric distance on $V(G)$. A nonempty $C \subseteq V(G)$ is said to be weakly G-bounded if any G-monotone sequence $(x_n)_{n\in\mathbb{N}}$ in C is bounded, i.e. $\delta((x_n)) = \sup\limits_{n,m\in\mathbb{N}} d(x_n,x_m) < \infty$.

It is clear that the motivation behind the introduction of the idea of G-compactness is the sequential characterization of metric compact sets as well as the use of monotone sequences in the study of fixed points of monotone mappings. Therefore, whenever a topological concept is characterized by sequences, it will have a similar extension to weighted graphs. The following definition illustrates this point.

Definition 10.4. Let G be a weighted digraph. Let d be a metric distance on $V(G)$. A nonempty set $C \subseteq V(G)$ is said to be G-complete (or a G-Cauchy space) if every G-monotone Cauchy sequence of vertices in C has a limit that is also in C.

Remark 10.1. It is surprising that in the Ran and Reurings extension [50] of the Banach contraction principle to partially metric spaces, one only needs to assume order-completeness, in the sense that monotone Cauchy sequences are convergent. Now one may ask whether such completeness is different from the metric completeness. A small modification of Example 10.1 will settle this question.

Example 10.2. Consider the set $C = \{(x,y) \in \mathbb{R}^2;\ 0 \leq x < 1 \textit{ and } 0 \leq y \leq 1\}$ and the family of intervals $(I_s)_{s\in[0,1)}$, in C defined by

$$I_s = \{(x,y)\ ;\ x = s \textit{ and } 0 \leq y \leq 1\}.$$

On C define the digraph G by $\Big((x,y),(a,b)\Big) \in E(G)$ if and only if (x,y) and (a,b) belong to some I_s, for $s \in [0,1)$, and $y \leq b$. It is clear that if $\Big((x_n,y_n)\Big)$ is a G-monotone sequence, then there exists $s_0 \in [0,1)$ such that $(x_n,y_n) \in I_{s_0}$, for all $n \in \mathbb{N}$. If τ is the Euclidean topology, then C is G-complete but not τ-complete.

10.2.2 Hyperbolic weighted graphs

In this Subsection, the concept of hyperbolic metric spaces defined on weighted graphs is introduced. Let (X,d) be a metric space. Suppose that there exists a family $\mathscr{F}$ of metric segments such that any two points x,y in X are endpoints of a unique metric segment $[x,y] \in \mathscr{F}$ (i.e., $[x,y]$ is an isometric image of the real line interval $[0,d(x,y)]$). We shall denote by $\beta x \oplus (1-\beta)y$ the unique point z of $[x,y]$ which satisfies

$$d(x,z) = (1-\beta)d(x,y), \text{ and } d(z,y) = \beta d(x,y),$$

where $\beta \in [0,1]$. Such metric spaces with a family $\mathscr{F}$ of metric segments are usually called *convex metric spaces* [46]. Moreover, under the assumption

$$d\Big(\alpha p \oplus (1-\alpha)x, \alpha q \oplus (1-\alpha)y\Big) \leq \alpha d(p,q) + (1-\alpha)d(x,y),$$

for all p,q,x,y in X, and $\alpha \in [0,1]$, X is said to be a *hyperbolic metric space* (see [53]).

Obviously, normed linear spaces are hyperbolic spaces. Hadamard manifolds [18] are nonlinear examples of hyperbolic metric spaces, as are the Hilbert open unit ball equipped with the hyperbolic metric [29], and the CAT(0) spaces [38, 39, 44]. A subset C of a hyperbolic metric space X is said to be convex if $[x,y] \subset C$ whenever x,y are in C.

Definition 10.5. Let (X,d) be a hyperbolic metric space. A graph G on X is said to be convex if and only if, for any $x,y,z,w \in X$ and $\alpha \in [0,1]$, it holds that

$$(x,z) \in E(G) \text{ and } (y,w) \in E(G) \Longrightarrow (\alpha x \oplus (1-\alpha)y, \alpha z \oplus (1-\alpha)w) \in E(G).$$

10.2.3 Modular weighted graphs

Let X be a nonempty set. Throughout this Chapter the following notation is agreed upon: for a function $\omega : (0,\infty) \times X \times X \to (0,\infty)$, we will write

$$\omega_\lambda(x,y) = \omega(\lambda,x,y),$$

for all $\lambda > 0$ and $x,y \in X$.

Definition 10.6. [20, 21] A function $\omega : (0,\infty) \times X \times X \to [0,\infty]$ is said to be a modular metric on X, if it satisfies the following axioms:

(a) $x = y$ if and only if $\omega_\lambda(x,y) = 0$, for all $\lambda > 0$;

(b) $\omega_\lambda(x,y) = \omega_\lambda(y,x)$, for all $\lambda > 0$, and $x,y \in X$;

(c) $\omega_{\lambda+\mu}(x,y) \leq \omega_\lambda(x,z) + \omega_\mu(z,y)$, for all $\lambda,\mu > 0$ and $x,y,z \in X$.

If condition (a), is replaced with (a'):

$$\omega_\lambda(x,x) = 0, \text{ for all } \lambda > 0,\ x \in X,$$

then ω is said to be a pseudomodular (metric) on X. A modular metric ω on X is said to be regular if the following weaker version of (a) is satisfied

$$x = y \text{ if and only if } \omega_\lambda(x,y) = 0, \text{ for some } \lambda > 0.$$

Finally, ω is said to be convex, if for $\lambda,\mu > 0$ and $x,y,z \in X$, it satisfies the inequality

$$\omega_{\lambda+\mu}(x,y) \leq \frac{\lambda}{\lambda+\mu}\omega_\lambda(x,z) + \frac{\mu}{\lambda+\mu}\omega_\mu(z,y).$$

Note that for a metric pseudomodular ω on a set X, and any $x,y \in X$, the function $\lambda \to \omega_\lambda(x,y)$ is nonincreasing on $(0,\infty)$. Indeed, if $0 < \mu < \lambda$, then

$$\omega_\lambda(x,y) \le \omega_{\lambda-\mu}(x,x) + \omega_\mu(x,y) = \omega_\mu(x,y).$$

Definition 10.7. [20, 21] Let ω be a pseudomodular on X. Fix $x_0 \in X$. The two sets

$$X_\omega = X_\omega(x_0) = \{x \in X : \omega_\lambda(x,x_0) \to 0 \ \ as \ \ \lambda \to \infty\},$$

and

$$X^*_\omega = X^*_\omega(x_0) = \{x \in X : \exists \lambda = \lambda(x) > 0 \text{ such that } \ \omega_\lambda(x,x_0) < \infty\}$$

are said to be modular spaces (around x_0).

It is obvious that $X_\omega \subset X^*_\omega$; in general this inclusion may be proper. It follows from [20, 21] that if ω is a modular on X, then the modular space X_ω can be equipped with a (nontrivial) metric, generated by ω and given by

$$d_\omega(x,y) = \inf\{\lambda > 0 : \omega_\lambda(x,y) \le \lambda\},$$

for any $x,y \in X_\omega$. If ω is a convex modular on X, according to [20, 21] the two modular spaces coincide, i.e. $X^*_\omega = X_\omega$, and this common set can be endowed with the metric d^*_ω given by

$$d^*_\omega(x,y) = \inf\{\lambda > 0 : \omega_\lambda(x,y) \le 1\},$$

for any $x,y \in X_\omega$. These distances will be called Luxemburg distances.

First attempts to generalize the classical Lebesgue function spaces L^p were made in the early 1930's by Orlicz and Birnbaum in connection with orthogonal expansions. With such generalization in mind, they considered spaces of functions with growth properties different from the power-type growth control provided by the L^p-norms. More precisely, they introduced the function spaces defined as follows:

$$L^\varphi = \left\{f : \mathbb{R} \to \mathbb{R};\ \exists \lambda > 0\ : \rho(\lambda f) = \int_{\mathbb{R}} \varphi\big(\lambda |f(x)|\big)\, dx < \infty\right\},$$

where $\varphi : [0,\infty] \to [0,\infty]$ was assumed to be a convex function increasing to infinity. In particular, the power function $\varphi(t) = t^p$, satisfies these conditions.

The notion of modular function spaces L^φ provides a wonderful example of a modular metric space. Indeed, define the function ω by

$$\omega_\lambda(f,g) = \rho\left(\frac{f-g}{\lambda}\right) = \int_{\mathbb{R}} \varphi\left(\frac{|f(x)-g(x)|}{\lambda}\right)\, dx,$$

for all $\lambda > 0$, and $f,g \in L^\varphi$, then ω is a modular metric on L^φ. Moreover the distance d^*_ω is exactly the distance generated by the Luxemburg norm on L^φ.

For more examples on modular function spaces, the reader my consult the book of Kozlowski [41]. For an exhaustive treatment of modular metric spaces we refer the reader to [20, 21].

Definition 10.8. Let X_ω be a modular metric space.

(1) The sequence $\{x_n\}_{n\in\mathbb{N}}$ in X_ω is said to be ω-convergent to $x \in X_\omega$, if and only if $\omega_1(x_n,x) \to 0$, as $n \to \infty$. The element x will be called the ω-limit of $\{x_n\}$.

(2) The sequence $\{x_n\}_{n\in N}$ in X_ω is said to be ω-Cauchy, if $\omega_1(x_m,x_n) \to 0$, as $m,n \to \infty$.

(3) A subset M of X_ω is said to be ω-closed, if the ω-limit of a ω-convergent sequence of M always belong to M.

(4) A subset M of X_ω is said to be ω-complete, if any ω-Cauchy sequence in M is a ω-convergent sequence and its ω-limit is in M.

(5) A subset M of X_ω is said to be ω-bounded, if it satisfies

$$\delta_\omega(M) = \sup\{\omega_1(x,y); x,y \in M\} < \infty.$$

(7) ω is said to satisfy the Fatou property, if and only if for any sequence $\{x_n\}_{n\in\mathbb{N}}$ in X_ω ω-convergent to x, we have

$$\omega_1(x,y) \le \liminf_{n\to\infty} \omega_1(x_n,y),$$

for any $y \in X_\omega$.

In general if $\lim_{n\to\infty} \omega_\lambda(x_n,x) = 0$, for some $\lambda > 0$, then it might not hold that $\lim_{n\to\infty} \omega_\lambda(x_n,x) = 0$, for all $\lambda > 0$. Therefore, following the usual terminology used in the theory of function spaces, ω is said to satisfy the Δ_2-condition, if $\lim_{n\to\infty} \omega_\lambda(x_n,x) = 0$, for some $\lambda > 0$ implies $\lim_{n\to\infty} \omega_\lambda(x_n,x) = 0$, for all $\lambda > 0$. The interested reader is referred to [20] and [21] for a discussion on the connection between ω-convergence and metric convergence with respect to the Luxemburg distances. In particular, one has

$$\lim_{n\to\infty} d_\omega(x_n,x) = 0 \text{ if and only if } \lim_{n\to\infty} \omega_\lambda(x_n,x) = 0, \text{ for all } \lambda > 0,$$

for any $\{x_n\} \in X_\omega$ and $x \in X_\omega$. Also, ω-convergence and d_ω convergence are equivalent if and only if the modular ω satisfies the Δ_2-condition. Moreover, if the modular ω is convex, it is well known that d^*_ω and d_ω are equivalent, which implies

$$\lim_{n\to\infty} d^*_\omega(x_n,x) = 0 \text{ if and only if } \lim_{n\to\infty} \omega_\lambda(x_n,x) = 0, \text{ for all } \lambda > 0,$$

for any $\{x_n\} \in X_\omega$ and $x \in X_\omega$ [20, 21]. It will be assumed in the sequel that ω satisfies the Fatou property.

10.3 Fixed point theory in weighted graphs

This Section is devoted to a detailed discussion of the latest fixed point results for monotone mappings defined on weighted graphs. Since the publication of the work by Ran and Reurings [50], the interest in the fixed point theory of such mappings has experienced unprecedented growth. The applications of this area are multiple. For example, the classical fixed point results are inadequate to handle the problem of finding a positive (or negative) solution of some equations.

The Section is organized as follows: First, basic definitions pertaining to the theory of monotone mappings on weighted graphs are given, secondly, some elementary facts needed for the proof of the main result will be presented.

10.3.1 Monotone mappings

Definition 10.9. Let G be a weighted digraph, d be a metric distance on $V(G)$ and C be a nonempty subset of $V(G)$. A mapping $T : C \to C$ is called

(a) G-monotone, if T is edge preserving, i.e., $(T(x),T(y)) \in E(G)$ whenever $(x,y) \in E(G)$, for any $x,y \in C$.

(b) G-contraction, if T is G-monotone and there exists $k \in [0,1)$ such that

$$\forall x,y \in C, (x,y) \in E(G) \Rightarrow d(T(x),T(y)) \leq k\, d(x,y).$$

(c) G-nonexpansive, if T is G-monotone and

$$\forall x,y \in C, (x,y) \in E(G) \Rightarrow d(T(x),T(y)) \leq d(x,y).$$

The point $x \in C$ is called a fixed point of T if $T(x) = x$.

Next, the concept of monotone multivalued mappings is discussed, [13].

Definition 10.10. ([13], Def. 2.6). Let $F : X \to 2^X$ be a multivalued mapping with nonempty closed and bounded values. The mapping F is said to be a G-contraction if there exists $k \in [0,1)$ such that

$$H(F(x),F(y)) \leq k\, d(x,y), \text{ for all } (x,y) \in E(G)$$

and such that, whenever $u \in F(x)$ and $v \in F(y)$ satisfy

$$d(u,v) \leq k\, d(x,y) + \alpha, \text{ for each } \alpha > 0,$$

then $(u,v) \in E(G)$.

In particular, this definition implies that if $u \in F(x)$ and $v \in F(y)$ are such that

$$d(u,v) \leq k\, d(x,y),$$

then $(u,v) \in E(G)$, which is very restrictive. In fact in the proof of Theorem 3.1 in [13], the authors tried unsuccessfully to construct an orbit (x_n) such that $(x_n,x_{n+1}) \in E(G)$, for any $n \geq 1$, which is impossible according to Definition 10.10. Our definition of G-contraction multivalued mappings is more appropriate. It finds its roots in [37].

Definition 10.11. [4] Let (X,d) be a metric space defined on a weighted graph G i.e., $V(G) = X$ and C be a nonempty subset of $V(G)$. A multivalued mapping $T : C \to 2^C$ is said to be a monotone increasing (resp. decreasing) G-contraction, if there exists $\alpha \in [0,1)$ such that for any $x,y \in C$ with $(x,y) \in E(G)$ and any $u \in T(x)$, there exists $v \in T(y)$ such that

$$(u,v) \in E(G)\ (resp.\ (v,u) \in E(G)) \ \ and \ \ d(u,v) \leq \alpha\, d(x,y).$$

Similarly, we will say that the multivalued mapping $T : C \to 2^C$ is monotone increasing (resp. decreasing) G-nonexpansive, if for any $x, y \in C$ with $(x,y) \in E(G)$ and any $u \in T(x)$, there exists $v \in T(y)$ such that

$$(u,v) \in E(G) \ (resp.\ (v,u) \in E(G)) \ \ and \ \ d(u,v) \leq d(x,y).$$

For a multivalued mapping T, x is a fixed point if and only if $x \in T(x)$. The set of all fixed points of a mapping T is denoted by $Fix(T)$.

10.3.2 Contraction monotone mappings

We begin with the following known theorems on the existence of a fixed point for monotone single-valued and multi-valued contraction mappings on metric spaces endowed with a graph structure.

Theorem 10.1. *[33] Let (X,d) be a complete metric space, and assume that the triple (X,d,G) has the following property:*

> *(J^*) For any $(x_n)_{n\geq 1}$ in X, if $x_n \to x$ and $(x_n, x_{n+1}) \in E(G)$, for $n \geq 1$, then there is a subsequence $(x_{k_n})_{n\geq 1}$ with $(x_{k_n}, x) \in E(G)$, for $n \geq 1$.*

Let $T : X \to X$ be a G-contraction and set $X_T := \{x \in X :\ (x, T(x)) \in E(G)\}$. Then the following statements hold:

1. $\big|Fix(T)\big| = \big|\{[x]_{\widetilde{G}} \ :\ x \in X_T\}\big|$.
2. *$Fix(T) \neq \emptyset$ if and only if $X_T \neq \emptyset$.*
3. *T has a unique fixed point if and only if there exists an $x_0 \in X_T$ such that $X_T \subseteq [x_0]_{\widetilde{G}}$.*
4. *For any $x \in X_T$, $T|_{[x]_{\widetilde{G}}}$ is a PO, that is, T has a unique fixed point $x^* \in [x]_{\widetilde{G}}$ and for each $x \in [x]_{\widetilde{G}}$, $\lim_{n\to\infty} T^n(x) = x^*$.*
5. *If $X_T \neq \emptyset$ and G is weakly connected, then T is a PO, that is T has a unique fixed point $x^* \in X$ and for each $x \in X$, $\lim_{n\to\infty} T^n(x) = x^*$.*

The multivalued version of Theorem 10.1 may be stated as follows:

Theorem 10.2. *[4] Let (X,d) be a complete metric space and suppose that the triple (X,d,G) has property (J^*). We denote by $\mathscr{CB}(X)$ the collection of all nonempty closed and bounded subsets of X. Let $T : X \to \mathscr{CB}(X)$ be a monotone increasing G-contraction mapping and $X_T := \{x \in X;\ (x,u) \in E(G)\ for\ some\ u \in T(x)\}$. If $X_T \neq \emptyset$, then the following statements hold:*

1. *For any $x \in X_T$, $T|_{[x]_{\widetilde{G}}}$ has a fixed point.*

2. *If $x \in X$ with $(x, \bar{x}) \in E(G)$ where $\bar{x}$ is a fixed point of T, then $\{T^n(x)\}$ converges to $\bar{x}$.*

3. *If G is weakly connected, then T has a fixed point in G.*

4. *If $X' := \bigcup\{[x]_{\widetilde{G}} : x \in X_T\}$, then $T|_{X'}$ has a fixed point in X.*

5. *If $T(X) \subseteq E(G)$, then T has a fixed point.*

6. *$Fix(T) \neq \emptyset$ if and only if $X_T \neq \emptyset$.*

Remark 10.2. The missing information in Theorem 10.2 is the uniqueness of the fixed point. In fact we do have a partial positive answer to this question. Indeed if $\bar{u}$ and $\bar{w}$ are two fixed points of T such that $(\bar{u}, \bar{w}) \in E(G)$, then necessarily $\bar{u} = \bar{w}$. In general T may have more than one fixed point.

Remark 10.3. Assuming G is such that $E(G) := X \times X$, then clearly G is connected and Theorem 10.2 yields Nadler's theorem [48].

The following is a direct consequence of Theorem 10.2.

Corollary 10.1. *Let (X, d) be a complete metric space. Let G be a graph on X such that the triple (X, d, G) has the Property **(J*)**. If G is weakly connected, then every G-contraction $T : X \to \mathscr{C}\mathscr{B}(X)$ such that $(x_0, x_1) \in E(G)$, for some $x_0 \in X$ and $x_1 \in T(x_0)$, has a fixed point.*

The following property, initially introduced in [49] for partially ordered sets and in [33] (Property (J^*) above) in metric spaces endowed with a graph, will be assumed in the sequel.

> (**Property ***) Let G be a weighted digraph and C be a nonempty subset of $V(G)$. C is said to have Property (*), if for any G-increasing (resp. G-decreasing) sequence $\{x_n\}$ in C which converges to x, there is a subsequence $\{x_{k_n}\}$ with $(x_{k_n}, x) \in E(G)$ (resp. $(x, x_{k_n}) \in E(G)$), for $n \in \mathbb{N}$.

Note that if G is a reflexive, transitive weighted graph, then the Property (*) implies the following property:

> (**Property ****) For any G-increasing (resp. G-decreasing) sequence $\{x_n\}$ in C, if $x_n \to x$, then $(x_n, x) \in E(G)$ (resp. $(x, x_n) \in E(G)$), for $n \in \mathbb{N}$.

10.3.3 Other types of Lipschitzian mappings

Following the publication of the Banach contraction principle (BCP) [12], there emerged multiple efforts to weaken its main assumptions. Most of the attention was focused on the Lipschitz condition.

10.3.3.1 Monotone almost contractions on weighted graphs

One of the first attempts to extend BCP was carried out by Kannan [34], followed by Chatterjea [19], Rus [54, 55], Taskovic [59] and Zamfirescu [61]. Berinde [14] was able to give a condition that captures most of the new concepts, which he called weak contraction and later on, almost contraction [15].

Definition 10.12. [15] Let (X,d) be a metric space. A map $T : X \to X$ is said to be an almost contraction if there exists $k < 1$ and $\theta \geq 0$ such that

$$d(T(x),T(y)) \leq k\, d(x,y) + \theta\, d(y,T(x)), \tag{AC}$$

for any $x,y \in X$.

It is obvious by symmetry 5hat condition (AC) is equivalent to

$$d(T(x),T(y)) \leq k\, d(x,y) + \theta\, d(x,T(y)),$$

for any $x,y \in X$. The following definition of an almost contraction is introduced because it captures most of the ideas behind the proofs of the existence of fixed points of such mappings.

Definition 10.13. Let (X,d) be a metric space. A map $T : X \to X$ is said to be a generalized almost contraction if there exist $k < 1$ and a function $\theta : [0,+\infty) \to [0,+\infty)$ which satisfies $\lim\limits_{t\to 0^+} \theta(t) = \theta(0) = 0$, such that

$$d(T(x),T(y)) \leq k\, d(x,y) + \min\Big\{\theta\, \Big(d(x,T(y))\Big), \theta\, \Big(d(y,T(x))\Big)\Big\}, \tag{GAC}$$

for any $x,y \in X$.

Remark 10.4. In [60], the authors gave an example of an almost contraction given by a convex averaging of nonexpansive mappings and for the identity. In particular, they considered the so-called (XU)- property. Specifically, let (X,d) be a metric space. A map $T : X \to X$ is said to satisfy the property (XU) in a nonempty subset K of X, if there exists $C \geq 1$ such that

$$d(x,y) \leq d(x,T(y)) \quad \textit{implies} \quad d(x,T(y)) \leq C\, d(x,y),$$

for any $x \neq y$ in K. In fact, in [60] the authors take K to be an open subset of X. In this case, the only map which satisfies the property (XU) is the identity map of K. To prove this claime notice that since $C \geq 1$, we have

$$d(y,T(x)) \leq C\, d(x,y),$$

for any $x \neq y$ in K. Hence,

$$d(x,T(x)) \leq d(x,y) + d(y,T(x)) \leq (1+C)\, d(x,y),$$

for any $x \neq y$ in K. Since K is open, we may choose $y \neq x$, arbitrarily close to x. Therefore, $d(x,T(x))$ smaller than any positive quantity. In other words, $T(x) = x$, for any $x \in K$.

Next, we discuss the concept of monotone almost contractions defined on weighted graphs. Aiming at facilitating the understanding of the main definitions, we denote by $\mathscr{C}_\prime(\mathbb{R}_+)$ the family of functions defined by $\theta \in \mathscr{C}_\prime(\mathbb{R}_+)$ if and only if $\theta : [0,+\infty) \to [0,+\infty)$ it satisfies $\lim_{t\to 0^+} \theta(t) = \theta(0) = 0$.

Definition 10.14. [9] Let (G,d) be a weighted digraph. A map $T : V(G) \to V(G)$ is said to be a monotone generalized almost contraction if

(a) T is G-monotone;

(b) there exist $k < 1$ and a function $\theta \in \mathscr{C}_\prime(\mathbb{R}_+)$ such that

$$d(T(x),T(y)) \leq k\, d(x,y) + \min\left\{\theta\, \Big(d(x,T(y))\Big), \theta\, \Big(d(y,T(x))\Big)\right\},$$

for any $x,y \in V(G)$ such that $(x,y) \in E(G)$.

Theorem 10.3. *[9] Let (G,d) be a weighted digraph. Assume G is G-complete and satisfies the property $(**)$. Let $T : V(G) \to V(G)$ be a monotone generalized almost contraction. Then T has a fixed point provided there exists $x_0 \in V(G)$ such that $(x_0, T(x_0)) \in E(\tilde{G})$.*

Proof. It can be assumed without loss of generality that $(x_0,T(x_0)) \in E(G)$. Since T is G-monotone , it follows that $(T^n(x_0), T^{n+1}(x_0)) \in E(G)$, for any $n \in \mathbb{N}$. Therefore $\{T^n(x_0)\}$ is a G-monotone sequence. Moreover, since T is a monotone generalized almost contraction, there exist $k < 1$ and a function $\theta \in \mathscr{C}_\prime(\mathbb{R}_+)$ such that

$$d(T(x),T(y)) \leq k\, d(x,y) + \min\left\{\theta\, \Big(d(x,T(y))\Big), \theta\, \Big(d(y,T(x))\Big)\right\},$$

for any $x,y \in V(G)$ such that $(x,y) \in E(G)$. Hence

$$d(T^{n+2}(x_0), T^{n+1}(x_0)) \leq k\, d(T^{n+1}(x_0), T^n(x_0)),$$

for any $n \in \mathbb{N}$. Obviously this yields

$$d(T^{n+1}(x_0), T^n(x_0)) \leq k^n\, d(T(x_0), x_0),$$

for any $n \in \mathbb{N}$. Since $k \in [0,1)$, we conclude that $\{T^n(x_0)\}$ is a Cauchy,

G-monotone sequence. Since G is G-complete, $\{T^n(x_0)\}$ converges to some point $z \in V(G)$. We claim that z is a fixed point of T. Indeed, note that if T is continuous, then this conclusion is obvious. Otherwise, we use the property $(**)$ satisfied by G, for this property implies tat $(T^n(x_0), z) \in E(G)$, for any $n \in \mathbb{N}$. Hence

$$\begin{aligned} d(T^{n+1}(x_0), T(z)) \quad &\leq \quad k\, d(T^n(x_0), z) \\ &\qquad + \min\Big\{\theta(d(T^n(x_0), T(z))), \theta(d(T^{n+1}(x_0), z))\Big\}, \end{aligned}$$

for any $n \in \mathbb{N}$. Using the properties of the function θ, it is readily concluded that

$$\lim_{n \to +\infty} \theta(d(T^{n+1}(x_0), z)) = 0,$$

which implies $\lim\limits_{n \to +\infty} d(T^{n+1}(x_0), T(z)) = 0$, i.e., $\{T^{n+1}(x_0)\}$ converges to $T(z)$. The uniqueness of the limit implies $T(z) = z$. □

The multivalued version of the BCP was obtained by Nadler [48]. Extensions and generalizations of Nadler's fixed point theorem were obtained by many mathematicians [26, 40].

Let (X,d) be a metric space. The Hausdorff-Pompeiu distance H on $\mathscr{CB}(\mathscr{X})$, the set of nonempty bounded and closed subsets of X, is defined by

$$H(A,B) = \max\left\{\sup_{b \in B} \inf_{a \in A} d(b,a),\ \sup_{a \in A} \inf_{b \in B} d(a,b)\right\},$$

for any $A, B \in \mathscr{CB}(\mathscr{X})$. The following technical result will shed light on Definition 10.15.

Lemma 10.1. *[48] Let (X,d) be a metric space. For any $A, B \in \mathscr{CB}(X)$, $\varepsilon > 0$, and for any $a \in A$, there exists $b \in B$ such that*

$$d(a,b) \leq H(A,B) + \varepsilon.$$

Using Lemma 10.1, a simpler definition of a monotone multivalued almost contraction can be formulated, one that in particular, avoids the use of the Hausdorff-Pompeiu distance.

Definition 10.15. [9] Let (G,d) be a weighted digraph. A map $T : V(G) \to \mathscr{C}(V(G))$ is said to be

(a) monotone, if whenever $x, y \in V(G)$ and satisfy $(x,y) \in E(G)$, then for any $\alpha \in T(x)$, there exists $\beta \in T(y)$ such that $(\alpha, \beta) \in E(G)$.

(b) a monotone generalized almost contraction, if T is monotone and there exist $k<1$ and a function $\theta \in \mathscr{C}_\prime(\mathbb{R}_+)$ such that for any $x,y \in V(G)$ with $(x,y) \in E(G)$ and any $\alpha \in T(x)$, there exists $\beta \in T(y)$ such that $(\alpha,\beta) \in E(G)$ and

$$d(\alpha,\beta) \leq k\, d(x,y) + \min\Big\{\theta\,\Big(dis(x,T(y))\Big), \theta\,\Big(dis(y,T(x))\Big)\Big\},$$

where $dis(x,A) = \inf\{d(x,a);\ a \in A\}$.

We are now ready for the multi-valued version Theorem 10.3.

Theorem 10.4. *[9] Let (G,d) be a weighted digraph. Assume G is G-complete and satisfies the property $(I*)$. Let $T : V(G) \to \mathscr{C}(V(G))$ be a monotone generalized almost contraction. If*

$$E_T := \{x \in V(G);\ \textit{there exists } y \in T(x) \textit{ such that } (x,y) \in E(\tilde{G})\} \neq \emptyset,$$

then T has a fixed point.

Proof. Note that if T has a fixed point $z \in V(G)$, then $z \in E_T$. Assume E_T is nonempty and let $x_0 \in E_T$. Without loss of generality, we will assume that there exists $x_1 \in T(x_0)$ such that $(x_0,x_1) \in E(G)$. Since T is a monotone generalized almost contraction, there exist $k<1$ and a function $\theta \in \mathscr{C}_\prime(\mathbb{R}_+)$ such that for any $x,y \in V(G)$ with $(x,y) \in E(G)$ and any $a \in T(x)$, there exists $b \in T(y)$ such that $(a,b) \in E(G)$ and

$$d(a,b) \leq k\, d(x,y) + \min\Big\{\theta\,\Big(dis(x,T(y))\Big), \theta\,\Big(dis(y,T(x))\Big)\Big\}.$$

In this case, there exists $x_2 \in T(x_1)$ such that

$$d(x_1,x_2) \leq k\, d(x_0,x_1) + \min\Big\{\theta\,\Big(dis(x_1,T(x_0))\Big), \theta\,\Big(dis(x_0,T(x_1))\Big)\Big\}.$$

Since $x_1 \in T(x_0)$, we get $dis(x_1,T(x_0)) = 0$, which implies

$$d(x_1,x_2) \leq k\, d(x_0,x_1).$$

By induction, a sequence $\{x_n\}$ can be constructed in $V(G)$ such that

(1) $(x_n,x_{n+1}) \in E(G)$;

(2) $x_{n+1} \in T(x_n)$;

(3) $d(x_{n+1},x_{n+2}) \leq k\, d(x_n,x_{n+1})$;

for any $n \in \mathbb{N}$. Condition (2) implies $d(x_{n+1}, x_n) \leq k^n\ d(x_0, x_1)$, for any $n \in \mathbb{N}$. Since $k \in [0,1)$, it follows that $\{x_n\}$ is a Cauchy G-monotone sequence (because of (1)). On account of the fact that G is G-complete, $\{x_n\}$ converges to some point $z \in V(G)$. We claim that z is a fixed point of T. Indeed, using the property $(**)$ satisfied by G, it follows that $(x_n, z) \in V(G)$ holds for any $n \in \mathbb{N}$. Hence for any $n \in \mathbb{N}$, there exists $z_n \in T(z)$ such that

$$d(x_{n+1}, z_n) \leq k\ d(x_n, z) + \min\Big\{\theta(dis(x_n, T(z))), \theta(dis(z, T(x_n))\Big\},$$

for any $n \in \mathbb{N}$. Since $x_{n+1} \in T(x_n)$ it is clear that $dis(z, T(x_n)) \leq d(z, x_{n+1})$, for any $n \in \mathbb{N}$. Hence $\lim\limits_{n\to+\infty} dis(z, T(x_n)) = 0$. The main property satisfied by θ implies

$$\lim_{n\to+\infty} \min\Big\{\theta(dis(x_n, T(z))), \theta(dis(z, T(x_n))\Big\} \leq \lim_{n\to+\infty} \theta(d(z, T(x_n)) = 0.$$

Therefore, $\lim\limits_{n\to+\infty} dis(x_{n+1}, z_n) = 0$, from which it follows that $\{z_n\}$ also converges to z. Since $T(z)$ is closed, necessarily $z \in T(z)$, i.e., z is a fixed point of T. □

10.3.3.2 Monotone quasi-contraction mappings on weighted graphs

The concept of quasi-contraction mappings was introduced by Ćirić [23] in dealing with a generalization of the Banach contraction principle. In this Subsection, the notion of quasi-contraction mappings is investigated in the framework of monotone mappings. In what follows, it will be assumed that (X, d) is a metric space, that G is a reflexive transitive weighted graph defined on X, that $E(G)$ has property (*) and that G-intervals are closed.

Definition 10.16. [2] Let C be a nonempty subset of X. A mapping $T : C \to C$ is called a G-monotone quasi-contraction if T is G-monotone and there exists $k < 1$ such that for any $x, y \in C$, $(x, y) \in E(G)$, it holds that

$$\begin{aligned} d(T(x), T(y)) \quad &\leq \quad k\ \max\Big(d(x,y); d(x, T(x)); d(y, T(y)); \\ &\qquad\qquad d(x, T(y)); d(y, T(x))\Big). \end{aligned}$$

An existence fixed point theorem for such mappings will be proved. First, let T and C be as in Definition 10.16. For any $x \in C$, define the orbit $\mathscr{O}(x) = \{x, T(x), T^2(x), \cdots\}$, and its diameter by

$$\delta(x) = \sup\{d(T^n(x), T^m(x)) : n, m \in \mathbb{N}\}\ .$$

The following technical lemma is crucial in the proof of the main result of this Section.

Lemma 10.2. *[2] Let (X,d) and G be as above. Let C be a nonempty subset of X and $T : C \to C$ be a G-monotone quasi-contraction mapping. Let $x \in C$ be such that $(x,T(x)) \in E(G)$ and that $\delta(x) < \infty$. Then for any $n \geq 1$, one has:*

$$\delta(T^n(x)) \leq k^n \delta(x) \,,$$

where $k < 1$ is the constant given in Definition 10.16. Moreover, it holds that

$$d(T^n(x), T^{n+m}(x)) \leq k^n \delta(x),$$

for any $n,m \in \mathbb{N}$.

Proof. Since T is G-monotone, one has $(T^n(x), T^{n+1}(x)) \in E(G)$, for any $n \in \mathbb{N}$. The transitivity of the graph G yields $(T^n(x), T^m(x)) \in E(G)$, for any $n,m \in \mathbb{N}$. Hence

$$\begin{aligned} d(T^n(x),T^m(x)) \quad \leq \quad k \; \max \Big(& d(T^{n-1}(x),T^{m-1}(x)); d(T^{n-1}(x),T^n(x)); \\ & d(T^{m-1}(x),T^m(x)); d(T^{n-1}(x),T^m(x)); \\ & \quad d(T^n(x),T^{m-1}(x)) \Big), \end{aligned}$$

for any $n,m \geq 1$. This obviously implies that

$$\delta(T^n(x)) \leq k \; \delta(T^{n-1}(x)) \,, \;\; n \geq 1.$$

Hence

$$\delta(T^n(x)) \leq k^n \; \delta(x) \,, \;\; n \geq 1,$$

from which it follows that

$$d(T^n(x), T^{n+m}(x)) \leq \delta(T^n(x)) \leq k^n \; \delta(x),$$

for any $n,m \in \mathbb{N}$. □

Using Lemma 10.3, the main result of this Section can now be proved:

Theorem 10.5. *[2] Let (X,d) and G be as above. Assume that (X,d) is complete. Let C be a closed, nonempty subset of X and $T : C \to C$ be a G-monotone quasi-contraction mapping. Let $x \in C$ be such that $(x,T(x)) \in E(G)$ and that $\delta(x) < \infty$. Then*

(a) $\{T^n(x)\}$ converges to $\omega \in C$, which is a fixed point of T and $(x,\omega) \in E(G)$. Moreover, we have

$$d(T^n(x), \omega) \leq k^n \; \delta(x), \;\; n \geq 1.$$

(b) If z is a fixed point of T such that $(x,z) \in E(G)$, then $z = \omega$.

Proof. For (i), observe that Lemma 10.3 implies that $\{T^n(x)\}$ is Cauchy. Since X is complete and C is closed, there exists $\omega \in C$ such that $\{T^n(x)\}$ converges to ω. Since

$$d(T^n(x), T^{n+m}(x)) \leq k^n\ \delta(x),\ \ n, m \in \mathbb{N},$$

letting $m \to \infty$ one gets

$$d(T^n(x), \omega) \leq k^n\ \delta(x)\ ,\ \ n \geq 1.$$

It follows from the G-monotonicity of T that $(T^n(x), T^{n+1}(x)) \in E(G)$, for any $n \geq 1$. By property (**), $(T^n(x), \omega) \in E(G)$, for any $n \geq 1$. In particular, $(x, \omega) \in E(G)$. In order to show that ω is a fixed point of T, note that

$$\begin{aligned} d(T^n(x), T(\omega)) \quad \leq \quad & k\ \max\Big(d(T^{n-1}(x), \omega); d(T^{n-1}(x), T^n(x)); \\ & d(\omega, T(\omega)); d(T^{n-1}(x), T(\omega)); d(T^n(x), \omega)\Big), \end{aligned}$$

for any $n \geq 1$. Letting $n \to +\infty$, it follows that $d(\omega, T(\omega)) \leq k\, d(\omega, T(\omega))$, which forces $d(\omega, T(\omega)) = 0$, since $k < 1$. Therefore, $T(\omega) = \omega$.
As for (ii), let $z \in C$ be a fixed point of T such that $(x, z) \in E(G)$. Then ,

$$d(T^n(x), z) \leq k\ \max\Big(d(T^{n-1}(x), z); d(T^{n-1}(x), T^n(x)); d(T^n(x), z)\Big),$$

for any $n \geq 2$ and letting $n \to +\infty$ one readily concludes that

$$d(\omega, z) = \limsup_{n \to \infty} d(T^n(x), z) \leq k\ \limsup_{n \to \infty} d(T^n(x), z) = k\, d(\omega, z).$$

Since $k < 1$, it follows that $d(\omega, z) = 0$, i.e., $\omega = z$. □

In the next Subsection, the validity of Theorem 10.5 in modular metric spaces is discussed. This is a very important class since spaces in it are similar to metric spaces in their structure but without the triangle inequality. Modular metric spaces are used in a wide range of applications.

First, the existence of fixed points for multivalued monotone Ćirić quasi-contraction on weighted graphs is studied. In the sequel, it is assumed that (X, d) is a metric space, that $\mathscr{C}\mathscr{B}(X)$ is the class of all nonempty closed and bounded subsets of X and that G is a reflexive weighted graph defined on X. Moreover, we assume that the triple (X, G, d) has property (*) and that G-intervals are closed.

Definition 10.17. [2] Let (X, G, d) be as above. A multivalued mapping $J : X \to \mathscr{C}(X)$ is called a G-monotone quasi-contraction, if there exists $k \in$

$[0,1)$ such that for any $a,b \in X$ with $(a,b) \in E(G)$ and any $A \in J(a)$, there exists $B \in J(b)$ such that $(A,B) \in E(G)$ and

$$d(A,B) \leq k \ \max\Big(d(a,b); d(a,A); d(b,B); d(a,B); d(b,A)\Big). \tag{10.1}$$

The point $a \in X$ is called a fixed point of J if $a \in J(a)$.

Example 10.3. Let $X = \{0,1,2,3\}$ and $d(x,y) = |x-y|$, $\forall x,y \in X$. Define the multivalued map $J : X \to \mathscr{C}(\mathscr{X})$ by:

$$J(x) = \{0,2,3\} \text{ for } x \in \{0,1\} \text{ and } J(x) = \{3\} \text{ for } x \in \{2,3\}.$$

Then J is a G-monotone quasi-contraction with $k = \frac{1}{3}$, where

$$G = \{(0,0),(1,1),(2,2),(3,3),(0,1),(2,3)\},$$

but is not a multivalued quasi-contraction since

$$d(0,3) > \frac{1}{3} \ \max\Big(d(1,2); d(1,0); d(2,3); d(1,3); d(2,0)\Big).$$

Next we discuss the existence of fixed points for such mappings. First, let J be as in Definition 10.17. For any $u_0 \in X$, the sequence $\{u_n\}$ defines an orbit of J at u_0 if $u_n \in J(u_{n-1})$, $n = 1,2,\cdots$.

The following technical Lemma is of vital importance in the proof of the main result of this Section.

Lemma 10.3. *[2] Let (X,G,d) be as above. Let $J : X \to \mathscr{C}(\mathscr{X})$ be a G-monotone multivalued quasi-contraction mapping. Let $u_0 \in X$ be such that $(u_0,u_1) \in E(G)$ for some $u_1 \in J(u_0)$. Assume that $k < \frac{1}{2}$, where k is the constant introduced in Definition 10.17 (for J). Then, there exists an orbit $\{u_n\}$ of J at u_0 such that $(u_n,u_{n+1}) \in E(G)$, for any $n \in \mathbb{N}$ and such that*

$$d(u_n,u_{n+1}) \leq \left(\frac{k}{1-k}\right)^n d(u_0,u_1).$$

Proof. The orbit $\{u_n\}$ of J at u_0 will be constructed by induction. Assume that $\{u_0,u_1,\cdots,u_n\}$ have been found such that $u_{i+1} \in J(u_i)$, $(u_i,u_{i+1}) \in E(G)$ and that

$$d(u_i,u_{i+1}) \leq \left(\frac{k}{1-k}\right)^i d(u_0,u_1),\ i = 0,\cdots,n-1.$$

Since J is a G-monotone multivalued quasi-contraction mapping, there exists $u_{n+1} \in J(u_n)$ such that

$$d(u_n, u_{n+1}) \leq k \max\Big(d(u_{n-1}, u_n); d(u_n, u_{n+1}); d(u_{n-1}, u_{n+1}); d(u_n, u_n)\Big).$$

Obviously this implies

$$\begin{aligned} d(u_n, u_{n+1}) &\leq k \max\Big(d(u_{n-1}, u_n); d(u_{n-1}, u_{n+1})\Big) \\ &\leq k \max\Big(d(u_{n-1}, u_n); d(u_{n-1}, u_n) + d(u_n, u_{n+1})\Big) \\ &\leq k \Big(d(u_{n-1}, u_n) + d(u_n, u_{n+1})\Big). \end{aligned}$$

Hence

$$d(u_n, u_{n+1}) \leq \frac{k}{1-k}\, d(u_{n-1}, u_n) \leq \left(\frac{k}{1-k}\right)^n d(u_0, u_1).$$

The proof of Lemma 10.3 follows by induction. □

Next we state the main result of this Subsection.

Theorem 10.6. *[4] Let (X, d) be a complete metric space and G be a reflexive weighted graph defined on X such that (X, G, d) has Property (*). Let $J : X \to \mathscr{C}(\mathscr{X})$ be a G-monotone multivalued quasi-contraction mapping. Let $u_0 \in X$ be such that $(u_0, u_1) \in E(G)$ for some $u_1 \in J(u_0)$. Assume that $k < \frac{1}{2}$, where k is the constant given in Definition 10.17 for J. Then there exists an orbit $\{u_n\}$ of J at u_0, which converges to $\omega \in X$, a fixed point of J.*

Proof. The orbit sequence $\{u_n\}$ of J at u_0 obtained in Lemma 10.3, is Cauchy. Since X is complete, there exists $\omega \in X$ such that $\{u_n\}$ converges to ω. Since $(u_n, u_{n+1}) \in E(G)$, for any $n \geq 1$, Property (P) implies that there is a subsequence $\{u_{k_n}\}$ such that $(u_{k_n}, \omega) \in E(G)$, for any $n \geq 0$. It is next shown that ω is a fixed point of J, i.e., that $\omega \in J(\omega)$. Since $u_{k_n+1} \in J(u_{k_n})$ and $(u_{k_n}, \omega) \in E(G)$, there exists $\omega_n \in J(\omega)$ such that

$$\begin{aligned} d(u_{k_n+1}, \omega_n) \leq k \max\Big(&d(u_{k_n}, \omega); d(u_{k_n}, u_{k_n+1}); d(\omega, \omega_n); \\ &d(u_{k_n}, \omega_n); d(u_{k_n+1}, \omega)\Big), \end{aligned}$$

for any $n \geq 1$. In particular, we have

$$\begin{aligned} d(u_{k_n+1}, \omega_n) \leq k \Big(&d(u_{k_n}, \omega) + d(u_{k_n}, u_{k_n+1}) + d(\omega, \omega_n) \\ &+ d(u_{k_n}, \omega_n) + d(u_{k_n+1}, \omega)\Big), \end{aligned}$$

for any $n \geq 1$. Since $d(\omega, \omega_n) - d(\omega, u_{k_n+1}) \leq d(u_{k_n+1}, \omega_n)$, we get

$$\begin{aligned}(1-2k)d(\omega, \omega_n) - d(\omega, u_{k_n+1}) \quad &\leq \quad k\, \Big(d(u_{k_n}, \omega) + d(u_{k_n}, u_{k_n+1}) \\ &\qquad + d(u_{k_n}, \omega) + d(u_{k_n+1}, \omega)\Big),\end{aligned}$$

for any $n \geq 1$. Hence

$$(1-2k) \limsup_{n \to +\infty} d(\omega_n, \omega) \leq 0,$$

which implies $\lim_{n \to +\infty} d(\omega_n, \omega) = 0$, since $k < 1/2$. Therefore $\{\omega_n\}$ converges to ω and since $J(\omega)$ is closed, it follows that $\omega \in J(\omega)$. In conclusion, ω is a fixed point of J.

□

If G in Theorem 10.6 is assumed to be transitive, one has $(u_0, \omega) \in E(G)$.

Remark 10.5. It is not clear to the authors whether the conclusion of Theorem 10.6 is still valid when $k < 1$.

10.3.3.3 Monotone Gregus-Ćirić mappings on weighted graphs

In 1980, Gregus [30] proved the following result:

Theorem 10.7. *Let X be a Banach space and C be a nonempty, closed and convex subset of X. Let $T : C \to C$ be a mapping satisfying*

$$\|T(x) - T(y)\| \leq a\|x - y\| + p\|T(x) - x\| + p\|T(y) - y\|,$$

for all $x, y \in C$, where $0 < a < 1$, $p \geq 0$ and $a + 2p = 1$. Then T has a unique fixed point.

Ćirić [23] obtained the following generalization of Gregus' theorem.

Theorem 10.8. *Let (X, d) be a complete, convex, metric space and C be a nonempty, closed and convex subset of X. Let $T : C \to C$ be a mapping satisfying*

$$\begin{aligned}d(T(x), T(y)) \quad &\leq \quad a \, \max\Big\{d(x, y), c\, \Big[d(x, T(y)) + d(y, T(x))\Big]\Big\} \\ &\qquad + b\, \max\{d(x, T(x)), d(y, T(y))\},\end{aligned} \tag{CG}$$

for all $x, y \in C$, where $0 < a < 1$, $a + b = 1$ and $0 \leq c \leq \frac{4-a}{8-a}$. Then T has a unique fixed point.

Remark 10.6. Under the assumption that $a+b<1$ and $c \leq \frac{1}{2}$, any map T which satisfies the condition (CG) also satisfies the following condition:

$$\begin{aligned} d(T(x),T(y)) \quad \leq \quad & (a+b) \max\Big\{d(x,y), d(x,T(y)), d(y,T(x)), \\ & d(x,T(x)), d(y,T(y))\Big\}. \end{aligned}$$

In other words, T is a Ćirić quasi-contraction mapping. This concept was introduced by Ćirić [22] as an extension to the contraction condition. In [22], he proved an analogue to the Banach contraction principle for this type of mapping, without the use of convexity.

Recently, Djafari-Rouhani and Moradi [24] obtained the following improvement of Ćirić's result:

Theorem 10.9. *Let (X,d) be a complete, convex metric space and $T: X \to X$ be a mapping satisfying*

$$\begin{aligned} d(T(x),T(y)) \quad \leq \quad & a \max\{d(x,y), c[d(x,T(y))+d(y,T(x))]\} \\ & +b \max\{d(x,T(x)), d(y,T(y))\}, \end{aligned}$$

for all $x,y \in X$, where $0<a<1$, $a+b=1$ and $0 \leq c < \frac{1}{2}$. Then T has a unique fixed point.

In fact, in [24], the authors give a simple example showing that the conclusion of Theorem 10.9 does not hold if $c > \frac{1}{2}$ and asked whether the conclusion holds when $c = \frac{1}{2}$. This problem is still open.

In this work, we generalize Theorem 10.9 to the case of monotone self-mappings defined on a weighted graph.

Definition 10.18. [7] Let G be a weighted digraph and d be a metric distance on $V(G)$. Let C be a nonempty subset of $V(G)$. A mapping $T: C \to C$ is called

(a) G-monotone Gregus-Ćirić-mapping, if T is G-monotone and there exist $a,b,c \in [0,+\infty)$ such that

$$\begin{aligned} d(T(x),T(y)) \quad \leq \quad & a \max\Big\{d(x,y), c\Big[d(x,T(y))+d(y,T(x))\Big]\Big\} \\ & +b \max\Big\{d(x,T(x)), d(y,T(y))\Big\}, \end{aligned}$$

for any $x,y \in C$ with $(x,y) \in E(G)$.

(b) G-monotone Gregus-Ćirić contraction, if T is a G-monotone Gregus-Ćirić-mapping with $0<a<1$, $a+b=1$ and $c \leq \frac{1}{2}$.

Note that in the example given by the authors in [24], the mapping $T(x) = x+1$ is monotone for the corresponding order and may be seen as an example of a monotone Gregus-Ćirić mapping. Moreover, any monotone contraction is a monotone Gregus-Ćirić-contraction. The example studied by Ran and Reurings [50] shows that a monotone-contraction may fail to be a contraction.

Let C be a nonempty subset of $V(G)$ and $T : C \to C$ be G-monotone Gregus-Ćirić-contraction mapping. Then there exist positive numbers a, b, c, such that $0 < a < 1$, $a+b = 1$ and $c \leq \frac{1}{2}$ such that

$$\begin{aligned} d(T(x),T(y)) \quad &\leq \quad a \ \max\Big\{d(x,y), c\Big[d(x,T(y)) + d(y,T(x))\Big]\Big\} \\ &\qquad + b \ \max\Big\{d(x,T(x)), d(y,T(y))\Big\}, \end{aligned}$$

for any $x, y \in C$ with $(x,y) \in E(G)$.

The following technical results will be crucial to the establishment of the main Theorem of this Section.

Lemma 10.4. *[7] Under the above assumptions, it holds that*

$$d(x,y) \leq \frac{2-a}{1-a} \ \Big(d(x,T(x)) + d(y,T(y))\Big),$$

for any $x, y \in C$ with $(x,y) \in E(G)$ or $(y,x) \in E(G)$.

Proof. Without loss of generality, it may be assumed that $(x,y) \in E(G)$. It then follows that

$$\begin{aligned} d(T(x),T(y)) \quad &\leq \quad a \ \max\Big\{d(x,y), c\Big[d(x,T(y)) + d(y,T(x))\Big]\Big\} \\ &\qquad +b \ \max\Big\{d(x,T(x)), d(y,T(y))\Big\}. \end{aligned}$$

Since $c \leq \frac{1}{2}$, it is concluded that

$$\begin{aligned} c\Big[d(x,T(y)) + d(y,T(x))\Big] \quad &\leq \quad c\Big[d(x,T(x)) + 2d(T(x),T(y)) \\ &\qquad +d(y,T(y))\Big] \\ &\leq \quad d(x,T(x)) + d(T(x),T(y)) + d(y,T(y), \end{aligned}$$

which implies

$$\begin{array}{rcl} d(T(x),T(y)) & \leq & a \max\Big\{d(x,T(x))+d(T(x),T(y))+d(y,T(y)), \\ & & c\Big[d(x,T(y))+d(y,T(x))\Big]\Big\} \\ & & +b \max\Big\{d(x,T(x)),d(y,T(y))\Big\} \\ & \leq & a\,\Big\{d(x,T(x))+d(T(x),T(y))+d(y,T(y))\Big\} \\ & & +b\,\Big\{d(x,T(x))+d(y,T(y))\Big\} \\ & \leq & (a+b)\Big(d(x,T(x))+d(T(x),T(y))\Big) \\ & & +a\,d(T(x),T(y)). \end{array}$$

Next, the equality $a+b=1$ yields

$$d(T(x),T(y)) \leq \frac{1}{1-a}\,\Big(d(x,T(x))+d(T(x),T(y))\Big).$$

Hence

$$\begin{array}{rcl} d(x,y) & \leq & d(x,T(x))+d(T(x),T(y))+d(y,T(y), \\ & \leq & \big(1+\frac{1}{1-a}\big)\,\Big(d(x,T(x))+d(T(x),T(y))\Big), \end{array}$$

which implies

$$d(x,y) \leq \frac{2-a}{1-a}\,\Big(d(x,T(x))+d(y,T(y))\Big).$$

□

Lemma 10.5. *[7] Under the above assumptions, if $x \in C$ is such that $(x,T(x)) \in E(G)$ or $(T(x),x) \in E(G)$, then the sequence $\Big\{d(T^n(x),T^{n+1}(x))\Big\}$ is decreasing.*

Proof. Without loss of generality, assume that $(x,T(x)) \in E(G)$. It follows from the monotonicity of T that $(T^n(x),T^{n+1}(x)) \in E(G)$, for any $n \in \mathbb{N}$. Fix $n \geq 1$. Then

$$\begin{array}{rcl} d(T^n(x),T^{n+1}(x)) & \leq & a \max\Big\{d(T^{n-1}(x),T^n(x)), \\ & & c\,d(T^{n-1}(x),T^{n+1}(x))\Big\} \\ & & +\,b \max\Big\{d(T^{n-1}(x),T^n(x)), \\ & & d(T^n(x),T^{n+1}(x))\Big\}. \end{array}$$

Assume that $d(T^{n-1}(x),T^n(x)) < d(T^n(x),T^{n+1}(x))$ holds. Since

$$\begin{array}{rcl} d(T^{n-1}(x),T^{n+1}(x)) & \leq & d(T^{n-1}(x),T^n(x))+d(T^n(x),T^{n+1}(x)) \\ & < & 2\,d(T^n(x),T^{n+1}(x)), \end{array}$$

and $c \leq \frac{1}{2}$, one has

$$\begin{array}{rcl} d(T^n(x),T^{n+1}(x)) & < & a\,d(T^n(x),T^{n+1}(x))+b\,d(T^n(x),T^{n+1}(x)) \\ & = & d(T^n(x),T^{n+1}(x)). \end{array}$$

This contradiction implies $d(T^n(x),T^{n+1}(x)) \leq d(T^{n-1}(x),T^n(x))$. Since n was taken arbitrarily, it clear that $\left\{d(T^n(x),T^{n+1}(x))\right\}_{n\in\mathbb{N}}$ is decreasing. $\square$

Lemma 10.6. *[7] Under the above assumptions, if G is transitive and $x \in C$ satisfies $(x,T(x)) \in E(G)$ or $(T(x),x) \in E(G)$, then there exists $n \geq 1$ such that*

$$d(T^n(x),T^{n+2}(x)) \leq \frac{2}{2-a}\,d(x,T(x)).$$

Proof. The reasoning mimics the argument used by Djafari-Rouhani and Moradi in their proof of [[24], Theorem 2.2]. Without loss of generality, assume that $(x,T(x)) \in E(G)$. Since G is transitive and T is G-monotone, then $(T^n(x),T^{n+h}(x)) \in E(G)$, for any $n,h \in \mathbb{N}$. Fix $n \geq 1$. Then,

$$\begin{array}{rcl} d(T^n(x),T^{n+2}(x)) & \leq & a\,\max\Big\{d(T^{n-1}(x),T^{n+1}(x)), \\ & & \quad c\left[d(T^{n-1}(x),T^{n+2}(x))+\,d(T^n(x),T^{n+1}(x))\right]\Big\} \\ & & \quad +\,b\,\max\Big\{d(T^{n-1}(x),T^n(x)), \\ & & \quad d(T^{n+1}(x),T^{n+2}(x))\Big\}. \end{array}$$

Assume that for some $n \geq 1$ it holds that

$$d(T^{n-1}(x),T^{n+1}(x)) \leq c\left[d(T^{n-1}(x),T^{n+2}(x))+\,d(T^n(x),T^{n+1}(x))\right].$$

Since $\left\{d(T^n(x),T^{n+1}(x))\right\}_{n\in\mathbb{N}}$ is decreasing, it follows that

$$\begin{array}{rcl} d(T^n(x),T^{n+2}(x)) & \leq & ac\left[d(T^{n-1}(x),T^{n+2}(x))+\,d(T^n(x),T^{n+1}(x))\right] \\ & & \quad +\,b\,d(x,T(x)). \end{array}$$

From $d(T^{n-1}(x),T^{n+2}(x)) \leq d(T^{n-1}(x),T^n(x))+d(T^n(x),T^{n+2}(x))$ it follows that

$$\begin{array}{rcl} d(T^n(x),T^{n+2}(x)) & \leq & ac\left[2\,d(x,T(x))+\,d(T^n(x),T^{n+2}(x))\right] \\ & & \quad +\,b\,d(x,T(x)), \end{array}$$

which implies

$$d(T^n(x), T^{n+2}(x)) \leq \frac{2ac+b}{1-ac}\ d(x, T(x)).$$

The function $f(c) = \frac{2ac+b}{1-ac}$ is increasing in the interval $[0, \frac{1}{2}]$. Hence

$$\frac{2ac+b}{1-ac} \leq \frac{a+b}{1-a/2} = \frac{2}{2-a}.$$

It is therefore concluded that

$$d(T^n(x), T^{n+2}(x)) \leq \frac{2}{2-a}\ d(x, T(x)).$$

Next, assume that for any $n \geq 1$ it holds

$$c\ \Big[d(T^{n-1}(x), T^{n+2}(x)) +\ d(T^n(x), T^{n+1}(x))\Big] \leq d(T^{n-1}(x), T^{n+1}(x)).$$

In this case,

$$d(T^n(x), T^{n+2}(x)) \leq a\ d(T^{n-1}(x), T^{n+1}(x)) + b\ d(x, T(x)),$$

which easily implies

$$\begin{array}{rcl} d(T^n(x), T^{n+2}(x)) & \leq & a^{n-1}\ d(x, T^2(x)) + \frac{b}{1-a}\ d(x, T(x) \\ & = & a^{n-1}\ d(x, T^2(x)) + d(x, T(x). \end{array}$$

Since $d(x, T^2(x)) \leq d(x, T(x)) + d(T(x), T^2(x)) \leq 2\ d(x, T(x))$ one concludes thhat

$$d(T^n(x), T^{n+2}(x)) \leq (2\ a^{n-1}\ + 1)\ d(x, T(x),$$

for any $n \geq 1$. Since $0 < a < 1$, there exists $n \geq 1$ such that

$$2\ a^{n-1}\ + 1 \leq \frac{a}{2-a} + 1 = \frac{2}{2-a},$$

which implies

$$d(T^n(x), T^{n+2}(x)) \leq \frac{2}{2-a}\ d(x, T(x)).$$

$\square$

The final basic result of this Subsection is the following:

Lemma 10.7. *[7] Let a, b, c be positive numbers such that $0 < a < 1$, $a + b = 1$ and that $c < \frac{1}{2}$. Then for $\beta \geq 0$ such that $2c < \beta < 1$, it holds that*

$$K = \alpha\, a\, \max\left\{\alpha + \frac{2\beta}{2-a}, c\left[2 + \frac{2\beta}{2-a}\right]\right\} + \beta^2\, a + b < 1,$$

where $\alpha = 1 - \beta$.

Proof. Note that $K < 1$ if and only if

$$\alpha\, \max\left\{\alpha + \frac{2\beta}{2-a}, c\left[2 + \frac{2\beta}{2-a}\right]\right\} + \beta^2 < 1,$$

where $1 - b = a$ and $a > 0$. It follows from the equality $1 - \beta^2 = \alpha(1 + \beta)$ and from the fact that $\alpha > 0$, that $K < 1$ if and only if

$$\max\left\{\alpha + \frac{2\beta}{2-a}, c\left[2 + \frac{2\beta}{2-a}\right]\right\} < 1 + \beta.$$

Since $a < 1$, one has $1 < 2 - a$ which implies $\dfrac{2\beta}{2-a} < 2\beta$. Hence

$$\alpha + \frac{2\beta}{2-a} = 1 - \beta + \frac{2\beta}{2-a} < 1 + \beta.$$

Moreover, $\beta^2 < 1 < 2 - a$ and since $2c < \beta$, one readily concludes that

$$2c\left[1 + \frac{\beta}{2-a}\right] < \beta\left[1 + \frac{\beta}{2-a}\right] = \beta + \frac{\beta^2}{2-a} < 1 + \beta.$$

Therefore,

$$\max\left\{\alpha + \frac{2\beta}{2-a}, c\left[2 + \frac{2\beta}{2-a}\right]\right\} < 1 + \beta,$$

which completes the proof that $K < 1$. □

In what follows, the existence of fixed points of G-monotone Gregus-Ćirić mappings defined on weighted graphs is discussed. As mentioned earlier, the fixed point results for this type of mappings were obtained in the context of convex metric spaces. Throughout this Subsection, G will stand for a transitive weighted digraph with d a metric distance on $V(G)$. It will also be assumed that $V(G)$ is a convex metric space, such that G-intervals are convex. The next Theorem is the main fixed point result of this Section.

Theorem 10.10. *[7] Let C be a nonempty, G-complete and convex subset of $V(G)$ that satisfies the Property (*). Let $T : C \to C$ be a G-monotone*

Gregus-Ćirić contraction mapping, i.e., assume there exist positive numbers a, b, c with $0 < a < 1$, $a + b = 1$ and $c \leq \frac{1}{2}$, such that

$$\begin{aligned} d(T(x), T(y)) \quad &\leq \quad a \max\Big\{d(x,y), c\Big[d(x,T(y)) + d(y,T(x))\Big]\Big\} \\ &\qquad + b \max\Big\{d(x,T(x)), d(y,T(y))\Big\}, \end{aligned}$$

for any $x, y \in C$ with $(x,y) \in E(G)$. Assume that $c < \frac{1}{2}$. Let $x \in C$ be such that $(x, T(x)) \in E(G)$ (or $(T(x), x) \in E(G)$). Then T has a fixed point ω such that $(x, \omega) \in E(G)$ (or $(\omega, x) \in E(G)$). Moreover, if Ω is another fixed point of T with $(x, \Omega) \in E(G)$ (or $(\Omega, x) \in E(G)$), then necessarily $\omega = \Omega$.

Proof. It can be assumed without loss of generality that $(x, T(x)) \in E(G)$ and that x is not a fixed point of T. In this case, $(T^n(x), T^{n+1}(x)) \in E(G)$, for any $n \in \mathbb{N}$. Lemma 10.6 implies the existence of $n \geq 1$ such that

$$d(T^n(x), T^{n+2}(x)) \leq \frac{2}{2-a}\, d(x, T(x)).$$

Let $\beta < 1$ be the number obtained in Lemma 10.7. Set

$$z = \alpha\, T^{n+1}(x) \oplus \beta\, T^{n+2}(x) \in C,$$

since C is convex. Using the convexity of the G-intervals, it can be easily shown that $(T^{n+1}(x), z) \in E(G)$ and that $(z, T^{n+2}(x)) \in E(G)$. Since T is G-monotone and G is transitive, one must have that $(z, T(z)) \in E(G)$ and $(T^n(x), z) \in E(G)$. Moreover,

$$d(z, T(z)) \leq \alpha\, d(T^{n+1}(x), T(z)) + \beta\, d(T^{n+2}(x), T(z)).$$

Hence

$$\begin{aligned} d(T^{n+1}(x), T(z)) \quad &\leq \quad a \max\Big\{d(T^n(x), z), c\,\Big[d(T^{n+1}(x), z) \\ &\qquad + d(T^n(x), T(z))\Big]\Big\} \\ &\qquad + b \max\Big\{d(T^n(x), T^{n+1}(x)), d(z, T(z))\Big\}, \end{aligned}$$

and

$$\begin{aligned} d(T^{n+2}(x), T(z)) \quad &\leq \quad a \max\Big\{d(T^{n+1}(x), z), \\ &\qquad c\,\Big[d(T^{n+2}(x), z) + d(T^{n+1}(x), T(z))\Big]\Big\} \\ &\qquad + b \max\Big\{d(T^{n+1}(x), T^{n+2}(x)), d(z, T(z))\Big\}. \end{aligned}$$

First note that it holds that

$$\begin{aligned} d(T^n(x), z) \quad &\leq \quad \alpha\, d(T^n(x), T^{n+1}(x)) + \beta\, d(T^n(x), T^{n+2}(x)) \\ &\leq \quad \alpha\, d(x, T(x)) + \beta\, \tfrac{2}{2-a}\, d(x, T(x)), \end{aligned}$$

and that

$$\begin{aligned} d(T^{n+1}(x),z)+d(T^n(x),T(z)) &\leq \beta\, d(T^{n+1}(x),T^{n+2}(x)) \\ &\quad +d(T^n(x),z)+d(z,T(z)) \\ &\leq \beta\, d(x,T(x))+\alpha\, d(T^n(x),T^{n+1}(x)) \\ &\quad +\beta\, d(T^n(x),T^{n+2}(x))+d(z,T(z)), \\ &\leq d(x,T(x))+\beta\, \tfrac{2}{2-a}\, d(x,T(x)) \\ &\quad +d(z,T(z)), \end{aligned}$$

which implies

$$\begin{aligned} d(T^{n+1}(x),T(z)) &\leq a\, \max\Big\{\Big[\alpha+\tfrac{2\beta}{2-a}\Big]\, d(x,T(x)), \\ &\quad c\,\Big[\Big(1+\tfrac{2\beta}{2-a}\Big)\, d(x,T(x))+d(z,T(z))\Big]\Big\} \\ &\quad + b\, \max\Big\{d(T^n(x),T^{n+1}(x)),d(z,T(z))\Big\} \\ &\leq a\, \max\Big\{\Big[\alpha+\tfrac{2\beta}{2-a}\Big]\, d(x,T(x)), \\ &\quad c\,\Big[\Big(1+\tfrac{2\beta}{2-a}\Big)\, d(x,T(x))+d(z,T(z))\Big]\Big\} \\ &\quad +b\, \max\Big\{d(x,T(x)),d(z,T(z))\Big\}. \end{aligned}$$

Analogously,

$$\begin{aligned} d(T^{n+2}(x),T(z)) &\leq a\, \max\Big\{d(T^{n+1}(x),z),c\,\Big[d(T^{n+2}(x),z) \\ &\quad +d(T^{n+1}(x),T(z))\Big]\Big\} \\ &\quad + b\, \max\Big\{d(T^{n+1}(x),T^{n+2}(x)),d(z,T(z))\Big\} \\ &\leq a\, \max\Big\{\beta\, d(T^{n+1}(x),T^{n+2}(x)), \\ &\quad c\,\Big[\alpha\, d(T^{n+2}(x),T^{n+1}(x))+ d(T^{n+1}(x),z) \\ &\quad +d(z,T(z))\Big]\Big\}+b\, \max\Big\{d(x,T(x)),d(z,T(z))\Big\} \\ &\leq a\, \max\Big\{\beta\, d(x,T(x)),c\,\Big[\alpha\, d(x,T(x)) \\ &\quad +\beta\, d(T^{n+1}(x),T^{n+2}(x))+d(z,T(z))\Big]\Big\} \\ &\quad +b\, \max\Big\{d(x,T(x)),d(z,T(z))\Big\} \\ &\leq a\, \max\Big\{\beta\, d(x,T(x)),c\,\Big[d(x,T(x))+d(z,T(z))\Big]\Big\} \\ &\quad +b\, \max\Big\{d(x,T(x)),d(z,T(z))\Big\}. \end{aligned}$$

It follows from

$$d(z,T(z)) \leq \alpha\, d(T^{n+1}(x),T(z))+\beta\, d(T^{n+2}(x),T(z)),$$

that

$$\begin{aligned} d(z,T(z)) \quad \leq \quad & \alpha\, a\, \max\Big\{\Big[\alpha+\tfrac{2\beta}{2-a}\Big]\, d(x,T(x)), \\ & c\,\Big[\Big(1+\tfrac{2\beta}{2-a}\Big)\, d(x,T(x))+d(z,T(z))\Big]\Big\} \\ & +\beta\, a\, \max\Big\{\beta\, d(x,T(x)), c\,\Big[d(x,T(x))+d(z,T(z))\Big]\Big\} \\ & +b\, \max\Big\{d(x,T(x)), d(z,T(z))\Big\}. \end{aligned}$$

Assume that $d(x,T(x)) < d(z,T(z))$. It follows that

$$\begin{aligned} d(z,T(z)) \quad < \quad & \alpha\, a\, \max\Big\{\Big[\alpha+\tfrac{2\beta}{2-a}\Big], c\,\Big[\Big(1+\tfrac{2\beta}{2-a}\Big)+1\Big]\Big\} d(z,T(z) \\ & +\beta\, a\, \max\{\beta, 2\, c\}\, d(z,T(z))+\, b\, d(z,T(z)). \end{aligned}$$

The inequality $2c<\beta$ yields

$$d(z,T(z)) < \Big[\alpha\, a\, \max\Big\{\alpha+\frac{2\beta}{2-a}, c\,\Big(2+\frac{2\beta}{2-a}\Big)\Big\}+\beta^2\, a+b\Big]\, d(z,T(z)).$$

By virtue of Lemma 10.7 one has:

$$K=\alpha\, a\, \max\Big\{\alpha+\frac{2\beta}{2-a}, c\,\Big(2+\frac{2\beta}{2-a}\Big)\Big\}+\,\beta^2\, a+\, b<1,$$

which implies $d(z,T(z)) < K\, d(z,T(z))$ an obvious contradiction. Therefore it must hold that $d(z,T(z)) \leq d(x,T(x))$. Hence

$$d(z,T(z)) \leq \Big[\alpha\, a\, \max\Big\{\alpha+\frac{2\beta}{2-a}, c\,\Big(2+\frac{2\beta}{2-a}\Big)\Big\}+\beta^2\, a+b\Big]\, d(x,T(x)),$$

i.e., $d(T(z),z) \leq K\, d(x,T(x))$. A sequence $\{z_n\}$ in C will be next constructed by induction, such that

(a) $z_0=x$ and z_1 is the point constructed before;

(b) $(z_n,z_{n+1}) \in E(G)$, for any $n \in \mathbb{N}$;

(c) $d(z_{n+1},T(z_{n+1})) \leq K\, d(z_n,T(z_n))$, for any $n \in \mathbb{N}$.

In particular, $d(z_{n+1},T(z_{n+1})) \leq K^n\, d(x,T(x))$, for any $n \in \mathbb{N}$. Since G is transitive, $(z_n,z_m) \in E(G)$ for any $n \leq m$. Using Lemma 10.4 it is immediately concluded that

$$d(z_n,z_m) \leq \frac{2-a}{1-a}\,\Big(d(z_n,T(z_n))+d(z_m,T(z_m))\Big).$$

Since $K<1$, it is clear that $\{z_n\}$ is Cauchy and G-increasing. Hence it is convergent to some point $\omega \in C$, because C is G-complete. Since C satisfies Property (*), it is clear that $(z_n, \omega) \in E(G)$, for any $n \in \mathbb{N}$. In particular, it must hold that $(x, \omega) \in E(G)$. It will be proved next that ω is a fixed point of T. Since $(z_n, \omega) \in E(G)$, for any $n \in \mathbb{N}$, one has

$$\begin{aligned} d(T(z_n), T(\omega)) \quad \leq \quad & a \max\Big\{d(z_n, \omega), c\, \Big[d(z_n, T(\omega)) + d(T(z_n), \omega)\Big]\Big\} \\ & + b \max\Big\{d(z_n, T(z_n)), d(\omega, T(\omega))\Big\}. \end{aligned}$$

On account of the equality $\lim_{n\to+\infty} d(z_n, T(z_n)) = \lim_{n\to+\infty} d(z_n, \omega) = 0$, it follows that $\lim_{n\to+\infty} d(T(z_n), \omega) = 0$, which implies

$$d(\omega, T(\omega)) \leq a \max\Big\{0, c\, \Big[d(\omega, T(\omega)) + 0\Big]\Big\} + b \max\Big\{0, d(\omega, T(\omega))\Big\},$$

i.e. $d(\omega, T(\omega)) \leq a\, c\, d(\omega, T(\omega)) + b\, d(\omega, T(\omega))$. Since $ac + b < a + b = 1$, it follows that $d(\omega, T(\omega)) = 0$, i.e., that $T(\omega) = \omega$. Finally, let Ω be another fixed point of T such that $(x, \Omega) \in E(G)$. Since T is G-monotone, one must have $(T^n(x), \Omega) \in E(G)$. It follows from the convexity of the G-intervals that $(z_n, \Omega) \in E(G)$ for any $n \in \mathbb{N}$. The application of Lemma 10.4 yields

$$d(z_n, \Omega) \leq \frac{2-a}{1-a}\, \Big(d(z_n, T(z_n)) + d(\Omega, T(\Omega))\Big) = \frac{2-a}{1-a}\, d(z_n, T(z_n)),$$

for any $n \in \mathbb{N}$. Letting $n \to +\infty$, it is readily seen that $\{z_n\}$ converges to Ω. The uniqueness of the limit implies that $\omega = \Omega$. □

Remark 10.7. Assuming $a + b < 1$ and $c \leq \frac{1}{2}$ it follows that the map T is a quasi-contraction mapping [22]. In this case, Theorem 10.10 is similar to the main fixed point result found in [1, 11], without any convexity assumption on the weighted graph.

10.3.3.4 Monotone quasi-contraction mappings on modular weighted graphs

This Subsection deals with the existence of fixed points of G-monotone contraction mappings on modular function spaces. This result is the modular version of Jachymski's fixed point results for mappings defined on a metric space endowed with a weighted graph.

Definition 10.19. [2] Let (X, ω) be a modular metric space, G be a reflexive weighted graph defined on X and C be a nonempty subset of X. The mapping $T : C \to C$ is said to be a G-monotone ω-quasi-contraction if T is

G-monotone and there exists $k < 1$ such that for any $x, y \in C$, $(x, y) \in E(G)$, we have

$$\omega_1(T(x), T(y)) \leq k \max\Big(\omega_1(x,y); \omega_1(x, T(x)); \omega_1(y, T(y)); \omega_1(x, T(y)); \omega_1(y, T(x))\Big).$$

$E(G)$ is said to have Property (***) if

> (**Property** ***) for any G-increasing sequence $\{x_n\}$ in X such that x_n ω-converges to x, there is a subsequence $(x_{k_n})_{n\geq 1}$ with $(x_{k_n}, x) \in E(G)$, for $n \geq 1$.

Note that if G is a reflexive transitive weighted graph defined on X, then the Property (***) implies the following:

> for any G-increasing sequence $\{x_n\}$ in X such that x_n ω-converges to x, then $(x_n, x) \in E(G)$, for every $n \geq 1$.

Throughout this Subsection, (X, ω) will be assumed to be a modular metric space, G will stand for a reflexive transitive weighted graph defined on X and $E(G)$ will be assumed to have property (***).

An analogue to Theorem 10.5 in modular metric spaces will be proved. For any $x \in C$, define the orbit $\mathscr{O}(x) = \{x, T(x), T^2(x), \cdots\}$, and its diameter by

$$\delta_\omega(x) = \sup\{\omega_1(T^n(x), T^m(x)) : n, m \in \mathbb{N}\} .$$

Throughout the remainder of this discussion it is assumed that ω is regular and satisfies the Fatou property. The following technical Lemma is the crux of the proof of the main result of this Subsection. It is the modular version of Lemma 10.3 and its proof will be omitted.

Lemma 10.8. *[2] Let (X, ω) and G be as above. Let C be a nonempty subset of X and $T : C \to C$ be a G-monotone ω-quasi-contraction mapping. Let $x \in C$ be such that $(x, T(x)) \in E(G)$ and that $\delta_\omega(x) < \infty$. Then for any $n \geq 1$, it holds that*

$$\delta_\omega(T^n(x)) \leq k^n \delta_\omega(x) ,$$

where $k < 1$ is the constant associated with the definition of G-monotone ω-quasi-contraction. Moreover we have

$$\omega_1(T^n(x), T^{n+m}(x)) \leq k^n \ \delta_\omega(x)$$

for any $n, m \in \mathbb{N}$.

The following theorem is the main result of this Subsection.

Theorem 10.11. *[2] Let (X,ω) and G be as above. Let C be a ω-complete, nonempty subset of X. Let $T : C \to C$ be a G-monotone, ω-quasi-contraction mapping. Let $x \in C$ be such that $(x, T(x)) \in E(G)$ and assume that $\delta_\omega(x) < \infty$. Then*

(a) $\{T^n(x)\}$ ω-converges to $z \in C$, which is a fixed point of T and $(x,z) \in E(G)$, provided $\omega_1(z,T(z)) < \infty$ and $\omega_1(x,T(z)) < \infty$. Moreover,

$$\omega_1(T^n(x),z) \leq k^n\ \delta_\omega(x),\ \ n \geq 1.$$

(b) If w is a fixed point of T such that $(x,w) \in E(G)$ and that $\omega_1(T^n(x),w) < \infty$, for any $n \geq 1$, then $z = w$.

Proof. For the proof of (a), observe that Lemma 10.8 implies that $\{T^n(x)\}$ is ω-Cauchy. Since C is ω-complete, there exists $z \in C$ such that $\{T^n(x)\}$ ω-converges to z. It follows from the inequality

$$\omega_1(T^n(x),T^{n+m}(x)) \leq k^n\ \delta_\omega(x),$$

valid for any $n,m \in \mathbb{N}$ and from the Fatou property (once we let $m \to \infty$) that

$$\omega_1(T^n(x),z) \leq k^n\ \delta_\omega(x)\ ,\ \ n \geq 1.$$

Since T is G-monotone, one readily concludes that $(T^n(x),T^{n+1}(x)) \in E(G)$, for any $n \geq 1$. On account of property (***), one has $(T^n(x),z) \in E(G)$, for any $n \geq \mathbb{N}$. In particular, $(x,z) \in E(G)$. Next, assume that $\omega_1(z,T(z)) < \infty$ and that $\omega_1(x,T(z)) < \infty$. It will follow that z is a fixed point of T. Indeed, it follows by induction that $\omega_1(T^n(x),T(z)) < \infty$ and that, for any $n \geq 1$,

$$(\diamondsuit)\quad \omega_1(T^n(x),T(z)) \leq k\ \max\Big(\omega_1(T^{n-1}(x),z);\omega_1(T^{n-1}(x),T^n(x)); \omega_1(T(z),z);\omega_1(T^{n-1}(x),T(z)); \omega_1(T^n(x),z)\Big).$$

Consider $r(y) = \limsup\limits_{n\to+\infty} \omega_1(T^n(x),y)$, for $y \in C$. It is clear from $(\diamondsuit)$ that

$$\begin{aligned}\omega_1(T^n(x),T(z)) &\leq k\ \max\Big(k^{n-1}\ \delta_\omega(x);\omega_1(T(z),z);\omega_1(T^{n-1}(x),T(z)); k^n\ \delta_\omega(x)\Big)\\ &= k\ \max\Big(k^{n-1}\ \delta_\omega(x);\omega_1(T(z),z);\omega_1(T^{n-1}(x),T(z))\Big)\\ &\leq k^n\ \delta_\omega(x) + k\ \omega_1(T(z),z) + k\ \omega_1(T^{n-1}(x),T(z))\\ &\leq \delta_\omega(x) + \omega_1(T(z),z) + k\ \omega_1(T^{n-1}(x),T(z)),\end{aligned}$$

for any $n \geq 1$. One gets by induction:

$$\omega_1(T^n(x), T(z)) \quad \leq \quad \frac{1}{1-k}\Big(\delta_\omega(x) + \omega_1(T(z), z)\Big) + k^n \; \omega_1(x, T(z)),$$

for any $n \geq 1$, which implies

$$r(T(z)) \leq \frac{1}{1-k}\Big(\delta_\omega(x) + \omega_1(T(z), z)\Big) < +\infty.$$

So letting $n \to +\infty$ in the inequality

$$\omega_1(T^n(x), T(z)) \leq k \; \max\Big(k^{n-1} \; \delta_\omega(x); \omega_1(T(z), z); \omega_1(T^{n-1}(x), T(z))\Big),$$

one obtains

$$r(T(z)) \leq k \; \max\Big(\omega_1(z, T(z)), r(T(z))\Big).$$

Since ω satisfies the Fatou property, it follows that $\omega_1(z, T(z)) \leq r(T(z))$. This yields,

$$r(T(z)) \leq k \; \max\Big(\omega_1(z, T(z)), r(T(z))\Big) = k \; r(T(z)).$$

Since $k < 1$, it follows $r(T(z)) = 0$, from which it is clear that $\omega_1(z, T(z)) = 0$. Since ω is regular, one cocludes that $T(z) = z$.

Next the proof of (b) is tackled. Let $w \in C$ be a fixed point of T such that $(x, w) \in E(G)$ and that $\omega_1(T^n(x), w) < \infty$, for any $n \geq 1$. Using induction again it can be seen that

$$\omega_1(T^n(x), w) \leq k \; \max\Big(\omega_1(T^{n-1}(x), w); \omega_1(T^{n-1}(x), T^n(x)); \omega_1(T^n(x), w)\Big),$$

for any $n \geq 2$. Note that if, for some $n \geq 1$ it holds that

$$\max\Big(\omega_1(T^{n-1}(x), w); \omega_1(T^{n-1}(x), T^n(x)); \omega_1(T^n(x), w)\Big) = \omega_1(T^n(x), w),$$

then $\omega_1(T^n(x), w) \leq k \;\; \omega_1(T^n(x), w)$. Since $k < 1$ it is clear that $\omega_1(T^n(x), w) = 0$. So $T^n(x) = w$ which yields $T^{n+m}(x) = w$, for any $m \geq 0$, since w is a fixed point of T. This clearly forces $z = w$. Assume otherwise that

$$\max\Big(\omega_1(T^{n-1}(x), w); \omega_1(T^{n-1}(x), T^n(x)); \omega_1(T^n(x), w)\Big) \neq \omega_1(T^n(x), w),$$

for any $n \geq 2$. In this case, one has:

$$\omega_1(T^n(x), w) \leq k \; \max\Big(\omega_1(T^{n-1}(x), w); k^{n-1} \; \delta_\omega(x)\Big),$$

for any $n \geq 2$. Hence

$$\omega_1(T^n(x),w) \leq k\ \omega_1(T^{n-1}(x),w) + k^n\ \delta_\omega(x) \leq k\ \omega_1(T^{n-1}(x),w) + \delta_\omega(x),$$

which implies by induction

$$\omega_1(T^n(x),w) \leq k^n\ \omega_1(x,w) + \frac{1}{1-k}\ \delta_\omega(x),$$

for any $n \geq 1$. In particular, $\limsup_{n\to\infty} \omega_1(T^n(x),w) < +\infty$. The inequality

$$\omega_1(T^n(x),w) \leq k\ \max\Big(\omega_1(T^{n-1}(x),w); k^{n-1}\ \delta_\omega(x)\Big),$$

for any $n \geq 2$, yields

$$\limsup_{n\to\infty} \omega_1(T^n(x),w) \leq k\ \limsup_{n\to\infty} \omega_1(T^n(x),w).$$

On the other hand, $k < 1$, which yields $\limsup_{n\to\infty} \omega_1(T^n(x),w) = 0$, i.e., $\{T^n(x)\}$ converges to w. The uniqueness of the limit implies that $z = w$. Indeed,

$$\omega_2(z,w) \leq \omega_1(T^n(x),z) + \omega_1(T^n(x),w),\ \ n \geq 1.$$

Letting $n \to +\infty$ it follows that $\omega_2(z,w) = 0$. Since ω is regular, one must necessarily have $z = w$. □

Note that under the assumptions of Theorem 10.11, if w is another fixed point of T such that $(w,z) \in E(G)$ and that $\omega_1(z,w) < \infty$ it follows that

$$\omega_1(z,w) = \omega_1(T(z),T(w)) \leq k\ \omega_1(z,w),$$

which implies $z = w$, since $k < 1$.

10.3.3.5 Monotone Reich contraction mappings on weighted graphs

Following the Banach contraction principle, Nadler [48] gave the definition of multivalued contractions and established a multivalued contraction version of the classical Banach's fixed point theorem. Subsequently many mathematcians generalized Nadler's fixed point theorem in different ways. In this regard, we mention the work of Reich, [52] where he posed a (still open) problem on the existence of a fixed point of certain class of multivalued mappings (see Problem 10.1). Mizoguchi and Takahashi [47] gave partial answers to Reich' s problem. Following the publication of Ran and Reurings' fixed point theorem [50], seen as the Banach contraction principle in metric spaces endowed with a partial order, Sultana and Vetrivel [58] tried to extend the main results of [47] to metric spaces endowed with a graph. In particular, they used their ideas to discuss the iterate of the Bernstein operator. Moreover, they gave an example of a nonlinear version of the Bernstein operator and establish the Kelisky and Rivlin's theorem [35] for such operator.

In this Subsection, the definition of the Reich multivalued mappings given by the authors in [58] will be revisited and a fixed point theorem for these mappings will be proved. In addition, vector valued Bernstein operators will be introduced and a more general version of the Kelisky and Rivlin's theorem will be discussed. In particular, the conclusion of [58] regarding the Bernstein operator will be improved.

The difficulty encountered when dealing with multivalued mappings defined on a partially ordered set $(X,\preceq)$ is the problem of comparing two subsets with respect to the order. In fact, there are mainly three well-known pre-orders (reflexive, transitive but not necessarily antisymmetric), namely the *Smyth ordering*, the *Hoare ordering* and the *Egli-Milner ordering* [37, 31, 57], which have been proposed in the context of nondeterministic programming languages, for example.

For any nonempty subsets A and B of X:

1. $A \preceq_S B$, if and only if for any $b \in B$, there exists $a \in A$ such that $a \preceq b$ (Smyth ordering);

2. $A \preceq_H B$, if and only if for any $a \in A$, there exists $b \in B$ such that $a \preceq b$ (Hoare ordering);

3. $A \preceq_{EM} B$, if and only if $A \preceq_S B$ and $A \preceq_H B$ (Egli-Milner ordering).

Clearly, the Hoare order is equivalent to the Smyth order in the dual underlying lattice. Similarly, we follow Jachymski's extension [33] of the fixed point theorem of Ran and Reurings [50] to a metric space endowed with a graph instead of a partial order. Recall that a directed graph G consists of two sets: $V(G)$ a nonempty set of elements called vertices, and $E(G)$ a possibly empty set of elements in $V(G)\times V(G)$ called edges. If $E(G)$ contains all the loops (u,u), then G is reflexive. Let X be a set endowed with a reflexive digraph G such that $V(G)=X$. Given any nonempty subsets A and B of X, the following notation will be used:

1. $(A,B)_S \in E(G)$, if and only if for any $b \in B$, there exists $a \in A$ such that $(a,b) \in E(G)$;

2. $(A,B)_H \in E(G)$, if and only if for any $a \in A$, there exists $b \in B$ such that $(a,b) \in E(G)$;

3. $(A,B)_{EM} \in E(G)$, if and only if $(A,B)_S \in E(G)$ and $(A,B)_H \in E(G)$.

Throughout this Subsection, only the Hoare relationship will be used and the subscript H will be omitted. The following technical result is useful in the sequel.

Lemma 10.1, allows for a simpler formulation of the notion of multivalued contractions, which avoids the use of Hausdorff-Pompeiu distance. More precisely, let (X,d) be a metric space. The mapping $T : X \to \mathscr{C}\mathscr{B}(X)$ is a contraction mapping if there exists $\alpha \in [0,1)$ such that for any $x,y \in X$ and $a \in T(x)$, there exists $b \in T(y)$ such that

$$d(a,b) \leq \alpha\, d(x,y).$$

Clearly this definiton does not use the boundedness assumption of the considered subsets of X. Instead, the class $\mathscr{C}(X)$ of all nonempty closed subsets of X will be considered.

In their attempt to extend the fixed point theorem of Mizoguchi-Takahashi for Reich multivalued contraction mappings to the setting of metric spaces endowed with a graph, Sultana and Vetrivel [58] introduced the concept of Reich G-contractions, namely:

Definition 10.20. [58] Let (X,d,G) be a metric space endowed with a reflexive directed graph G with no parallel edges. The multivalued map $T : X \to \mathscr{C}\mathscr{B}(X)$ is called a Reich G-contraction if for any different $x,y \in X$ with $(x,y) \in E(G)$, it holds that

(a) $H(T(x),T(y)) \leq k(d(x,y))\, d(x,y)$,

(b) if $(u,v) \in T(x) \times T(y)$ is such that $d(u,v) \leq d(x,y)$, then $(u,v) \in E(G)$, for some $k : (0,+\infty) \to [0,1)$ which satisfies $\limsup_{s \to t+} k(s) < 1$, for any $t \in [0,+\infty)$.

This definition is not appropriate because of condition (b). The following example clarifies the reason behind this claim.

Example 10.4. Consider the space $\mathbb{R}^2$ endowed with the Euclidean distance d and the graph G obtained by the pointwise ordering of $\mathbb{R}^2$ defined by

$$(x,y) = ((x_1,x_2),(y_1,y_2)) \in E(G) \text{ iff } x_1 \leq y_1 \;\&\; x_2 \leq y_2.$$

Let A be the unit ball of $\mathbb{R}^2$, i.e., $A = \{(x_1,x_2) \in \mathbb{R}^2;\ d^2(x,0) = x_1^2 + x_2^2 \leq 1\}$. Consider the multivalued map $T : \mathbb{R}^2 \to \mathscr{C}\mathscr{B}(\mathbb{R}^2)$ defined by $T(x) = A$. Then $H(T(x),T(y) = 0$ for any $x,y \in \mathbb{R}^2$. Since T is a constant multivalued mapping, it is a contraction according to Nadler's definition. Therefore, T must be a Reich G-contraction. In this case, condition (a) is obviously satisfied but condition (b) fails. Indeed, set $x = (2,0)$ and $y = (2,2)$. Then $x \neq y$ and $(x,y) \in E(G)$. Since $d(x,y) = 2$, (b) will hold if and only if for any $u,v \in A$

such that $d(u,v) \leq 2$, it must hold $(u,v) \in E(G)$. This is not the case, as can be verified by considering

$$u = (1,0) \text{ and } v = (0,1),$$

then $u,v \in A$, $d(u,v) = \sqrt{2}$, $(u,v) \notin E(G)$ and $(v,u) \notin E(G)$.

Before the correct definition of Reich multivalued G-contractions is given, the following remark is in order:

Remark 10.8. Let (X,d) be a metric space. Let $T : X \to \mathscr{CB}(\mathscr{X})$. Assume there exists $\alpha : (0,+\infty) \to [0,1)$ with $\limsup_{s \to t+} \alpha(s) < 1$, for any $t \in [0,+\infty)$, such that

$$H(T(x),T(y)) \leq \alpha(d(x,y))\, d(x,y),$$

for any different $x,y \in X$. Using Lemma 10.1, it can be easily shown that, for any different $x,y \in X$ and $a \in T(x)$, there exists $b \in T(y)$ such that

$$d(a,b) \leq \beta(d(x,y))\, d(x,y),$$

where $\beta = \frac{1}{2}(1+\alpha)$ and satisfies $\limsup_{s \to t+} \beta(s) < 1$, for any $t \in [0,+\infty)$.

The following Definition is more appropriate than Definition 10.20.

Definition 10.21. [10] Let (X,d,G) be a metric space endowed with a reflexive directed graph G with no parallel edges. The multivalued map $T : X \to \mathscr{C}(X)$ is called a Reich G-contraction if there exists $k : (0,+\infty) \to [0,1)$ with $\limsup_{s \to t+} k(s) < 1$, for any $t \in [0,+\infty)$, such that for any different $x,y \in X$ with $(x,y) \in E(G)$ and any $a \in T(x)$, there exists $b \in T(y)$ for which $(a,b) \in E(G)$ and

$$d(a,b) \leq k(d(x,y))\, d(x,y).$$

A point $x \in X$ is a fixed point of T if $x \in T(x)$.

In [52], Reich raised the following question:

Problem 10.1. Let (X,d) be a complete metric space. Consider a multivalued map $T : X \to \mathscr{CB}(\mathscr{X})$ and assume that T satisfies the following condition: there exists $k : (0,+\infty) \to [0,1)$ with $\limsup_{s \to t+} k(s) < 1$, for any $t \in (0,+\infty)$, such that for any different $x,y \in X$,

$$H(T(x),T(y)) \leq k(d(x,y))\, d(x,y).$$

Does T have a fixed point?

In [51], Reich proved that such mappings have a fixed point provided they have compact values. Clearly, if $k(t)$ is a constant function, Nadler answers Reich's question in the affirmative. In [47], Mizoguchi and Takahashi gave a positive answer when the function $k(t)$ is defined on $[0,+\infty)$. It is still unclear whether Reich's problem can be answered in the affirmative. This Section is devoted to a discussion of the graph version of the fixed point theorem of Mizoguchi and Takahashi.

The first order of business is the statement and a simpler proof of the original theorem of [47], in the absence of boundedness.

Theorem 10.12. *[10] Let (X,d) be a complete metric space. Then any Reich-contraction mapping $T : X \to \mathscr{C}(X)$, has a fixed point.*

Proof. Since $T : X \to \mathscr{C}(X)$ is a Reich-contraction mapping, there exists $k : (0,+\infty) \to [0,1)$ with $\limsup_{s\to t+} k(s) < 1$, for any $t \in [0,+\infty)$, such that for any $x,y \in X$ and $a \in T(x)$, there exists $b \in T(y)$ for which

$$d(a,b) \leq k(d(x,y))\; d(x,y).$$

Fix $y_0 \in X$. If y_0 is a fixed point of T, then there is nothing to prove. Otherwise, choose $y_1 \in T(y_0)$ different from y_0. Using the contractive assumption of T, it can be seen that there exists $y_2 \in T(y_1)$ such that

$$d(y_1,y_2) \leq k(d(y_0,y_1))\; d(y_0,y_1).$$

By induction, construct a sequence $\{y_n\}$ in X such that $y_{n+1} \in T(y_n)$ and $y_n \neq y_{n+1}$ with

$$d(y_n,y_{n+1}) \leq k(d(y_{n-1},y_n))\; d(y_{n-1},y_n),$$

for any $n \geq 1$. Since $k(t) < 1$, for any $t \in [0,+\inf)$, it is clearly seen that $\{d(y_n,y_{n+1})\}$ is a decreasing sequence of positive numbers. Let

$$t_0 = \lim_{n\to+\infty} d(y_n,y_{n+1}) = \inf_{n\in\mathbb{N}} d(y_n,y_{n+1}).$$

Since $\limsup_{s\to t_0+} k(s) < 1$, there exist $\alpha < 1$ and $n_0 \geq 1$ such that $k(d(y_n,y_{n+1})) \leq \alpha$, for any $n \geq n_0$. It is then clear that

$$d(y_n,y_{n+1}) \leq \prod_{k=n_0}^{k=n} k(d(y_k,y_{k+1}))\; d(y_{n_0},y_{n_0+1}) \leq \alpha^{n-n_0}\; d(y_{n_0},y_{n_0+1}),$$

for any $n \geq n_0$. This implies that $\sum d(y_n,y_{n+1})$ is convergent. Hence $\{y_n\}$ is a Cauchy sequence. Since X is complete, $\{y_n\}$ converges to some point

$x \in X$. To see that x is a fixed point of T, observe that the contractivity of T implies the existence of $z_n \in T(x)$ such that

$$d(y_{n+1}, z_n) \leq k(d(y_n, x))\, d(y_n, x) < d(y_n, x),$$

for any $n \in \mathbb{N}$. This will force $\{z_n\}$ to also converge to x. Since $T(x)$ is closed, it is clear that $x \in T(x)$, i.e., x is a fixed point of T, as claimed. $\square$

Next, the extension of Theorem 10.12 to metric spaces endowed with a graph is analyzed.

Theorem 10.13. *[10] Let (X,d) be a complete metric space. Let G be a reflexive graph with no parallel edges, such that $E(G) = X$. Assume that (X,d,G) satisfies Property (*). Let $T : X \to \mathscr{C}(X)$ be a Reich G-contraction. Then there exists $k : (0,+\infty) \to [0,1)$ which satisfies $\limsup_{s \to t+} k(s) < 1$, for any $t \in [0,+\infty)$, such that for any different $x,y \in X$ with $(x,y) \in E(G)$ and any $a \in T(x)$, there exists $b \in T(y)$ such that $(a,b) \in E(G)$ and*

$$d(a,b) \leq k(d(x,y))\, d(x,y).$$

Set $X_T = \{x \in X;$ there exists $y \in T(x)$ such that $(x,y) \in E(G)\}$. If $X_T \neq \emptyset$, then T has a fixed point.

Proof. Assume $X_T \neq \emptyset$. Let $y_0 \in X_T$. Then there exists $y_1 \in T(y_0)$ such that $(y_0, y_1) \in E(G)$. If $y_1 = y_0$, then y_0 is a fixed point of T. Assume $y_0 \neq y_1$. Then, there exists $y_2 \in T(y_1)$ such that

$$d(y_1, y_2) \leq k(d(y_0, y_1))\, d(y_0, y_1).$$

Use induction to construct a sequence $\{y_n\}$ such that $y_n \neq y_{n+1}$, $y_{n+1} \in T(y_n)$, $(y_n, y_{n+1}) \in E(G)$ and

$$d(y_n, y_{n+1}) \leq k(d(y_{n-1}, y_n))\, d(y_{n-1}, y_n),$$

for any $n \geq 1$. As in the proof of Theorem 10.12, it can be shown that $\{y_n\}$ converges to some point $x \in X$. To see that x is a fixed point of T, observe that (X,d,G) satisfies Property (*), there exists a subsequence $\{y_{\varphi(n)}\}$ of $\{y_n\}$ such that $(y_{\varphi(n)}, x) \in E(G)$, for any $n \in \mathbb{N}$. Using the contractivity of T, it is clear that there exists $z_n \in T(x)$ such that

$$d(y_{\varphi(n)+1}, z_n) \leq k(d(y_{\varphi(n)}, x))\, d(y_{\varphi(n)}, x) < d(y_{\varphi(n)}, x),$$

for any $n \in \mathbb{N}$. This will force $\{z_n\}$ to also converge to x. Since $T(x)$ is closed, it follows that $x \in T(x)$, i.e., x is a fixed point of T, which is the desired conclusion. $\square$

Remark 10.9. Once Theorem 10.12 and Theorem 10.13 are established, it is easy to extend them to the case of uniformly locally contractive mappings in the sense of Edelstein [25], with or without a graph.

Application: A generalized Bernstein operator

In [35], Kelisky and Rivlin investigated the behavior of the iterates of the Bernstein polynomial of degree $n \geq 1$, defined by

$$B_n(f)(t) = \sum_{k=0}^{k=n} f\left(\frac{k}{n}\right)\binom{n}{k} t^k(1-t)^{n-k},$$

for any $f \in C([0,1])$ and $t \in [0,1]$, where $C([0,1])$ is the space of continuous functions defined on $[0,1]$. In particular, they proved that for any $f \in C([0,1])$, the equality

$$\lim_{j\to+\infty} B_n^j(f)(t) = f(0)(1-t) + f(1)t, \;\; 0 \leq t \leq 1 \tag{KRB}$$

holds.

Their proof relies on matrix-algebra techniques. Rus [56] was the first one to notice the existence of a proof of (KRB) that is metric in nature. In fact, his proof inspired Jachymski [33] to rephrase it using the language of graph theory. In [58], Sultana and Vetrivel modified the Bernstein operator to obtain a nonlinear version, which would not be suitable for the technique used by Kelisky and Rivlin. Indeed, Sultana and Vetrivel introduced the operator:

$$B'_n(f)(t) = \sum_{k=0}^{k=n} \left| f\left(\frac{k}{n}\right)\right| \binom{n}{k} t^k(1-t)^{n-k},$$

for any $f \in C([0,1])$ and they proved that

$$\lim_{j\to+\infty} (B'_n)^j(f)(t) = f(0)(1-t) + f(1)t, \;\; 0 \leq t \leq 1,$$

for any $f \in C([0,1])$, such that $f(0) \geq 0$ and $f(1) \geq 0$. A better conclusion will be presented here. Indeed, it is easy to see that $B'_n(f) = B_n(|f|)$, for any $f \in C([0,1])$, which yields the proof of the following more general result:

Proposition 10.1. *Let $f \in C([0,1])$. Then*

$$\lim_{j\to+\infty} (B'_n)^j(f)(t) = |f(0)|(1-t) + |f(1)|t, \;\; 0 \leq t \leq 1.$$

The classical Bernstein operator will be now extended to the vector-valued case. Let $(X, \|.\|_X)$ be a Banach space. Consider the Banach space $C([0,1],X)$ of all continuous functions defined on $[0,1]$ with values in X. The norm in $C([0,1],X)$ is given by

$$\|f\| = \sup\{\|f(t)\|_X;\; t \in [0,1]\}.$$

Fix $n \geq 1$ and define the generalized Bernstein operator $B_n : C([0,1],X) \to C([0,1],X)$ by

$$B_n(f)(t) = \sum_{k=0}^{k=n} \binom{n}{k} t^k (1-t)^{n-k} \, f\left(\frac{k}{n}\right),$$

for any $t \in [0,1]$. In this case a conclusion similar to Kelisky and Rivlin's result holds:

Theorem 10.14. *[4] For any $f \in C([0,1],X)$, we have*

$$\lim_{j \to +\infty} B_n^j(f)(t) = (1-t)f(0) + tf(1), \;\; 0 \leq t \leq 1.$$

Proof. This conclusion will be proved using the language of graph theory. Since the proof is identical to the one used by Rus [56], we prefer to give this proof instead. First notice that

$$\sum_{k=0}^{k=n} \binom{n}{k} t^k (1-t)^{n-k} = 1, \; and \; \sum_{k=0}^{k=n} \frac{k}{n} \binom{n}{k} t^k (1-t)^{n-k} = t,$$

for any $t \in [0,1]$. Set $g(t) = (1-t)f(0) + tf(1)$, for $t \in [0,1]$. Obviously $g \in C([0,1],X)$. We have $B_n(g) = g$. Since $f(0) = g(0)$ and $f(1) = g(1)$, we have

$$B_n(f)(t) - B_n(g)(t) = \sum_{k=1}^{n-1} \binom{n}{k} t^k (1-t)^{n-k} \left(f\left(\frac{k}{n}\right) - g\left(\frac{k}{n}\right) \right),$$

for any $t \in [0,1]$. Hence

$$\|B_n(f)(t) - B_n(g)(t)\| \leq \sum_{k=1}^{n-1} \binom{n}{k} t^k (1-t)^{n-k} \left\| f\left(\frac{k}{n}\right) - g\left(\frac{k}{n}\right) \right\|,$$

for any $t \in [0,1]$, which implies

$$\|B_n(f)(t) - B_n(g)(t)\| \leq \left(1 - \frac{1}{2^{n-1}}\right) \|f - g\|,$$

for any $t \in [0,1]$. Therefore,

$$\|B_n(f) - g\| = \|B_n(f) - B_n(g)\| \leq \left(1 - \frac{1}{2^{n-1}}\right) \|f - g\|.$$

It follow by induction that

$$\|B_n^j(f) - g\| \leq \left(1 - \frac{1}{2^{n-1}}\right)^j \|f - g\|,$$

for any $j \in \mathbb{N}$. This clearly implies the conclusion of Theorem 10.14. $\square$

Motivated by the example given by Sultana and Vetrivel, the Bernstein operator $B'_n : C([0,1],X) \to C([0,1],X)$ is introduced as

$$B'_n(f)(t) = \sum_{k=0}^{k=n} \binom{n}{k} t^k (1-t)^{n-k}\, T\left(f\left(\frac{k}{n}\right)\right), \qquad (10.2)$$

where $T : X \to X$ is continuous. Since $B'_n(f) = B_n(T \circ f)$, it follows :

Theorem 10.15. *[10] For any $f \in C([0,1],X)$, one has*

$$\lim_{j\to+\infty} (B'_n)^j(f)(t) = (1-t)T\Big(f(0)\Big) + tT\Big(f(1)\Big), \;\; 0 \le t \le 1.$$

Remark 10.10. A careful look at the definition reveals that B_n is actually a convex combination, because $\sum_{i=0}^{i=m} \binom{m}{i} t^i(1-t)^{m-i} = 1$, for any $m \ge 1$. Therefore, conclusions similar as those of Theorems 10.14 and 10.15 can be obtained by taking X to be a hyperbolic metric space, such as $CAT(0)$ spaces. For more on hyperbolic spaces, we refer the interested reader to [18, 29, 38, 39, 44, 53].

Next, the existence of fixed points for multivalued monotone Reich (a,b,c)-contraction on weighted graphs is investigated. Throughout the following discussion it will be assumed that (X,d) is a metric space, that $\mathscr{CB}(\mathscr{X})$ is the class of all nonempty closed and bounded subsets of X and that G is a reflexive digraph defined on X. It is also assumed that the triple (X,G,d) has property (P) and that G-intervals are closed. Recall that a G-interval is any of the subsets $[x,\to) = \{u \in X; (x,u) \in E(G)\}$ and that $(\leftarrow, y] = \{u \in X; (u,y) \in E(G)\}$, for any $x, y \in X$.

Reich in [51] proved that any multivalued Reich (a,b,c)-contraction on a complete space has a fixed point. In this section, the notation of multivalued monotone Reich contraction mappings on graphs is defined and a fixed point theorem for such mappings is proved.

Definition 10.22. [10] Let (X,d) be a metric space. A multivalued mapping $J : X \to \mathscr{C}(X)$ is called Reich (a,b,c)-contraction if there exist nonnegative numbers a,b,c with $a+b+c < 1$ such that for any $u,w \in X$ and any $U \in J(u)$, there exists $W \in J(w)$ for which

$$d(U,W) \le a\, d(u,w) + b\, d(u,U) + c\, d(w,W). \qquad (10.3)$$

Definition 10.23. [10] Let (X,G,d) be as above. A multivalued mapping $J : X \to \mathscr{C}(\mathscr{X})$ is called G-monotone Reich (a,b,c)-contraction if there exist nonnegative numbers a,b,c with $a+b+c < 1$ such that for any $u,w \in X$ with $(u,w) \in E(G)$ and any $U \in J(u)$, there exists $W \in J(w)$ such that $(U,W) \in E(G)$ and

$$d(U,W) \le a\, d(u,w) + b\, d(u,U) + c\, d(w,W). \qquad (10.4)$$

Example 10.5. Let $X = \{0,1,2,3\}$ and $d(x,y) = |x-y|$, $\forall x,y \in X$. Define the multivalued map $J : X \to \mathscr{C}(X)$ by:

$$J(x) = \{0,2,3\} \text{ for } x \in \{0,1\} \text{ and } J(x) = \{1,3\} \text{ for } x \in \{2,3\}.$$

Then J is a G-monotone Reich $(\frac{1}{3},0,\frac{1}{3})$-contraction, where

$$G = \{(0,0),(1,1),(2,2),(3,3),(0,1),(0,2),(2,3)\},$$

but J is not a multivalued Reich $(\frac{1}{3},0,\frac{1}{3})$-contraction since $d(0,1) > \frac{1}{3}\ d(1,2) + 0\ d(1,0) + \frac{1}{3}\ d(2,1)$ and $d(0,3) > \frac{1}{3}\ d(1,2) + 0\ d(1,0) + \frac{1}{3}\ d(2,3)$.

Such an example reinforces the idea that the study of multivalued G-monotone Reich contraction is worthy of consideration. The next Theorem is the main result of this Section.

Theorem 10.16. *[10] Let (X,d) be a complete metric space and G be a reflexive digraph defined on X such that (X,G,d) has Property (P). Let $J : X \to \mathscr{C}(\mathscr{X})$ be a multivalued G-monotone Reich (a,b,c)-contraction mapping. Let $u_0 \in X$ be such that $(u_0,u_1) \in E(G)$, for some $u_1 \in J(u_0)$. Then there exists an orbit $\{u_n\}$ of J at u_0 which converges to $\omega \in X$, a fixed point of J.*

Proof. Since $(u_0,u_1) \in E(G)$ and as J is a G-monotone Reich contraction mapping, there exists $u_2 \in J(u_1)$ such that $(u_1,u_2) \in E(G)$ and in addition

$$d(u_1,u_2) \le a\ d(u_0,u_1) + b\ d(u_0,u_1) + c\ d(u_1,u_2).$$

Thus,

$$d(u_1,u_2) \le \frac{(a+b)}{1-c}\ d(u_0,u_1).$$

Set $\alpha = \frac{(a+b)}{1-c}$. A straightforward inductive procedure yields a sequence $\{u_n\}_{n\in\mathbf{N}}$ such that $(u_n,u_{n+1}) \in E(G)$ with

$$d(u_n,u_{n+1}) \le \alpha^n\ d(u_0,u_1).$$

Clearly, $\{u_n\}_{n\in\mathbf{N}}$ is Cauchy. Since (X,d) is complete, there exists $\omega \in X$ such that $u_n \to \omega$. Since (X,G,d) has Property (P), there is a subsequence (u_{k_n}) such that $(u_{k_n},\omega) \in E(G)$, for every $n \ge 0$. Next, ω will be proven to be a fixed point of J, i.e., we will show that $\omega \in J(\omega)$. Since $u_{k_n+1} \in J(u_{k_n})$ and $(u_{k_n},\omega) \in E(G)$, there exists $\omega_n \in J(\omega)$ such that

$$d(u_{k_n+1},\omega_n) \le a\ d(u_{k_n},\omega) + b\ d(u_{k_n},u_{k_n+1}) + c\ d(\omega,\omega_n),$$

for any $n \geq 1$. Thus,

$$d(\omega, \omega_n) - d(u_{k_n+1}, \omega) \leq a\, d(u_{k_n}, \omega) + b\, d(u_{k_n}, u_{k_n+1}) + c\, d(\omega, \omega_n),$$

for any $n \geq 1$. Therefore,

$$(1-c)d(\omega, \omega_n) \leq a\, d(u_{k_n}, \omega) + b\, d(u_{k_n}, u_{k_n+1}) + d(u_{k_n+1}, \omega),$$

for any $n \geq 1$. Hence

$$(1-c) \limsup_{n \to +\infty} d(\omega, \omega_n) \leq 0,$$

which implies $\lim_{n \to +\infty} d(\omega, \omega_n) = 0$, since $c < 1$. Therefore $\{\omega_n\}$ converges to ω and since $J(\omega)$ is closed it follows that $\omega \in J(\omega)$, i.e., ω is a fixed point of J. □

10.3.4 Nonexpansive monotone mappings

For the rest of this Section, a Banach space $(X, \|.\|)$ is fixed. Let G be a weighted digraph such that $V(G) \subset X$. In this case, the weight of an edge (u,v) is given by $\|u-v\|$, for all $u, v \in V(G)$. The following linear convexity structure will be needed in the sequel:

(CG) If $(x,y) \in E(G)$ and $(u,v) \in E(G)$, then

$$(\alpha\, x + (1-\alpha)\, u, \alpha\, y + (1-\alpha)\, v) \in E(G)$$

for all $x, y, u, v \in C$ and $\alpha \in [0,1]$.

Remark 10.11. It is not difficult to show that a G-nonexpansive mapping may not be continuous. Therefore, in this case, it is quite difficult to expect any nice behavior that will imply the existence of a fixed point for this class of mappings.

The existence of fixed points of G-nonexpansive mappings is next tackled. As such mappings do not behave well on their entire domains, but only on connected vertices, our analysis is built on a constructive iteration method introduced by Mann [45] (see also [32, 42]).

Lemma 10.9. *[5] Let $T : C \to C$ be a G-monotone mapping, where $C \subseteq V(G)$ is nonempty and convex. Fix $\lambda \in (0,1)$ and $x_0 \in C$. Consider the Mann iteration sequence $(x_n)_{n \in \mathbb{N}} \subset C$ defined by*

$$x_{n+1} = (1-\lambda)x_n + \lambda T(x_n), \;\; n \in \mathbb{N}. \tag{MS}$$

(a) *If* $(x_0, T(x_0)) \in E(G)$, *then* $(x_n, x_{n+1}) \in E(G)$ *for any* $n \in \mathbb{N}$.

(b) *If* $(T(x_0), x_0) \in E(G)$, *then* $(x_{n+1}, x_n) \in E(G)$ *for any* $n \in \mathbb{N}$.

Therefore, if $(x_0, T(x_0)) \in E(\widetilde{G})$, *then* (x_n) *is G-monotone.*

Proof. (a). As $(x_0, T(x_0)) \in E(G)$ and $(x_0, x_0) \in E(G)$, it follows from property (CG) that

$$\Big((1-\lambda)x_0 + \lambda x_0, (1-\lambda)x_0 + \lambda T(x_0)\Big) \in E(G),$$

i.e., $(x_0, x_1) \in E(G)$. Now assume that $(x_{n-1}, x_n) \in E(G)$ for $n > 0$. As T is G-monotone, one has $(T(x_{n-1}), T(x_n)) \in E(G)$. Property (CG) again implies

$$\Big((1-\lambda)x_{n-1} + \lambda T(x_{n-1}), (1-\lambda)x_n + \lambda T(x_n)\Big) \in E(G),$$

i.e., $(x_n, x_{n+1}) \in E(G)$. Hence it follows by induction that $(x_n, x_{n+1}) \in E(G)$ for all $n \in \mathbb{N}$. The proof of (b) follows along similar lines. □

The following crucial inequality is essential to show that the sequence (MS) has the main property.

Lemma 10.10. *[1, 27, 28] Let* $T : C \to C$ *be a G-nonexpansive mapping, where* $C \subseteq V(G)$ *is nonempty and convex. Suppose that G is transitive and that there is* $x_0 \in C$ *with* $(x_0, T(x_0)) \in E(\widetilde{G})$. *Consider the sequence* $(x_n)_{n\in\mathbb{N}}$ *defined by (MS). Then:*

$$(GK) \quad (1+n\lambda)\|T(x_i) - x_i\| \le \|T(x_{i+n}) - x_i\| + (1-\lambda)^{-n}\Big(\|T(x_i) - x_i\| - \|T(x_{i+n}) - x_{i+n}\|\Big),$$

for any $i, n \in \mathbb{N}$.

Theorem 10.17. *[5] Let* $T : C \to C$ *be a G-nonexpansive mapping, where* $C \subseteq V(G)$ *is nonempty, convex and weakly G-bounded. Suppose that G is transitive and that there is* $x_0 \in C$ *with* $(x_0, T(x_0)) \in E(\widetilde{G})$. *Consider the sequence* $(x_n)_{n\in\mathbb{N}}$, *defined by (MS). Then,* $\lim_{n\to\infty} \|x_n - T(x_n)\| = 0$.

Proof. It follows from Lemma 10.10 that for any $i, n \in \mathbb{N}$, one has

$$(1+n\lambda)\|T(x_i) - x_i\| \le \|T(x_{i+n}) - x_i\| + (1-\lambda)^{-n}(\|T(x_i) - x_i\| - \|T(x_{i+n}) - x_{i+n}\|). \tag{10.5}$$

Notice that $(\|x_n - T(x_n)\|)_{n\in\mathbb{N}}$ is decreasing. Indeed,

$$x_{n+1} - x_n = \lambda\ (T(x_n) - x_n),$$

which implies that $\|x_{n+1} - x_n\| = \lambda\ \|x_n - T(x_n)\|$, for any $n \in \mathbb{N}$. In order to see that $(\|x_n - T(x_n)\|)_{n\in\mathbb{N}}$ is decreasing, it is enough to show that $(\|x_{n+1} - x_n\|)_{n\in\mathbb{N}}$ is decreasing. For any $n \in \mathbb{N}$ it is clear that

$$\begin{aligned}
\|x_{n+2} - x_{n+1}\| &= \Big\|(1-\lambda)\ (x_{n+1} - x_n) + \lambda\ (T(x_{n+1}) - T(x_n))\Big\| \\
&\leq (1-\lambda)\ \|x_{n+1} - x_n\| + \lambda\ \|T(x_{n+1}) - T(x_n)\| \\
&\leq (1-\lambda)\ \|x_{n+1} - x_n\| + \lambda\ \|x_{n+1} - x_n\| \\
&= \|x_{n+1} - x_n\|,
\end{aligned}$$

where we used the G-monotonicity of T. Therefore, $(\|x_{n+1} - x_n\|)_{n\in\mathbb{N}}$ is decreasing. Set $\lim\limits_{n\to+\infty} \|x_n - T(x_n)\| = R$. Note that for any $i, n \in \mathbb{N}$, we have

$$\|T(x_{i+n}) - x_i\| \leq \|T(x_{i+n}) - x_{i+n}\| + \|x_{i+n} - x_i\| \leq \|T(x_0) - x_0\| + \delta((x_n)).$$

Lemma 10.9 implies that (x_n) is G-monotone. Since C is weakly G-bounded, (x_n) must be bounded. Now, letting $i \to +\infty$ in (10.5), we have

$$(1 + n\lambda)R \leq \|T(x_0) - x_0\| + \delta((x_n)),$$

for any $n \in \mathbb{N}$. This clearly implies that $R = 0$, i.e.

$$\lim_{n\to+\infty} \|x_n - T(x_n)\| = 0.$$

$\square$

The above results lead to a fixed point result for G-monotone mappings. This is the topic of the next discussion. The Banach spaces under consideration will be $L^p([0,1])$, $1 \leq p < +\infty$ and ℓ^p, $1 \leq p < +\infty$. It will be shown that the new results are an improvement over the results given in [1]. Recall that $L^p([0,1])$ is the set of real valued functions defined on $[0,1]$ with Lebesgue-integrable absolute value, i.e., such that $\int_{[0,1]} |f(x)|^p dx < +\infty$. In $L^p([0,1])$, $p \geq 1$, τ is the almost everywhere convergence and in ℓ^p, $p \geq 1$, τ is the coordinatewise convergence. In the next example, a digraph G will be constructed on $L^1([0,1])$ for which G_τ-compactness and τ-compactness are different.

Example 10.6. [5] Set $I_n = \left(\frac{1}{n+1}, \frac{1}{n}\right)$, for $n \geq 1$. Define the digraph G on $L^1([0,1])$ by $(f,g) \in E(G)$ if and only if there exists $n_0 \geq 1$ such that $0 \leq f(t) \leq g(t) \leq n_0$, for almost any $t \in I_{n_0}$ and that $f(t) = g(t) = 0$, for almost any $t \notin I_{n_0}$. We claim that $L^1([0,1])$ is G_τ-compact. Let (f_n) be a G-increasing sequence. Since $(f_1, f_2) \in E(G)$, there exists $n_0 \geq 1$ such that

$$\begin{cases} 0 \leq f_1(t) \leq f_2(t) \leq n_0, & \text{for almost any } t \in I_{n_0}; \\ f_1(t) = f_2(t) = 0, & \text{for almost any } t \notin I_{n_0}. \end{cases}$$

Now, $(f_2, f_3) \in E(G)$ implies that there exists $n_1 \geq 1$ such that

$$\begin{cases} 0 \leq f_2(t) \leq f_3(t) \leq n_1, & \text{for almost any } t \in I_{n_1}; \\ f_2(t) = f_3(t) = 0, & \text{for almost any } t \notin I_{n_1}. \end{cases}$$

If $n_0 \neq n_1$, then $f_2 = 0$ and hence $f_1 = 0$. Either all $f_n = 0$ or there exists $f_{m_0} \neq 0$, for some $m_0 \geq 1$. Assume without loss of generality that $m_0 = 1$, i.e., $f_1 \neq 0$. In this case, there exists $n_0 \geq 1$ such that

$$\begin{cases} 0 \leq f_n(t) \leq f_{n+1}(t) \leq n_0, & \text{for almost any } t \in I_{n_0}, \\ f_n(t) = 0, & \text{for almost any } t \notin I_{n_0}, \end{cases}$$

for all $n \geq 1$. Hence the sequence (f_n) is bounded and converges almost everywhere, i.e., τ-converges, to an element $f \in L^1[0,1]$ such that

$$\begin{cases} 0 \leq f_n(t) \leq f(t) \leq n_0, & \text{for almost any } t \in I_{n_0}, \\ f_n(t) = f(t) = 0, & \text{for almost any } t \notin I_{n_0}, \end{cases}$$

i.e., $(f_n, f) \in E(G)$, for all $n \geq 1$. Next, let (g_n) be a G-decreasing sequence. Since $(g_2, g_1) \in E(G)$, there exists $n_0 \geq 1$ such that

$$\begin{cases} 0 \leq g_2(t) \leq g_1(t) \leq n_0, & \text{for almost any } t \in I_{n_0}, \\ g_1(t) = g_2(t) = 0, & \text{for almost any } t \notin I_{n_0}. \end{cases}$$

Obviously this implies that $g_n(t) = 0$ for almost any $t \notin I_{n_0}$ and $n \geq 1$. Therefore, either $g_n = 0$ for any $n \geq 2$, or there exists $n_1 \geq 2$ such that $g_{n_1} \neq 0$. As before,it can be shown that

$$0 \leq g_{n+1}(t) \leq g_n(t) \leq n_0, \text{ for almost any } t \in I_{n_0},$$

for any $n \geq n_1$. Hence the sequence (g_n) is bounded and converges almost everywhere, i.e., τ-converges, to an element $g \in L^1[0,1]$ for which there exists $n_1 \geq 1$ such that

$$\begin{cases} 0 \leq g(t) \leq g_n(t) \leq n_0, & \text{for almost any } t \in I_{n_0}, \\ g_n(t) = g(t) = 0, & \text{for almost any } t \notin I_{n_0} \end{cases}$$

In other words, $(g, g_n) \in E(G)$, for all $n \geq n_1$. Therefore, $L^1[0,1]$ is G_τ-compact. We remark that it is obvious that $L^1[0,1]$ is not τ-compact.

In the general theory of nonexpansive mappings, the iteration sequence defined by (MS) is recognized as an approximate fixed point sequence of T (see e.g. [32]). It is quite remarkable that this result holds for G-nonexpansive mappings as well. To prove the next result, the following crucial Lemma is indispensable.

Lemma 10.11. *[17] If $(f_n)_{n\geq 1}$ is a sequence of L^p-uniformly bounded functions on a measure space, and $f_n \overset{a.e.}{\to} f$, then*

$$\liminf_{n\to\infty} \|f_n\|^p = \liminf_{n\to\infty} \|f_n - f\|^p + \|f\|^p,$$

for all $p \in (0,\infty)$.

The main result of this subsection is the following Theorem.

Theorem 10.18. *[5] Let G be a weighted, reflexive and transitive digraph such that $V(G) \subseteq L^p([0,1])$, $p \geq 1$. Let $C \subset V(G)$ be nonempty, convex and weakly G-bounded. Let $T : C \to C$ be a G-nonexpansive mapping. Assume there exists $f_0 \in C$ with $(f_0, T(f_0)) \in E(\widetilde{G})$. Consider the sequence $(f_n)_{n\mathbb{N}}$, defined by f_0 and (MS). Then any a.e.-cluster point f of $(f_n)_{n\geq 1}$ is a fixed point of T, i.e., $T(f) = f$.*

Proof. Without loss of generality, it may assumed that $(f_0, T(f_0)) \in E(G)$. Let f be an a.e.-cluster point of (f_n). Since C is G_τ-compact and as for any $n \in \mathbb{N}$, $(f_n, f_{n+1}) \in E(G)$, there is a subsequence $(f_{\phi(n)})$ of (f_n) with $f_{\phi(n)} \overset{a.e.}{\to} f$ and $(f_n, f) \in E(G)$, for any $n \in \mathbb{N}$. Since C is weakly G-bounded, Lemma 10.11 implies

$$\liminf_{n\to\infty} \|f_{\phi(n)} - T(f)\|^p = \liminf_{n\to\infty} \|f_{\phi(n)} - f\|^p + \|f - T(f)\|^p.$$

It follows from $\lim\limits_{n\to+\infty} \|f_{\phi(n)} - T(f_{\phi(n)})\| = 0$ that

$$\liminf_{n\to\infty} \|f_{\phi(n)} - T(f)\|^p = \liminf_{n\to\infty} \|T(f_{\phi(n)}) - T(f)\|^p,$$

which implies

$$\liminf_{n\to\infty} \|T(f_{\phi(n)}) - T(f)\|^p = \liminf_{n\to\infty} \|f_{\phi(n)} - f\|^p + \|f - T(f)\|^p.$$

On the other hand, as $(f_{\phi(n)}, f) \in E(G)$, for every $n \in \mathbb{N}$; thus:

$$\begin{aligned}
\liminf_{n\to\infty} \|f_{\phi(n)} - f\|^p + \|f - T(f)\|^p &= \liminf_{n\to\infty} \|T(f_{\phi(n)}) - T(f)\|^p \\
&\leq \liminf_{n\to\infty} \|f_{\phi(n)} - f\|^p,
\end{aligned}$$

which obviously implies $\|f - T(f)\|^p = 0$, or $T(f) = f$. $\square$

Theorem 10.18 implies the following Theorem whose proof will be omitted.

Theorem 10.19. *[5] Let G be a weighted, reflexive and transitive digraph such that $V(G) \subseteq L^p([0,1])$, $p \geq 1$. Let $C \subset V(G)$ be nonempty, convex, weakly G-bounded and G_τ-compact. Let $T : C \to C$ be a G-nonexpansive mapping. Assume there exists $f_0 \in C$ such that $(f_0, T(f_0)) \in E(\widetilde{G})$. Then T has a fixed point.*

Theorem 10.19 is a generalization of the original existence theorem [16, 43] for nonexpansive mappings that are not necessarily monotone. It is at the same time an extension of the main result of [1] and an improvement of the main result of [36]

Since a lemma similar to Lemma 10.11 exists in ℓ^p, $p \geq 1$, for the coordinatewise convergence, we have the following result:

Theorem 10.20. *[5] Let G be a weighted reflexive and transitive digraph such that $V(G) \subseteq l^p([0,1])$, $p \geq 1$. Let $C \subset V(G)$ be nonempty, convex, weakly G-bounded and G_τ-compact. Let $T : C \to C$ be a G-nonexpansive mapping. Assume there exists $f_0 \in C$ with $(f_0, T(f_0)) \in E(\widetilde{G})$. Then T has a fixed point.*

Next, some existence results for nonexpansive, single-valued and multivalued G-monotone mappings defined on hyperbolic metric spaces will be studied. To the best of our knowledge, the following results were never investigated for such mappings.

Theorem 10.21. *[5] Let (X,d) be a complete hyperbolic metric space and suppose that the triple (X,d,G) has property* **(*)**. *Assume G is convex. Let C be a nonempty, closed, convex and bounded subset of X. Let $T : C \to C$ be a G-nonexpansive mapping. Assume $C_T := \{x \in C : \ (x, T(x)) \in E(G)\} \neq \emptyset$. Then*

$$\inf\{d(x,T(x);\ x \in C\} = 0.$$

In particular, there exists an approximate fixed point sequence (x_n)] of T in C i.e., such that

$$\lim_{n\to\infty} d(x_n, T(x_n)) = 0.$$

Proof. Fix $a \in C$. Let $\lambda \in (0,1)$ and define $T_\lambda : C \to C$ by

$$T_\lambda(x) = \lambda a \oplus (1-\lambda)T(x).$$

If $(x,y) \in E(G)$, then necessarily $(T(x),T(y)) \in E(G)$, since T is G-edge preserving. Moreover, since G is convex and $(a,a) \in E(G)$ it is clear that

$$(T_\lambda(x), T_\lambda(y)) = (\lambda a \oplus (1-\lambda)T(x), \lambda a \oplus (1-\lambda)T(y)) \in E(G),$$

i.e., T_λ is G-edge preserving, and

$$\begin{aligned} d(\lambda a \oplus (1-\lambda)T(x), \lambda a \oplus (1-\lambda)T(y)) &\leq (1-\lambda)d(T(x),T(y)) \\ &\leq (1-\lambda)d(x,y) \end{aligned}$$

i.e., $d(T_\lambda(x),T_\lambda(y)) \leq (1-\lambda)d(x,y)$. In other words, T_λ is a G-contraction. It is easy to see that $C_T \subset C_{T_\lambda}$. Hence C_{T_λ} is not empty. Theorem 10.1 implies the existence of a fixed point ω_λ of T_λ in C. Thus,

$$\omega_\lambda = \lambda a \oplus (1-\lambda)T(\omega_\lambda),$$

which yields

$$d(\omega_\lambda, T(\omega_\lambda)) \leq \lambda d(a, T(\omega_\lambda)) \leq \lambda\ \delta(C),$$

where $\delta(C) = \sup\{d(x,y);\ x,y \in C\}$ is the diameter of C. Set $x_n = \omega_{1/n}$, for $n \geq 1$. It is then clear that $d(x_n, T(x_n)) \leq \delta(C)/n$ for $n \geq 1$. In particular, it follows that

$$\inf\{d(x,T(x);\ x \in X\} \leq \lim_{n\to\infty} d(x_n, T(x_n)) = 0.$$

The proof of Theorem 10.21 is therefore complete. □

Theorem 10.22. *[5] Let (X,d) be a complete hyperbolic metric space and suppose that the triple (X,d,G) has property* **(*)**. *Assume G is convex and transitive. Let C be a nonempty, G-compact and convex subset of X. Let $T : C \to C$ be a G-nonexpansive mapping. Assume $C_T := \{x \in C :\ (x,T(x)) \in E(G)\} \neq \emptyset$. Then T has a fixed point.*

Proof. Choose $x_0 \in C_T$. Let (λ_n) be a sequence of numbers in $(0,1)$ such that $\lim\limits_{n\to\infty} \lambda_n = 0$. As in the proof of Theorem 10.21, define the mapping $T_1 : C \to C$ by

$$T_1(x) = \lambda_1 x_0 \oplus (1-\lambda_1)T(x).$$

Since $(x_0, T(x_0)) \in E(G)$, one has $(x_0, T_1(x_0)) \in E(G)$, and since T_1 is G-edge preserving one concludes that $(T_1^n(x_0), T_1^{n+1}(x_0)) \in E(G)$ and that

$$d(T_1^n(x_0), T_1^{n+1}(x_0)) \leq \lambda_1^n\ d(x_0, T_1(x_0)),\ \ for\, n \geq 1.$$

Hence $(T_1^n(x_0))$ is a Cauchy sequence. Since C is G-compact, $(T_1^n(x_0))$ must be convergent. Set $\lim\limits_{n\to\infty} T_1^n(x_0) = x_1$. Property **(**)** implies that $(x_0,x_1) \in E(G)$. A sequence (x_n) can be constructed by induction in such a way that x_{n+1} is a fixed point of $T_{n+1} : C \to C$, defined by

$$T_{n+1}(x) = \lambda_{n+1} x_n \oplus (1-\lambda_{n+1})T(x),$$

obtained as the limit of $(T_{n+1}^k(x_n))_{k\geq 1}$. In particular, $(x_n, x_{n+1}) \in E(G)$, for any $n \geq 1$. Since C is G-compact, there exists a subsequence (x_{k_n}) which converges to $\omega \in C$. Since G is transitive, property **(**)** implies that $(x_{k_n}, \omega) \in E(G)$. Using the G-nonexpansiveness of T, one concludes that

$$d(T(x_{k_n}), T(\omega)) \leq d(x_{k_n}, \omega),\ for\ n \geq 1.$$

Hence $(T(x_{k_n}))$ converges to $T(\omega)$, and since x_{n+1} is a fixed point of T_{n+1}, it follows that $x_{n+1} = \lambda_{n+1} x_n \oplus (1-\lambda_{n+1}) T(x_{n+1})$, which implies

$$d(x_{n+1}, T(x_{n+1})) \ \leq\ \lambda_{n+1}\, d(x_n, T(x_{n+1})) \leq \lambda_{n+1}\ \delta(C),\ for\ n \geq 1,$$

from which it follows $\lim\limits_{n\to\infty} d(x_n, T(x_n)) = 0$. Hence $(T(x_{k_n}))$ converges to ω as well. Therefore it must hold that $T(\omega) = \omega$, i.e., T has a fixed point. $\square$

Next the above results are investigated for mutlivalued mappings. The first claim for these mappings is the analogue to Theorem 10.21.

Theorem 10.23. *[5] Let (X,d) be a complete hyperbolic metric space and suppose that the triple (X,d,G) has property* **(*)**. *Assume G is convex. Let C be a nonempty, closed, convex and bounded subset of X. Set $\mathscr{C}(C)$ to be the set of all nonempty, closed subsets of C. Let $T : C \to \mathscr{C}(C)$ be a monotone increasings G-nonexpansive mapping. If $C_T := \{x \in C;\ (x,y) \in E(G)\ for\ some\ y \in T(x)\}$ is not empty, then T has an approximate fixed point sequence $(x_n) \in C$, that is, for any $n \geq 1$, there exists $y_n \in T(x_n)$ such that*

$$\lim_{n\to\infty} d(x_n, y_n) = 0.$$

In particular, $\lim\limits_{n\to\infty} dist(x_n, T(x_n)) = 0$, where

$$dist(x_n, T(x_n)) = \inf\{d(x_n, y);\ y \in T(x_n)\}.$$

Proof. Fix $\lambda \in (0,1)$ and $x_0 \in C$. Define the multivalued map T_λ on C by

$$T_\lambda(x) = \lambda\ x_0 \oplus (1-\lambda)\ T(x) = \{\lambda\ x_0 \oplus (1-\lambda)\ y;\ y \in T(x)\}.$$

Note that $T_\lambda(x)$ is a nonempty and closed subset of C. It will be shown that T_λ is a monotone increasing G-contraction. Let $x, y \in C$ be such that $(x,y) \in E(G)$. Since T is a monotone increasing G-nonexpansive mapping, for any $x^* \in T(x)$ there exists $y^* \in T(y)$ such that $(x^*, y^*) \in E(G)$ and $d(x^*, y^*) \leq d(x,y)$. Since

$$d\Big(\lambda\, x_0 \oplus (1-\lambda)\, x^*, \lambda\, x_0 \oplus (1-\lambda)\, y^*\Big) \leq (1-\lambda)\, d(x^*, y^*) \leq (1-\lambda)\, d(x,y),$$

the claim follows. Since G is convex, we get $(\lambda\ x_0 \oplus (1-\lambda)\ x^*, \lambda\ x_0 \oplus (1-\lambda)\ y^*) \in E(G)$. This clearly shows that T_λ is a monotone increasing G-contraction, as claimed. Note that $C_T \subset C_{T_\lambda}$, which implies that C_{T_λ} is nonempty. Using Theorem 10.2 it is readily concluded that T_λ has a fixed point $x_\lambda \in C$. Thus there exists $y_\lambda \in T(x_\lambda)$ such that

$$x_\lambda = \lambda\ x_0 \oplus (1-\lambda)\ y_\lambda.$$

In particular ,

$$d(x_\lambda, y_\lambda)\ \le\ \lambda\ d(x_0, y_\lambda)) \le \lambda\ \delta(C),$$

which implies $dist(x_\lambda, T(x_\lambda)) \le \lambda\ \delta(C)$. Choosing $\lambda = \frac{1}{n}$, for $n \ge 1$, it is easily seen that there exists $x_n \in C$ and $y_n \in T(x_n)$ such that $d(x_n, y_n) \le \delta(C)/n$, which implies

$$dist(x_n, T(x_n)) \le \frac{1}{n}\ \delta(C).$$

The proof of Theorem 10.23 is therefore complete. □

The multivalued version of Theorem 10.22 may be stated as:

Theorem 10.24. *[5] Let (X,d) be a complete hyperbolic metric space and suppose that the triple (X,d,G) has property* **(**)**. *Assume G is convex and transitive. Let C be a nonempty, G-compact and convex subset of X. Then any monotone increasing G-nonexpansive mapping $T : C \to \mathscr{C}(C)$, has a fixed point, provided that $C_T := \{x \in C;\ (x,y) \in E(G)\ for\ some\ y \in T(x)\}$ is not empty.*

Proof. Choose $x_0 \in C_T$. Let (λ_n) be a sequence of numbers in $(0,1)$ such that $\lim\limits_{n\to\infty} \lambda_n = 0$. As in the proof of Theorem 10.23, define the mapping $T_1 : C \to C$ by

$$T_1(x) = \lambda_1 x_0 \oplus (1-\lambda_1)T(x).$$

Since $C_T \subset C_{T_1}$, there exists $y_0 \in T_1(x_0)$ such that $(x_0, y_0) \in E(G)$. The properties of T_1 imply that there exists $y_2 \in T_1(y_1)$ such that $(y_1, y_2) \in E(G)$ and

$$d(y_1, y_2) \le (1-\lambda_1) d(x_0, y_1).$$

By induction, build a sequence (y_n), with $y_0 = x_0$, such that $y_{n+1} \in T_1(y_n)$, $(y_n, y_{n+1}) \in E(G)$ and that

$$d(y_n, y_{n+1}) \le (1-\lambda_1)\ d(y_{n-1}, y_n) \le (1-\lambda_1)^n\ d(x_0, y_1) \le (1-\lambda_1)^n\ \delta(C),$$

for $n \ge 1$. So (y_n) is Cauchy. Set $\lim\limits_{n\to+\infty} y_n = x_1 \in C$. The property **(**)** implies that $(y_n, x_1) \in E(G)$, for any n. In particular, $(x_0, x_1) \in E(G)$. Using the properties of T_1 it is readily seen that for any n, there exists $z_n \in T(x_1)$ such that

$$d(y_{n+1}, z_n) \leq (1-\lambda_1)\, d(y_n, x_1).$$

Clearly this implies that (z_n) converges to x_1 as well. Since $T(x_1)$ is closed, one concludes that $x_1 \in T(x_1)$, i.e., x_1 is a fixed point of T_1. Inductively, construct a sequence (x_n) in C such that x_{n+1} is a fixed point of $T_{n+1} : C \to \mathscr{C}(C)$, defined by

$$T_{n+1}(x) = \lambda_{n+1} x_n \oplus (1-\lambda_{n+1}) T(x),$$

and in such a way that $(x_n, x_{n+1}) \in E(G)$. Since C is G-compact, there exists a subsequence (x_{k_n}) that converges to $\omega \in C$. Since G is transitive, property **(**)** implies that $(x_n, \omega) \in E(G)$. Since x_n is a fixed point of T_n, there exists $z_n \in T(x_n)$ such that

$$x_n = \lambda_n\, x_{n-1} \oplus (1-\lambda_n)\, z_n,$$

for any $n \geq 1$. Notice that $d(x_n, z_n) \leq \ \lambda_n\ d(x_{n_1}, z_n) \leq \lambda_n\ \delta(C)$, for any $n \geq 1$. In particular, it is clear that $\lim\limits_{n\to\infty} d(x_n, z_n) = 0$. It can be inferred from the G-compactness of C that there exists a subsequence (x_{k_n}) which converges to some point $\omega \in C$. Clearly (z_{k_n}) also converges to ω. Using the G-nonexpansiveness of T, since $(x_{k_n}, \omega) \in E(G)$, it is immediate that there exists $\omega_n \in T(\omega)$ such that $d(z_{k_n}, \omega_n) \leq d(x_{k_n}, \omega)$, for any n. Therefore, (ω_n) converges to ω. Since $T(\omega)$ is closed, it follows that $\omega \in T(\omega)$, i.e. ω is a fixed point of T. $\square$

Bibliography

[1] M. R. Alfuraidan, *Fixed points of monotone nonexpansive mappings with a graph*, Fixed Point Theory and Applications **49** (2015) doi: 10.1186/s13663-015-0299-0.

[2] M. R. Alfuraidan, *On monotone Ćirić Quasi-Contraction Mappings with a Graph*, Fixed Point Theory and Applications **93** (2015) doi: 10.1186/s13663-015-0341-2.

[3] M. R. Alfuraidan, *On Monotone Pointwise Contractions in Banach Spaces with a Graph*, Fixed Point Theory and Applications **139** (2015) doi: 10.1186/s13663-015-0390-6.

[4] M. R. Alfuraidan, *Remarks on monotone multivalued mappings on a metric space with a graph*, J. Ineq. Appl. **202** (2015) doi:10.1186/s13660-015-0712-6.

[5] M. R. Alfuraidan, *Topological aspects of weighted graphs with application to fixed point theory*, Appl. Math. Comput., **314** (2017), 287-29.

[6] M. R. Alfuraidan, M. A. Khamsi, *Fixed points of monotone nonexpansive mappings on a hyperbolic metric space with a graph*, Fixed Point Theory and Applications **44** (2015) doi:10.1186/s13663-015-0294-5.

[7] M. R. Alfuraidan, M. A. Khamsi, *On Gregus-'Ciri'c mappings on weighted graphs*, Fixed Point Theory **20:1** (2019), 19-30.

[8] M. R. Alfuraidan, M. Bachar, M. A. Khamsi, *On Monotone Contraction Mappings in Modular Function Spaces*, Fixed Point Theory and Applications **28** (2015) doi:10.1186/s13663-015-0274-9.

[9] M. R. Alfuraidan, M. Bachar, M. A. Khamsi, *Almost Monotone Contractions on Weighted Graphs*, J. Nonlinear Sci. Appl. **8** (2016), 5189-5195.

[10] M. R. Alfuraidan, M. Bachar, M. A. Khamsi, *A graphical version of Reich's fixed point theorem*, J. Nonlinear Sci. Appl. **9** (2016), 3931-3938.

[11] M. Bachar, M. A. Khamsi, *On monotone Ćirić quasi-contraction mappings,* J. Math. Inequal. **10:2** (2016), 511-519.

[12] S. Banach, *Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales*, Fund. Math **3** (1922), 133-181.

[13] I. Beg, A. R. Butt, S. Radojević, *The contraction principle for set valued mappings on a metric space with a graph*, Computers and Mathematics with Applications **60** (2010), 1214-1219.

[14] V. Berinde, *On the approximation of fixed points of weak contractive mappings*, Carpathian J. Math. **19:1** (2003), 7-22.

[15] V. Berinde, M. Păcurar, *Fixed points and continuity of almost contraction*, Fixed Point Theory, **9:1** (2008), 23-34.

[16] M. Besbes, *Points fixes des contractions définies sur un convexe L^0-fermé de L^1*, C.R.A. Sc. de Paris, Tome 311, Serie I (1990), 243-246.

[17] H. Brezis, E. Lieb, *A relation between pointwise convergence of functions and convergence of functionals*, Proc. Amer. Math. Soc., **88:3** (1983), 486-490.

[18] H. Busemann, *Spaces with non-positive curvature*, Acta. Math. **80** (1948), 259-310.

[19] S. K. Chatterjea, *Fixed-point theorems*, C.R. Acad. Bulgare Sci. **25** (1972), 727-730.

[20] V.V. Chistyakov, *Modular metric spaces, I: Basic concepts*, Nonlinear Anal. **72:1** (2010), 1-14.

[21] V.V. Chistyakov, *Modular metric spaces, II: Application to superposition operators*, Nonlinear Anal. **72:1** (2010), 15-30.

[22] Lj. B. Ćirić, *A generalization of Banach's contraction principle*, Proc. Amer. Math. Soc. **45** (1974), 267-273.

[23] Lj. B. Ćirić, *On a generalization of Gregus fixed point theorem*, Czechoslovak. Math. J. **50** (2000), 449-458.

[24] B. Djafari-Rouhani, S. Moradi, *On the existence and approximation of fixed points for Ćirić type contractive mappings*, Quaest. Math. **37:2** (2014), 179-189.

[25] M. Edelstein, *An extension of Banach's contraction principle*, Proc. Amer. Math. Soc. **12** (1961), 7-10.

[26] Y. Feng, S. Liu, *Fixed point theorems for multivalued contractive mappings and multivalued Caristi type mappings*, J. Math. Anal. Appl. **317** (2006), 103-112.

[27] K. Goebel, W. A. Kirk, *Iteration processes for nonexpansive mappings*, Contemp. Math. **21** (1983), 115-123.

[28] K. Goebel, W. A. Kirk, *Topics in Metric Fixed Point Theory*, Cambridge University Press, Cambridge, 1990.

[29] K. Goebel, S. Reich, *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings*, Series of Monographs and Textbooks in Pure and Applied Mathematics, **83**, Dekker, New York, 1984.

[30] M. Gregus Jr., *A fixed point theorem in Banach space*, Boll. Unione Mat. Ital. Sez. A Mat.Soc. Cult. **80:517** (1980), 193-198.

[31] C. A. R. Hoare, *A Model for Communicating Sequential Processes*, Technical Report PRG-22, Programming Research Group, Oxford University Computing Lab. (1981).

[32] S. Ishikawa, *Fixed point by a new iteration*, Proc. Amer. Math. Soc. **44** (1974), 147-150.

[33] J. Jachymski, *The Contraction Principle for Mappings on a Metric Space with a Graph*, Proc. Amer. Math. Soc. **136** (2008), 1359-1373.

[34] R. Kannan, *Some results on fixed points*, Bull. Calcutta Math. Soc., **10** (1968), 71-76.

[35] R. P. Kelisky, T. J. Rivlin, *Iterates of Bernstein polynomials*, Pacific J. Math. **21** (1967), 511-520.

[36] M. A. Khamsi, A. R. Khan, *On monotone nonexpansive mappings in $L^1([0,1])$*, Fixed Point Theory and Applications **94** (2015), doi: 10.1186/s13663-015-0346-x.

[37] M. A. Khamsi, D. Misane, *Disjunctive Signed Logic Programs*, Fundamenta Informaticae, **32** (1996), 349-357.

[38] W. A. Kirk, *Fixed point theory for nonexpansive mappings, I and II*, Lecture Notes in Mathematics, Springer, Berlin, **886** (1981), 485-505.

[39] W. A. Kirk, *A fixed point theorem in CAT(0) spaces and $\mathbb{R}$-trees*, Fixed Point Theory Appl. **4** (2004), 309-316.

[40] D. Klim, D. Wardowski, *Fixed point theorems for set-valued contractions in complete metric spaces*, J. Math. Anal. Appl. **334** (2007), 132-139.

[41] W. M. Kozlowski, *Modular Function Spaces*, Series of Monographs and Textbooks in Pure and Applied Mathematics, **122**, Dekker, New York/Basel, 1988.

[42] M. A. Krasnoselskii, *Two observations about the method of successive approximations*, Uspehi Mat. Nauk. **10** (1955), 123-127.

[43] C. Lennard, *A new convexity property that implies a fixed point property for L_1*, Studia Math. **100:2** (1991), 95-108.

[44] L. Leustean, *A quadratic rate of asymptotic regularity for CAT(0)-spaces*, J. Math. Anal. Appl. **325** (2007), 386-399.

[45] W. R. Mann, *Mean value methods in iterations*, Proc. Amer. Math. Soc. **4** (1953), 506-510.

[46] K. Menger, *Untersuchungen über allgemeine Metrik*, Math. Ann. **100** (1928), 75-163.

[47] N. Mizoguchi, W. Takahashi, *Fixed Point Theorems for Multivalued Mappings on Complete Metric Spaces*, J. Math. Anal. Appl. **141** (1989), 177-188.

[48] S. B. Nadler, *Multivalued contraction mappings*, Pacific J. Math. **30** (1969), 475-488.

[49] J. J. Nieto, R. Rodríguez-López, *Contractive mapping theorems in partially ordered sets and applications to ordinary differential equations*, Order **22:3** (2005), 223-239.

[50] A. C. M. Ran, M. C. B. Reurings, *A fixed point theorem in partially ordered sets and some applications to matrix equations*, Proc. Amer. Math. Soc. **132** (2004), 1435-1443.

[51] S. Reich, *Fixed points of contractive functions*, Bollettino della Unione Mathematica Italiana, **5** (1972), 26-42.

[52] S. Reich, *Some fixed point problems*, Atri. Acad. Nuz. Lincei **57** (1974), 194-198.

[53] S. Reich, I. Shafrir, *Nonexpansive iterations in hyperbolic spaces*, Nonlinear Anal. **15** (1990), 537-558.

[54] I. A. Rus, *Generalized contractions*, Seminar on Fixed Point Theory, **3** (1983), 1-130.

[55] I. A. Rus, *Generalized Contractions and Applications*, Cluj University Press, Cluj-Napoca, 2001.

[56] I. A. Rus, *Iterates of Bernstein operators, via contraction principle*, J. Math. Anal. Appl. **292** (2004), 259-261.

[57] M. Smyth, *Power domains*, J. Comput. System Sci. **16:1** (1978), 23-26.

[58] A. Sultana, V. Vetrivel, *Fixed points of Mizoguchi-Takahashi contraction on a metric space with a graph and applications*, J. Math. Anal. Appl. **417** (2014), 336-344.

[59] M. Taskovic, *Osnove teorije fiksne tacke (Fundamental Elements of Fixed Point Theory)*, Matematicka biblioteka **50**, Belgrad, 1986.

[60] X. A. Udo-utun, Z. U. Siddiqui, M. Y. Balla, *An extension of the contraction mapping principle to Lipschitzian mappings*, Fixed Point Theory Appl. **162** (2015), doi:10.1186/s13663-015-0295-4.

[61] T. Zamfirescu, *Fix point theorems in metric spaces*, Arch. Math. Basel, **23** (1972), 292-298.

Chapter 11

Geometric aspects of generalized metric spaces: Relations with graphs, ordered sets and automata

M. Kabil

Laboratory of Mathematics, Computer Science and Applications, Department of Mathematics, Faculty of Sciences and Technologies Mohammedia, Hassan II University, Casablanca, Morocco,
E-mail : kabilfstm@gmail.com

M. Pouzet

ICJ, Mathématiques, Université Claude-Bernard Lyon1, France and Mathematics & Statistics Department, University of Calgary, Calgary, Canada,
E-mail : pouzet@univ-lyon1.fr

In this Chapter we present a generalization of the notion of metric space and some applications to discrete structures as graphs, ordered sets and transition systems. Results in this direction started in the middle eighties and were based on the impulse given by Quilliot (1983). Graphs and ordered sets were considered as generalized metric spaces equipped with distance functions d that are not real-valued, but are valued on an ordered semigroup equipped with an involution. In this frame, the class of maps preserving graphs or posets coincides with the family of nonexpansive mappings (that is, with the class of maps f such that $d(f(x), f(y)) \leq d(x,y)$, for all x,y). It was observed that many known results on retractions and fixed point property for classical metric spaces (whose morphisms are the nonexpansive mappings) are also valid for these spaces. For example, the characterization of absolute retracts, by Aronszajn and Panitchpakdi (1956), the construction of the injective envelope by Isbell (1965) and the fixed point theorem of Sine and Soardi (1979), translate into the Banaschewski-Bruns theorem (1967), the MacNeille completion of a poset (1933) and the famous Tarski fixed point theorem (1955). This prompted an analysis of several classes of discrete structures from a metric point of view. In this paper, we report the results obtained over the years, with a particular emphasis on the fixed point property.

11.1 Introduction

This Chapter delves into a generalization of metric spaces and its applications to discrete structures as graphs, ordered sets and transition systems. The results presented here originate in a paper by the second author [76], motivated by the work of Quilliot [80, 81]. The genesis of this topic is to be found in two theses [46], [66] and a paper [47]. The theme was subsequently developped in [52], [75], [53], [85], [54], [55], [56], [10], [11] and [61].

Since its introduction by Fréchet (1906), the notion of metric space has motivated many extensions (cf. the encyclopedia [29], also [64, 14, 15, 16], and recently [24]). In the sequel, a *generalized metric space* (see [29] p. 82) is a set E equipped with a *distance*, that is, with a map d from the direct product $E \times E$ into an ordered monoid, say $\mathscr{H}$, equipped with an involution $-$ preserving the order and reversing the monoid operation.This operation will be denoted by $\oplus$ (despite that it is not necessarily commutative) and its neutral element will be denoted by 0.

The conditions on d are the following:

$$d(x,y) = 0 \text{ if and only if } x = y; \tag{11.1}$$

$$d(x,y) = \overline{d(y,x)}; \tag{11.2}$$

$$d(x,y) \leq d(x,z) \oplus d(z,y); \tag{11.3}$$

for all $x,y,z \in E$.

The focus in this Chapter will be on the special case in which the following assumptions are imposed on $\mathscr{H}$.

1. 0 is the least element of $\mathscr{H}$; in which case, condition (i) for d reduces to $d(x,y) = 0$ if and only if $x = y$.

2. $\mathscr{H}$ is a complete lattice and the following distributivity condition holds:
$$\bigwedge_{\alpha \in A, \beta \in B} (p_\alpha \oplus q_\beta) = \bigwedge_{\alpha \in A} p_\alpha \oplus \bigwedge_{\beta \in B} q_\beta$$
for all $p_\alpha \in \mathscr{H}$ $(\alpha \in A)$ and $p_\beta \in \mathscr{H}$ $(\beta \in B)$.

In previous papers (e.g. [76]) such a structure has been called a *Heyting algebra*, or an *involutive Heyting algebra*. This terminology will be retained in this Chapter despite the fact that a more appropriate term could be *dual of an integral involutive quantale*, to refer to the notion of quantale introduced by Mulvey [67] in 1985. Indeed, according to the terminology of [51] (see also [31, 84]), a *quantale* is an ordered monoid satisfying the dual of the distributivity condition stated in (2); it is *involutive* if it is equipped with an involution, and it is *integral* if the largest element is the neutral element of the monoid.

Besides ordinary metric spaces, there are plenty of examples of this generalized structure. Reflexive graphs, undirected as well as directed, ordered sets, involutive and reflexive transition systems are the basic ones. Due to the conditions imposed on $\mathscr{H}$, there are important classes of objects that fall beyond this framework. For example, metric spaces with distances in Boolean algebras, as introduced in [14] (except if the Boolean algebra is the power set of a set); ultrametric spaces with values in an arbitrary poset; graphs which are not necessarily reflexive, or arbitrary transition systems. Attempts to capture these situations have been made in [75]; the case of generalized metric spaces over a Heyting algebra for which the least element is not necessarily the neutral element (cf. condition (1) above) being particularly studied.

We have restricted the scope of our approach to generalized metric spaces over a Heyting algebra because in this class there are significant results, easy to present and with the potential to be extended to more general situations.

The emphasis of this presentation is on retracts and on the fixed point property. Considering the class of generalized metric spaces over a Heyting algebra $\mathscr{H}$, we introduce the *nonexpansive maps* as maps f from a metric space $\mathbf{E} := (E, d)$ into another, say $\mathbf{E}' := (E', d')$, such that

$$d'(f(x), f(y)) \leq d(x, y) \text{ for all } x, y \in E. \tag{11.4}$$

From this starting point, we derive the notions of *isometry*, *retraction* and *coretraction*. Since the Heyting algebra under consideration is a complete lattice, arbitrary products of spaces can be defined, hence, as Duffus and Rival did [30] for graphs and posets, we may introduce *varieties* of metric spaces as classes of metric spaces closed under products and retracts. Among generalized metric spaces, those having the *fixed point property* (fpp), that is, those spaces such that every nonexpansive map f has a fixed point, are of particular interest. As in any category, (fpp) is preserved under retractions. This elementary fact has a significant consequence. Indeed, observing that coretractions are isometric embeddings, those generalized metric spaces for which this necessary condition is sufficient, which are called *absolute retracts*, play a special role. If there are *enough absolute retracts*, meaning that every generalized metric space isometrically embeds into an absolute retract, then absolute retracts are the natural candidates to have the fixed point property. Indeed, it suffices for them to embed into some space with the fixed point property. This point of view is illustrated by the fact that in the category of ordered sets with ordered maps as morphisms, absolute retracts coincide with complete lattices (Banaschewski, Bruns [6]) and according to the famous theorem of Tarski [90], these lattices have the fixed point property. In the category of (ordinary) metric spaces with nonexpansive mappings as morphisms, the absolute retracts are the hyperconvex metric spaces introduced by Aronszajn and Panitchpakdi [3] and on account of the Theorem Sine-Soardi [88, 89], the bounded ones have the fixed point property.

These results being expressible in terms of generalized metric spaces, it was natural to look at absolute retracts in the category of generalized metric spaces over a Heyting algebra. Four basic facts obtained in [47] are presented in this paper. First, we show that on the Heyting algebra $\mathscr{H}$, there is a distance $d_{\mathscr{H}}$ and that every metric space over $\mathscr{H}$ embeds isometrically into some power of the space $\mathbf{H} := (\mathscr{H}, d_{\mathscr{H}})$, equipped with the sup-distance (cf. Theorem 11.1). Next, we show that in this case, the notion of absolute retract is much simpler than in other categories. It coincides with three other notions: *extension property*, *injectivity* and *hyperconvexity* (cf Theorem 11.3). This yields a straightforward extension of the characterization of absolute retracts, due to Aronszajn and Panitchpakdi [3] for ordinary metric spaces. The latter in conjunction with the fact that $\mathbf{H} := (\mathscr{H}, d_{\mathscr{H}})$ is hyperconvex (Theorem 11.2), implies that every generalized metric space embeds isometrically into an absolute retract (cf. (4) of Theorem 11.3). The third fact is the existence of an *injective envelope*, that is, of a minimal injective space extending an arbitrary space isometrically (cf Theorem 11.4). For ordinary metric spaces, this was done by Isbell [45], while for posets, Banaschewski and Bruns [6] showed that the injective envelope of a poset is its MacNeille completion. This last fact is based on the observation that, in general, coretractions are more than isometries. Coretractions preserve *holes*, that is, families of balls with empty intersection. Considering the hole-preserving maps, introduced by Duffus and Rival for posets under the name of *gap preserving maps* [30], and then by Hell an Rival for graphs [40], we show that for the hole-preserving maps, the absolute retracts and the injectives coincide, that every generalized metric space embeds in one of them -by a hole-preserving map- and consequently, that they form a variety (Theorem 11.5).

We illustrate the results about generalized metric spaces presented above with metric spaces, graphs, posets and transition systems. We start with absolute retracts. We mention that the Aronszjan-Panitchpakdi characterization of absolute retracts was extended to ultrametric spaces by Bayod and Martinez [12]. We also refer the reader to Ackerman [2]. Considering reflexive and symmetric graphs, with the usual distance of the shortest path, paths are absolute retracts and every graph isometrically embeds into a product of paths (a result due independently to Quilliot [80], Nowakowski and Rival [71]). Furthermore, it has a minimal retract of product of paths (this last fact has been obtained independently by Pesch [74]). This extends to directed graphs: Quilliot [80] introduced a new kind of distance, the *zigzag distance*, on a directed graph $\mathbf{G}$. This distance takes into account all oriented paths joining two vertices of $\mathbf{G}$. The values of this distance are final segments of the monoid Λ^* of words over the two-letter alphabet $\Lambda := \{+,-\}$. The set $\mathbf{F}(\Lambda^*)$ of these final segments can be viewed as a Heyting algebra. It turns out that this Heyting algebra, not only possesses a metric structure, but it also has a graph structure, rendering it an absolute retract into the category of graphs. Every directed graph embeds isometrically into a power of itself, and the absolute retracts are retracts of products of that graph. The notion of injective envelope of two-element metric spaces was used to produce a family of finite directed graphs generating the variety of absolute retracts. A specialization to posets of the zigzag distance yields the notion of *fence distance* (Quilliot [80]); in this case, absolute retracts of posets are retracts of product of fences (Nevermann, Rival [68]). A graph is a *zigzag* if it symmetrisation is a path. Oriented zigzag graphs are absolute retracts in the variety of directed graphs, but are too simple to generate all absolute retracts in the variety of directed graphs. The full description was given in [55]. As shown in [11], zigzags generate the variety of absolute retracts in the category of oriented graphs. Considering the hole-preserving maps, posets that are absolute retracts are those with the *strong selection property* (notion introduced by Rival and Wille [83] for lattices and extended to posets by Nevermann and Wille [69]). For posets and graphs considered with the fence distance and the graph distance, Theorem 11.5 is due to Nevermann, Rival [68] and Hell, Rival [40], respectively. Of course, Theorem 11.5 applies to directed graphs equipped with the zigzag distance and to classical metric spaces as well.

It appears that the zigzag distance between two vertices x and y of a directed graph $\mathbf{G} := (V, \mathscr{E})$ is the language accepted by the automaton having V as set of states, $T := \{(p,+,q) : (p,q) \in \mathscr{E}\} \cup \{(p,-,q) : (q,p) \in \mathscr{E}\}$ as set of transitions, x as initial state and y as final state. This fact leads to the consideration of transition systems over an arbitrary alphabet Λ as a subclass of metric spaces, the distance between two states being the language accepted between these two states. If the alphabet is equipped with an involution, we may consider *reflexive* and *involutive* transition systems. The distance function takes values in the set $\mathbf{F}(\Lambda^*)$ of final segments of the set

Λ^* of words over the alphabet Λ. As for the two-letter alphabet, $\mathbf{F}(\Lambda^*)$ is a Heyting algebra, and our transition systems are generalized metric spaces, thus the above results apply. The existence of the injective envelope of a two- element metric space was used to prove that $\mathbf{F}(\Lambda^*)$ is a free monoid [56]. A presentation of this result is given in Section 11.7.

Turning to the fixed property, we might say that over the years, fixed point results for discrete of for continuous structures have proliferated. The theorem by Sine-Soardi has been extended to metric spaces endowed with a compact normal structure in the sense of Penot (Kirk's Theorem, [62]). It has also been extended to bounded hyperconvex generalized metric spaces, with an appropriate notion of boundedness [47]. Baillon [5] proved that arbitrary sets of commuting maps on a bounded hyperconvex metric space, have a common fixed point. Khamsi [60] extended the conclusion to metric spaces with a compact normal structure. Quite recently, Khamsi and the second author [61] extended it to generalized metric spaces endowed with a compact normal structure. As a consequence, every set of commuting order-preserving maps on a retract of a power of a finite fence, has a fixed point (the case of one map follows from a result due to I. Rival [82] for finite posets, and from the result by Baclawski and Björner [4] in the case of infinite posets). This applies in the same way to directed graphs (reflexive and antisymmetric) equipped with the zigzag distance and substantially completes the results of Quilliot [80] (Theorem 11.15).

Some aspects of generalized metric spaces are not dealt with in this Chapter. An important such aspect left untouched by us concerns homogeneity and amalgamation. In 1927, Urysohn [91] discovered a separable metric space having the property that every isometry between finite subsets of it extends to an isometry on the whole space and that every finite metric space embeds into it. Later on, Fraïssé [34] and then Jónsson [48], identified the notion of homogeneity and the test of amalgamation, showing that several classes of structures, now called Fraïssé classes (that includes the class of metric spaces), has an homogeneous structure. The existence of the Urysohn space follows then a special case. Then, in 2005, Kechris, Pestov and Todorcevic [59] characterized the classes introduced by Fraïssé with the Ramsey property. This characterisation led to numerous papers on homogeneity and particularly on (ordinary) homogeneous metric and ultrametric spaces [25, 26, 27, 70]. As indicated in [47] (Fact 4 of page 181), the class of metric space over a Heyting algebra has the amalgamation property, thus it may have homogeneous structures (e.g. when the algebra is countable). Independently of our work, some research has been devoted to generalized metric spaces that are also homogeneous [19, 42, 87].

11.2 Metric space over a Heyting algebra

In what follows we introduce the basic terminology to be used in the sequel, see [13, 18, 35]. Let $\mathscr{H}$ be a complete lattice, with a least element, denoted by 0 and a greatest element denoted by 1, equipped both with a monoid operation $\oplus$ and with an involution $-$ satisfying the following properties:
(i) The monoid operation is compatible with the ordering, that is, $p \leq p'$ and $q \leq q'$ imply $p \oplus q \leq p' \oplus q'$ for every $p, p', q, q' \in \mathscr{H}$.
(ii) The involution is order-preserving and reverses the monoid operation, that is,

$$\overline{p \oplus q} = \bar{q} \oplus \bar{p} \text{ holds for every } p, q \in \mathscr{H}.$$

We say that $\mathscr{H}$ is a *Heyting algebra* if it satisfies the following distributivity condition:

$$\bigwedge_{\alpha \in A, \beta \in B} (p_\alpha \oplus q_\beta) = \bigwedge_{\alpha \in A} p_\alpha \oplus \bigwedge_{\beta \in B} q_\beta, \tag{11.5}$$

for all $p_\alpha \in \mathscr{H}$ $(\alpha \in A)$ and $p_\beta \in \mathscr{H}$ $(\beta \in B)$ or equivalently, (because of the involution), if, for all $p_\alpha \in \mathscr{H}$ $(\alpha \in A)$ and $q \in \mathscr{H}$, it holds that

$$\bigwedge_{\alpha \in A} (p_\alpha \oplus q) = \bigwedge_{\alpha \in A} p_\alpha \oplus q. \tag{11.6}$$

Note that the above distributivity condition entails the compatibility of the monoid operation and the ordering.

In the sequel, the following assumption is made:

The least element 0 of $\mathscr{H}$ is the neutral element of the operation $\oplus$.

Let E be a set. A *distance* on E is a map $d : E \times E \to \mathscr{H}$ satisfying the following properties for all $x, y, z \in E$:
d1) $d(x,y) = 0$ if and only if $x = y$;
d2) $d(x,y) \leq \underline{d(x,z)} \oplus d(z,y)$;
d3) $d(x,y) = \overline{d(y,x)}$.
The pair $\mathbf{E} := (E, d)$ is a *metric space* over $\mathscr{H}$. If no confusion arises, the metric space will be denoted simply by the underlying set, E. If we replace the monoid operation $\oplus$ by its reverse, that is by the operation $(x, y) \mapsto y \oplus x$, and leave the ordering and the involution unchanged, then the new structure $\mathscr{H}'$ satisfies the same properties as $\mathscr{H}$; hence we can define distances over $\mathscr{H}'$. For example, if $d : E \times E \to \mathscr{H}$ is a distance then $\bar{d} : E \times E \to \mathscr{H}'$ defined by $\bar{d}(x,y) = d(y,x)$ is a distance over $\mathscr{H}'$; it is called the *dual distance*

to d. We write $\bar{\mathbf{E}} := (E, \bar{d})$ or simply $\bar{E}$ to denote the corresponding space. For typographical reasons, we will use $\bar{d}(x,y)$ instead of $\overline{d(x,y)}$. This will cause no confusion in the sequel.

Let $\mathbf{E} := (E,d)$ be a metric space over $\mathscr{H}$. For all $x \in E$ and $r \in \mathscr{H}$, we define the *ball with center x and radius r*, as the set $B_{\mathbf{E}}(x,r) = \{y \in E : d(x,y) \leq r\}$; if there is no danger of confusion we will denote it simply by $B(x,r)$.

If $\mathbf{E} := (E,d)$ and $\mathbf{E}' := (E',d')$ are two metric spaces over $\mathscr{H}$, then a map $f : E \to E^{'}$ is said to be *nonexpansive* (or *contracting*) provided that

$$d'(f(x), f(y)) \leq d(x,y) \text{ for all } x,y \in E. \tag{11.7}$$

If equality holds in inequality (11.7) for all $x,y \in E$, then f is an *isometry* of $\mathbf{E}$ into $\mathbf{E}'$. Hence, in our terminology, an isometry is not necessarily surjective. We say that $\mathbf{E}$ and $\mathbf{E}'$ are *isomorphic*, and in this case we write $\mathbf{E} \cong \mathbf{E}'$, if there is a surjective isometry from $\mathbf{E}$ onto $\mathbf{E}'$. If E is a subset of E' and the identity map $\mathrm{id} : E \to E'$ is nonexpansive, we say that $\mathbf{E}$ is a *subspace* of $\mathbf{E}'$, or that $\mathbf{E}'$ is an *extension* of $\mathbf{E}$. If, moreover, this map is an isometry (that is, if d is the restriction of d' to $E' \times E'$), then we call $\mathbf{E}$ an *isometric subspace* of $\mathbf{E}'$ and $\mathbf{E}'$ is said to be an *isometric extension* of $\mathbf{E}$. The *restriction of d' to* $\mathbf{E}$, denoted by $d'_{\restriction E}$, is the restriction of the map d' to $E \times E$. This is a distance, the resulting space, denoted by $\mathbf{E}'_{\restriction E} := (E, d'_{\restriction E})$, is the *restriction of* $\mathbf{E}'$ *to* $\mathbf{E}$; this is an isometric subspace of $\mathbf{E}'$. As usual in categories, $Hom(\mathbf{E}, \mathbf{E}')$ denote the set of all nonexpansive maps from $\mathbf{E}$ to $\mathbf{E}'$.

The fact that $\mathscr{H}$ is a complete lattice allows us to define arbitrary products of metric spaces. If $(\mathbf{E}_i)_{i \in I}$, where $\mathbf{E}_i := (E_i, d_i)$, is a family of metric spaces over $\mathscr{H}$, then the direct product $\mathbf{E} := \prod_{i \in I} \mathbf{E}_i$, is the cartesian product $E := \prod_{i \in I} E_i$, equipped with the "sup" (or ℓ^∞) distance $d : E \times E \to \mathscr{H}$ defined by:

$$d\Big((x_i)_{i \in I}, (y_i)_{i \in I}\Big) := \bigvee_{i \in I} d_i(x_i, y_i).$$

The distributivity condition on $\mathscr{H}$ allows to define a distance on the space of values $\mathscr{H}$. This fact relies on the classical notion of *residuation* (see [17, 92]).

Let $v \in \mathscr{H}$. Given $\gamma \in \mathscr{H}$, the sets $\{r \in \mathscr{H} : v \leq r \oplus \gamma\}$ and $\{r \in \mathscr{H} : v \leq \gamma \oplus r\}$ have least elements, that we denote respectively by $\lceil v - \gamma \rceil$ and by $\lceil -\gamma \oplus v \rceil$ (in fact, $\overline{\lceil -\gamma \oplus v \rceil} = \lceil \bar{v} - \bar{\gamma} \rceil$). It follows that for all $p, q \in \mathscr{H}$, the set

$$D(p,q) := \{r \in \mathscr{H} : p \leq q \oplus \bar{r} \text{ and } q \leq p \oplus r\}$$

has a least element, namely $\lceil \bar{p} - \bar{q} \rceil \vee \lceil -p \oplus q \rceil$. We set

$$d_{\mathscr{H}}(p,q) := MinD(p,q). \tag{11.8}$$

As shown in [47]:

Theorem 11.1. *The map $(p,q) \mapsto d_{\mathscr{H}}(p,q)$ is a distance on $\mathscr{H}$ and every metric space over $\mathscr{H}$ embeds isometrically into a power of the space $\mathbf{H} := (\mathscr{H}, d_{\mathscr{H}})$.*

This result follows from the fact that for every metric space $\mathbf{E} := (E,d)$ over $\mathscr{H}$, and for all $x, y \in E$, the following equality holds:

$$d(x,y) = \bigvee_{z \in E} d_{\mathscr{H}}\left(d(z,x), d(z,y)\right). \tag{11.9}$$

Indeed, for each $x \in E$, let $\bar{\delta}(x) : E \to \mathscr{H}$ be the map defined by $\bar{\delta}(x)(z) = d(z,x)$ for all $z \in E$; the equality above reflects the fact that the map from $\mathbf{E}$ into the power $\mathbf{H}^E$ is an isometric embedding (on the other hand, this equality expresses the fact that $\bar{\delta}(x)$ is a nonexpansive map from $\mathbf{E}$ into $\bar{\mathscr{H}} := (\mathscr{H}', d_{\mathscr{H}'})$).

11.3 Examples

11.3.1 Ordinary metric and ultrametric spaces

Let $\mathscr{H} := \mathbb{R}^+ \cup \{+\infty\}$ with addition on the set of non-negative reals extended to $\mathscr{H}$ in the obvious way,and the involution being defined as the identity. The metric spaces we get are just unions of disjoint copies of ordinary metric spaces. The fact that we add the point at infinity to $\mathbb{R}^+$ is an inessential difference. The point at infinity is adjoined in order to make $\mathscr{H}$ a complete poset and to have infinite products, thus avoiding ℓ^∞ type constructions. The distance between to elements in $\mathbf{H} := (\mathscr{H}, d_{\mathscr{H}})$ is given by the absolute value of their difference if both elements are finite; the distance from ∞ to any other element is ∞. Every space in our sense embeds isometrically into a power of $\mathbf{H}$ and, in fact, into a power of $\mathbb{R}^+$ equipped with the absolute value. On the other hand, every ordinary metric space embeds isometrically into some $\ell^\infty_{\mathbb{R}}(I)$, the space of bounded families $(x_i)_{i \in I}$ of real numbers, endowed with the sup-distance.

If the monoid operation on $\mathbb{R}^+ \cup \{+\infty\}$ is the join and the involution is the identity, distances are called *ultrametric distances* and metric spaces are called *ultrametric spaces* (see [12]). The notion of ultrametric spaces has been generalized by several authors (see [78, 79], [2], [19]). The general setting for the space of values is a join semilattice with a least element.

A *join-semilattice* is an ordered set in which two arbitrary elements x and y have a join, denoted by $x \vee y$, defined as the least element of the set of common upper bounds of x and y.

Let $\mathscr{H}$ be a join-semilattice with a least element, denoted by 0. A *pre-ultrametric space* over $\mathscr{H}$ is a pair $\mathbf{D} := (E, d)$ where d is a map from $E \times E$ into $\mathscr{H}$ such that for all $x, y, z \in E$:

$$d(x,x) = 0,\ d(x,y) = d(y,x) \text{ and } d(x,y) \leq d(x,z) \vee d(z,y). \tag{11.10}$$

The map d is an *ultrametric distance* over $\mathscr{H}$ and $\mathbf{D}$ is an *ultrametric space* over $\mathscr{H}$ if $\mathbf{D}$ is a pre-ultrametric space and d satisfies *the separation axiom*:

$$d(x,y) = 0 \text{ implies } x = y. \tag{11.11}$$

Any binary relational structure $\mathbf{M} := (E, (\mathscr{E}_i)_{i \in I})$, in which each $\mathscr{E}_i$ is an equivalence relation on the set E, can be viewed as a pre-ultrametric space on E. Indeed, given a set I, let $\wp(I)$ be its the power set. Then $\wp(I)$, ordered by inclusion, is a join-semilattice (in fact a complete Boolean algebra) , in which the join is the union and the least element 0 is the empty set. For $x, y \in E$, set $d_{\mathbf{M}}(x,y) := \{i \in I : (x,y) \notin \mathscr{E}_i\}$. Then the pair $\mathbf{D}_{\mathbf{M}} := (E, d_{\mathbf{M}})$ is a pre-ultrametric space over $\wp(I)$. Conversely, let $\mathbf{D} := (E, d)$ be a pre-ultrametric space over $\wp(I)$. For every $i \in I$ set $\mathscr{E}_i := \{(x,y) \in E \times E : i \notin d(x,y)\}$ and let $\mathbf{M} := (E, (\mathscr{E}_i)_{i \in I})$. Then $\mathscr{E}_i$ is an equivalence relation on E and $d_{\mathbf{M}} = d$. Furthermore, $\mathbf{D}_{\mathbf{M}}$ is an ultrametric space if and only if $\bigcap_{i \in I} \mathscr{E}_i = \Delta_E := \{(x,x) : x \in E\}$.

The congruences of an algebra form an important class of equivalence relations; they can be studied in terms of ultrametric spaces (see Section 11.8 for an example). If we suppose that our distributivity condition holds, which is for example the case if the set of values is a finite distributive lattice, the study of these ultrametric spaces fits into the analysis of metric spaces over a Heyting algebra. This case was particularly studied in [75] and more recently in [2, 19, 77].

11.3.2 Graphs and digraphs

A *binary relation* on a set E is a subset $\mathscr{E}$ of $E \times E$, the set of ordered pairs (x,y) of elements of E. The *inverse* of $\mathscr{E}$ is the binary relation $\mathscr{E}^{-1} := \{(x,y) : (y,x) \in \mathscr{E}\}$. The *diagonal* of E is the set $\Delta_E = \{(x,x) : x \in E\}$. A *directed graph* $\mathbf{G}$ is a pair $(E, \mathscr{E})$, where $\mathscr{E}$ is a binary relation on E. We say that $\mathbf{G}$ is *reflexive* if $\mathscr{E}$ is *reflexive*, that is, if $\mathscr{E}$ contains the diagonal Δ_E; G is said to be *oriented* if $\mathscr{E}$ is *antisymmetric*, that is, (x,y) and (y,x) cannot be in $\mathscr{E}$ simultaneously except if $x = y$. If $\mathscr{E}$ is *symmetric*, that is in the case that $\mathscr{E} = \mathscr{E}^{-1}$, we identify it with a subset of pairs of E and we say that the graph is *undirected*. If $\mathbf{G} := (E, \mathscr{E})$ and $\mathbf{G}' := (E', \mathscr{E}')$ are two directed graphs, a *homomorphism from* $\mathbf{G}$ *to* $\mathbf{G}'$ is a map $h : E \to E'$ such that $(h(x), h(y)) \in \mathscr{E}'$ whenever $(x, y) \in \mathscr{E}$, for every $(x, y) \in E \times E$.

In the sequel, all graphs we consider will be reflexive. Hence, graph-homomorphisms can send edges or arcs on loops. We refer to [18] for the terminology on graphs.

11.3.2.1 Reflexive graphs

Let $\mathscr{H}$ be the complete lattice consisting of three elements such that "$0 < \frac{1}{2} < 1$".

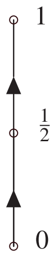


FIGURE 11.1: The ordered monoid $\mathscr{H}$.

The monoid operation is defined by $x \oplus y = \min\{x + y, 1\}$ and the involution is the identity.
Every symmetric reflexive graph $\mathbf{G} := (E, \mathscr{E})$ is a metric space over $\mathscr{H}$. The distance $d : E \times E \longrightarrow \mathscr{H}$ is defined by:

1. $d(x, y) = 1$ if $(x, y) \notin \mathscr{E}$;
2. $d(x, y) = \frac{1}{2}$ if $(x, y) \in \mathscr{E}$ and $x \neq y$;
3. $d(x, y) = 0$ if $x = y$.

Conversely, every metric space $\mathbf{E} := (E, d)$ over $\mathscr{H}$ can be viewed as a symmetric reflexive graph; the vertices are the elements of E and the set of edges $\mathscr{E}$ (including the loops) is defined as follows:

$$(x, y) \in \mathscr{E} \Longleftrightarrow d(x, y) \leq \frac{1}{2}.$$

Nonexpansive maps correspond to graph-homomorphisms (provided that edges are sent to edges or loops).

The distance $d_{\mathscr{H}}$ on the Heyting algebra takes on the value $\frac{1}{2}$ on the pairs $(x,y) \in \{(0,\frac{1}{2}),(\frac{1}{2},0),(\frac{1}{2},1),(1,\frac{1}{2})\}$, the value 1 on the pairs $(0,1)$ and $(1,0)$, and the 0 on the diagonal. The corresponding graph $\mathbf{G}_{\mathscr{H}}$ is the path P_3 on three vertices with $\frac{1}{2}$ as a middle point.

11.3.2.2 Reflexive digraphs

Let $\mathscr{H}$ be the complete lattice consisting of five elements $\{0,\frac{1}{2},+,-,1\}$, represented below:

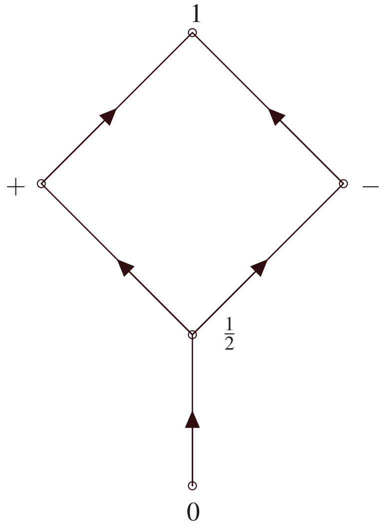


FIGURE 11.2: The ordered monoid $\mathscr{H}$.

The monoid operation is defined by

$$\begin{cases} x \oplus y = 1 \ \text{ if } x,y \geq \frac{1}{2}; \\ x \oplus y = \max(x,y) \ \text{ otherwise.} \end{cases}$$

The involution exchanges $+$ and $-$ and fixes 0, $\frac{1}{2}$ and 1.
If $\mathbf{G} := (E,\mathscr{E})$ is a reflexive directed graph, the application $d : E \times E \longrightarrow \mathscr{H}$ defined by

1. $d(x,y) = 1$ if $(x,y) \notin \mathscr{E} \cup \mathscr{E}^{-1}$;
2. $d(x,y) = +$ if $(x,y) \in \mathscr{E} \backslash \mathscr{E}^{-1}$;
3. $d(x,y) = -$ if $(x,y) \in \mathscr{E}^{-1} \backslash \mathscr{E}$;

4. $d(x,y) = \frac{1}{2}$ if $(x,y) \in \mathscr{E} \bigcap \mathscr{E}^{-1} \backslash \Delta_E$;

5. $d(x,y) = 0$ if $(x,y) \in \Delta_E$,

is a distance on E.
Conversely, every metric space (E,d) over $\mathscr{H}$ can be viewed as a reflexive digraph; the vertices are the elements of E and the set of arcs $\mathscr{E}$ is defined as follows:

$$(x,y) \in \mathscr{E} \Longleftrightarrow d(x,y) \leq +.$$

11.3.2.3 The graphic distance

A graph **P** is a *path* if we can enumerate the vertices in a non-repetitive sequence $(x_i)_{i \in I}$, where either $I = \{0,1,...,n\}$, $I = \mathbb{N}$, or $I = \mathbb{Z}$, in such a way(x_i, x_j) forms an edge if and only if $|j-i| \leq 1$; the path **P** is said to be *finite* if $I = \{0,1,...,n\}$ and in this case n is its *length*, whereas **P** is said to be*infinite* if $I = \mathbb{N}$, and *doubly infinite* if $I = \mathbb{Z}$. If $\mathbf{G} := (V, \mathscr{E})$ is an (undirected) graph, the *graphic distance* is the map $d_G : V \times V \longrightarrow \mathbb{N} \cup \{+\infty\}$, for which $d_G(x,y)$ is the length of the shortest path connecting x to y (if there is a such a path) and $+\infty$ otherwise. This is a distance on $\mathscr{H} := (\mathbb{N} \cup \{+\infty\}, \oplus)$, where $\oplus$ is the ordinary sum. The distance on $\mathscr{H}$ defined by means of Formula (11.8) is the graphic distance associated with the graph $\mathbf{G}_{\mathscr{H}}$, made of a one-way infinite path and an isolated vertex. Not every metric space over $\mathscr{H}$ comes from a graph. Still, with the fact that $\mathbf{G}_{\mathscr{H}}$ embeds isometrically into an infinite product of finite paths, it follows from Theorem 11.1 that *every graph embeds into a product of finite paths*, a result due to Nowakowski-Rival [71] and Quilliot [80].

11.3.2.4 The zigzag distance

A *reflexive zigzag* is a reflexive graph **L** whose symmetric hull is a path. If $\mathbf{L} := (L, \mathscr{L})$ is a finite reflexive oriented zigzag, we may enumerate the vertices in a non-repeating sequence $v_0 := x, \ldots, v_n := y$ and to this enumeration we may associate the finite sequence $ev(\mathbf{L}) := \alpha_0 \cdots \alpha_i \cdots \alpha_{n-1}$ of $+$ and $-$, where $\alpha_i := +$ if $(v_i, v_{i+1}) \in \mathscr{L}$ and $\alpha_i := -$ if $(v_{i+1}, v_i) \in \mathscr{L}$. We call such a sequence a *word* over the *alphabet* $\Lambda := \{+,-\}$. If the path has just one vertex, the corresponding word is the empty word, that we denote by $\square$. Conversely, to a finite word $u := \alpha_0 \cdots \alpha_i \cdots \alpha_{n-1}$ over Λ we may associate the reflexive oriented zigzag $\mathbf{L}_u := (\{0, \ldots n\}, \mathscr{L}_u)$ with end-points 0 and n (where n is the length $\mid u \mid$ of u) such that

$$\mathscr{L}_u = \{(i,i+1) : \alpha_i = +\} \cup \{(i+1,i) : \alpha_i = -\} \cup \Delta_{\{0,\ldots,n\}}.$$

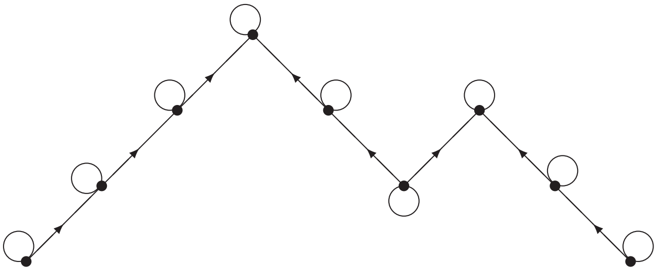


FIGURE 11.3: A reflexive oriented zigzag.

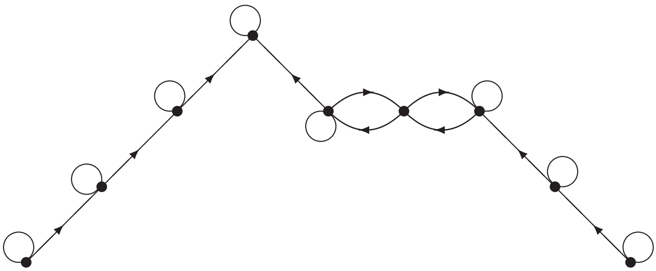


FIGURE 11.4: A reflexive directed zigzag.

Let $\mathbf{G} := (E, \mathscr{E})$ be a reflexive directed graph. For each pair $(x,y) \in E \times E$, the *zigzag distance* from x to y is the set $d_{\mathbf{G}}(x,y)$ of words u such that there is a nonexpansive map h from $\mathbf{L}_u$ into $\mathbf{G}$, which sends 0 to x and $\mid u \mid$ to y.

Because of the reflexivity of $\mathbf{G}$, every word obtained from a word belonging to $d_{\mathbf{G}}(x,y)$ by inserting letters into it, will also be into $d_{\mathbf{G}}(x,y)$. This leads to the following discussion.

Let Λ^* be collection of words over the alphabet $\Lambda := \{+,-\}$. Extend the involution on Λ to Λ^* by setting $\overline{\Box} := \Box$ and $\overline{u_0 \cdots u_{n-1}} := \overline{u_{n-1}} \cdots \overline{u_0}$, for every word in Λ^*. Order Λ^* by the *subword ordering*, denoted by $\leq$ and defined in the following way: If $u := \alpha_1\alpha_2\ldots\alpha_m, v := \beta_1\beta_2\ldots\beta_n \in \Lambda^*$ set:

$$u \leq v \text{ if and only if } \alpha_j = \beta_{i_j} \text{ for all } j = 1,\ldots m, \text{ with some } 1 \leq j_1 < \ldots j_m \leq n. \tag{11.12}$$

Let $\mathbf{F}(\Lambda^*)$ be the set of final segments of Λ^*. A final segment of Λ^* is a subset F of Λ^* such that $u \in F$ and $u \leq v$ implies $v \in F$. Setting $\overline{X} := \{\overline{u} : u \in X\}$ for a set X of words, we observe that $\overline{X}$ belongs to $\mathbf{F}(\Lambda^*)$. Order $\mathbf{F}(\Lambda^*)$ by the reverse of the inclusion, denote its least element by 0 (observe that this is Λ^*, the final segment generated by the empty word), set uv for the concatenation of two words $u, v \in \Lambda^*$ and $X \oplus Y$ for the concatenation $XY := \{uv : u \in X, v \in Y\}$. Then, it is easy to see that $\mathscr{H}_\Lambda := (\mathbf{F}(\Lambda^*), \oplus, \supseteq, 0, -)$ is an involutive Heyting algebra. This leads us to the consideration of distances and metric spaces over $\mathscr{H}_\Lambda$.

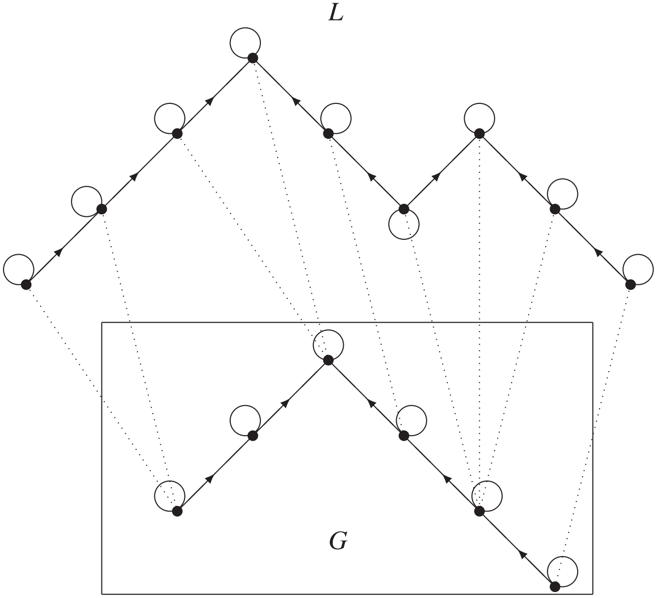


FIGURE 11.5: A morphism from an oriented zigzag L into a directed graph G.

There are two simple and crucial facts in the consideration of the zigzag distance (see [47]).

Lemma 11.1. *A map from a reflexive directed graph* **G** *into another is a graph-homomorphism, iff it is nonexpansive.*

Lemma 11.2. *The distance d of a metric space $\mathbf{E} := (E, d)$ over $\mathscr{H}_\Lambda$ is the zigzag distance of a reflexive directed graph $\mathbf{G} := (E, \mathscr{E})$ iff it satisfies the following property: for all $x, y, z \in E$, $u, v \in \Lambda^*$, $uv \in d(x, y)$ implies $u \in d(x, z)$ and $v \in d(z, y)$ for some $z \in E$. When this condition holds, $(x, y) \in \mathscr{E}$ iff $+ \in d(x, y)$.*

On account of Lemma 11.2, the various metric spaces mentioned in the introduction (injective, absolute retracts, etc.) are graphs equipped with the zigzag distance; in particular, the distance $d_{\mathscr{H}_\Lambda}$ defined on $\mathscr{H}_\Lambda$ is the zigzag

distance of some graph, say $\mathbf{G}_{\mathscr{H}_\Lambda}$. According to Theorem 11.1, every graph embeds isometrically into some power of $\mathbf{G}_{\mathscr{H}_\Lambda}$. This graph is countably infinite (this follows from Higman's theorem on words [41]) but it is not easy to describe. From the study of hyperconvexity (see Section 11.4.2) it follows that it embeds isometrically (w.r.t. the zigzag distance) into a product of its restrictions to principal initial segments of $\mathscr{H}_\Lambda$. Hence every graph isometrically embeds into a product of these finite graphs. The latter fact leads to a fairly precise description of absolute retracts in the category of reflexive directed graphs (see [55]).
The notion of zigzag distance is due to Quilliot [80, 81]. He considered reflexive directed graphs, not necessarily oriented and, in defining the distance, considered only oriented paths. The consideration of the set of values of the distance, namely $\mathscr{H}_\Lambda$, is in [76]. A general study is presented in [47]; some developments appear in [85] and [55].

11.3.3 Ordered sets

Let $\mathscr{H}$ be the following structure. The domain is the set $\{0,+,-,1\}$. The order is $0 \leq +,- \leq 1$, with $+$ incomparable to $-$; the involution exchanges $+$ and $-$ and fixes 0 and 1; the operation $\oplus$ is defined by $p \oplus q := p \vee q$ for every $p,q \in V$. As it is easy to check, $\mathscr{H}$ is an involutive Heyting algebra. If (E,d) is metric space over $\mathscr{H}$, then $\mathbf{P}_d := (E,\delta_+)$, where $\delta_+ := \{(x,y) : d(x,y) \leq +\}$, is an ordered set. Conversely, if $\mathbf{P} := (E,\leq)$ is an ordered set, then the map $d : E \times E \to \mathscr{H}$ defined by $d(x,y) := 0$ if $x = y$, $d(x,y) := +$ if $x < y$, $d(x,y) := -$ if $y < x$ and $d(x,y) := 1$ if x and y are incomparable, is a distance over $\mathscr{H}$. Clearly, if $\mathbf{E} := (E,d)$ and $\mathbf{E}' := (E',d')$ are two metric spaces over $\mathscr{H}$, a map $f : E \to E'$ is nonexpansive from $\mathbf{E}$ into $\mathbf{E}'$ iff it is order-preserving as a map from $\mathbf{P}_d$ into $\mathbf{P}_{d'}$. Depending on the value of their radius, $v \in \mathscr{H}$, a metric space over $\mathscr{H}$ has four types of balls: singletons, corresponding to $v = 0$, the full space, corresponding to $v = 1$, the principal final segments, $\uparrow x := \{y \in E : x \leq y\}$, corresponding to balls $B(x,+)$, and principal initial segments, $\downarrow x := \{y \in E : y \leq x\}$, corresponding to balls $B(x,-)$. The set $\mathscr{H}$ can be equipped with the distance $d_{\mathscr{H}}$ given by formula (11.8). The corresponding poset is the four element lattice $\{-,0,1,+\}$, with $0 < -,+ < 1$. The retracts of powers of this lattice are all complete lattices.

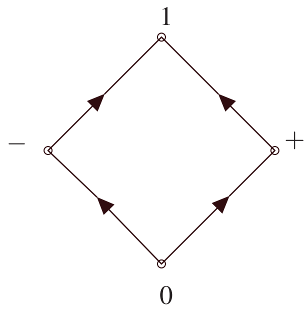
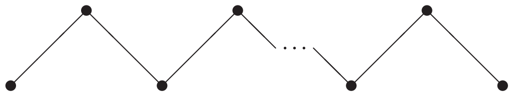


FIGURE 11.6: The ordered monoid $\mathscr{H}$.

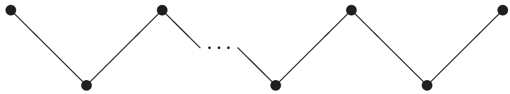
The fact, due to Birkhoff, that every poset embeds into a power of the two-element chain $\underline{\mathbf{2}} := \{0,1\}$ is the translation in terms of posets of Theorem 11.1.

11.3.4 The fence distance on posets

If we view an ordered set as a directed graph, we may associate its zigzag distance to it. In this case, the reflexive oriented zizags defined at the begining of Subsubsection 11.3.2.4 reduce to *fences*. Indeed, a *fence* is a poset whose comparability graph is a path. For example, a two-element chain is a fence. Each larger fence has two orientations, for example on the three vertices path, these orientations yield the $\bigvee$ and the $\bigwedge$. The $\bigvee$ is the 3-element poset consisting of $0,+,-$, with $0 < +,-$ and $+$ incomparable to $-$. The $\bigwedge$ is its dual. More generally, for each integer n, there are two fences of length n: the *up-fence* and the *down-fence*. The first one starts with $x_0 < x_1 > \ldots$, the second with $x_0 > x_1 < \ldots$ For $n := 2$ one gets $\bigwedge$ and $\bigvee$ respectively.



Up-fence



Down-fence

FIGURE 11.7: Up-fence and Down-fence.

Let $\mathbf{P} := (E, \leq)$ be a poset. If two vertices x and y are connected in the comparability graph of $\mathbf{P}$, one may map some fence into $\mathbf{P}$ by an order-preserving map sending the extremities of the fence onto x and y. One can then define the distance $d_{\mathbf{P}}(x,y)$ between x and y as the pair (n,m) of integers such that n (resp. m), is the shortest length of an up-fence (resp. a down fence), whose extremities can be mapped onto x and y. If x and y are not connected in the comparability graph of $\mathbf{P}$, one sets $d_{\mathbf{P}}(x,y) = +\infty$. For example, if $x < y$ then $d_{\mathbf{P}}(x,y) = (1,2)$. This distance is defined in [68], an alternative definition is in [47].

Let $\mathscr{H} := \{(n,m) \in (\mathbb{N} \setminus \{0\})^2 : |n-m| \leq 1\} \cup \{(0,0), +\infty\} \setminus \{(1,1)\}$, the pairs being ordered componentwise and $+\infty$ being at the top. The involution transforms (n,m) into (m,n). The sum $(n,m) \oplus (n',m')$ is $(n \oplus n', m \oplus m')$ where $n \oplus n'$ is $n+n'-1$ if n is odd, or $n+n'$ otherwise. With this operation, $\mathscr{H}$ forms a Heyting algebra. If $\mathbf{P} := (E, \leq)$ is a poset then $d_{\mathbf{P}} : E \times E \to \mathscr{H}$ is a distance over $\mathscr{H}$. According to Theorem 11.1, this Heyting algebra has a metric structure $\mathbf{H}$ and every metric space over $\mathscr{H}$ embeds isometrically into a power of $\mathbf{H}$. It turns out that $\mathbf{H}$ is the metric space associated to a poset $\mathbf{P}_{\mathscr{H}}$ (to see this, set $x \leq y$ if $x = y$ or 1 is the first component of $d_{\mathscr{H}}(x,y)$). This poset is represented below. Hence every poset embeds isometrically into a power of $\mathbf{P}_{\mathscr{H}}$. From the study of hyperconvexity in Section 11.4.2 it follows that this poset embeds isometrically into a product of fences, hence every poset embeds isometrically into a retract of fences ([80]). For more, see Nevermann-Rival, 1985 and Jawhari-al 1986.

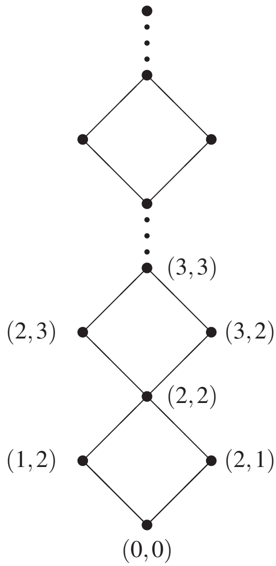


FIGURE 11.8: The ordered monoid $\mathscr{H}$.

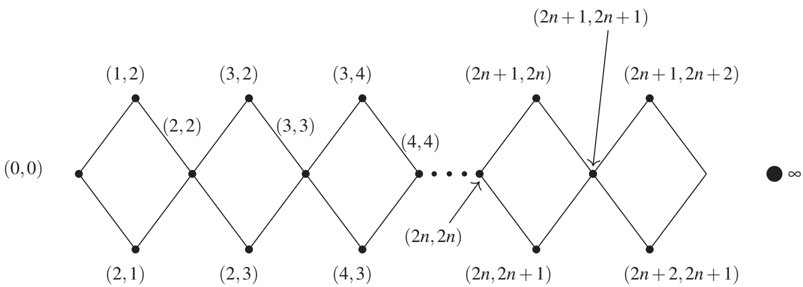


FIGURE 11.9: The poset $P_{\mathscr{H}}$.

11.3.5 Transitions systems

The zigzag distance is a special case of distance defined on transition systems. Indeed, it $\mathbf{M}$ is a transition system on an alphabet Λ, we may define the distance $d_{\mathbf{M}}(x,y)$ from a state x to a state y as the language accepted by the automaton $\mathscr{A}_{x,y} := (\mathbf{M},\{x\},\{y\})$, whose initial and final states are x y, respectively. Once the alphabet is equipped with an involution, this distance takes values in a Heyting algebra in which the neutral element is no longer the least element and satisfies conditions (11.1) in the introduction. As it turns out, if we view a reflexive graph as a transition system of a special form, the zigzag distance is the distance on that transition system. Next, we present the details of this claim.

Let Λ be a set. Consider Λ as an *alphabet* whose members are *letters* and extend the above discussion for two-letter alphabets to Λ. We write a word α with a mere juxtaposition of its letters as $\alpha = a_0 \dots a_{n-1}$, where a_i are letters from Λ for $0 \leq i \leq n-1$. The integer n is the *length* of the word α; it is denoted by $|\alpha|$. Hence we identify letters with words of length 1. We denote the empty word by $\square$, which is the unique word of length zero, by $\square$. The *concatenation* of two word $\alpha := a_0 \cdots a_{n-1}$ and $\beta := b_0 \cdots b_{m-1}$ is the word $\alpha\beta := a_0 \cdots a_{n-1} b_0 \cdots b_{m-1}$. We denote by Λ^* the set of all words on the alphabet Λ. Once equipped with the concatenation of words, Λ^* is a monoid, whose neutral element is the empty word, in fact Λ^* is the *free monoid* on Λ. A *language* is any subset X of Λ^*. We denote by $\wp(\Lambda^*)$ the set of languages. We will use capital letters for languages. If $X,Y \in \wp(\Lambda^*)$ the *concatenation* of X and Y is the set $XY := \{\alpha\beta : \alpha \in X, \beta \in Y\}$ (and we will use Xy and xYinstead of $X\{y\}$ and $\{x\}Y$). This operation extends the concatenation operation on Λ^*; with it, the set $\wp(\Lambda^*)$ is a monoid whose neutral element is the set $\{\square\}$.

Ordered by inclusion, this is a (join) lattice ordered monoid. Indeed, concatenation distributes over arbitrary unions, namely:

$$(\bigcup_{i\in I} X_i)Y = \bigcup_{i\in I} X_i Y.$$

But concatenation does not distribute over intersections (for a simple example, let $\Lambda := \{a,b,c\}$, $I := \{1,2\}$, $X_1 := \{ab\}$, $X_2 := \{a\}$, $Y := \{c,bc\}$, then $\emptyset = (X_1 \cap X_2)Y \neq X_1Y \cap X_2Y = \{abc\}$). Ordered by the reverse of the inclusion, the monoid $\wp(\Lambda^*)$ becomes a Heyting algebra (ordered by inclusion, however, it is not), in the sense that it satisfies the distributivity condition (11.5). If $-$ is an involution on Λ, it extends to an involution on Λ^*, by setting $\overline{\square} := \square$, and $\overline{\alpha} = \overline{a_{n-1}} \dots \overline{a_0}$ if $\alpha = a_0 \dots a_{n-1}$. This involution reverses the concatenation of words. Extended to $\wp(\Lambda^*)$ by setting $\overline{X} := \{\overline{\alpha} : \alpha \in X\}$, it reverses the concatenation of languages and preserves the inclusion order on languages. The set $\wp(\Lambda^*)$, with the concatenation of languages as a monoid operation, the reverse of the inclusion as order and the extension of the involution, is a Heyting algebra. But in this Heyting algebra, the neutral element (namely $\{\Box\}$), is not the least element.

We suppose from now on that the alphabet Λ is ordered.

We order Λ^* with the Higman ordering [41] that is, if α and β are two elements in Λ^* such $\alpha := a_0 \cdots a_{n-1}$ and $\beta := b_0 \cdots b_{m-1}$ then $\alpha \leq \beta$ if there is an injective and increasing map h from $\{0, ..., n-1\}$ to $\{0, ..., m-1\}$ such that for each i, $0 \leq i \leq n-1$, we have $a_i \leq b_{h(i)}$. Then Λ^* is an ordered monoid with respect to the concatenation of words. A *final segment* of Λ^* is any subset $F \subseteq \Lambda^*$ such that $\alpha \leq \beta, \alpha \in F$ implies $\beta \in F$. Initial segments are defined dually.

Let $\mathbf{F}(\Lambda^*)$ be the collection of final segments of Λ^*. The set $\mathbf{F}(\Lambda^*)$ is stable with respect to the concatenation of languages: if $X, Y \in \mathbf{F}(\Lambda^*)$, then $XY \in \mathbf{F}(\Lambda^*)$ (indeed, if $u, v, w \in \Lambda^*$ with $uv \leq w$, then $w = u'v'$ with $u \leq u'$ and $v \leq v'$). Clearly, the neutral element is Λ^*. The set $\mathbf{F}(\Lambda^*)$ ordered by inclusion is a complete lattice (the join is the union, the meet is the intersection). Concatenation distributes over unions. If we order $\mathbf{F}(\Lambda^*)$ by the reverse of the inclusion, denote $X \leq Y$ instead of $X \supseteq Y$, and set $\mathbf{1} := \Lambda^*$, we have the exact generalization obtained for a two-letter alphabet.

Lemma 11.3. *The set $\mathscr{H}_\Lambda := (\mathbf{F}(\Lambda^*), \oplus, \supseteq, \mathbf{1}, -)$, where $\oplus$ denotes the concatenation of languages, is a Heyting algebra and $\mathbf{1}$ is its least element.*

In contrast to the case of the power set, in $\mathbf{F}(\Lambda^*)$ concatenation distributes over intersections:

Lemma 11.4. $(\bigcap_{i \in I} X_i)Y = \bigcap_{i \in I} X_i Y$ *for all final segments X_i and Y of Λ^*.*

Proof. The inclusion $(\bigcap_{i \in I} X_i)Y \subseteq \bigcap_{i \in I} X_i Y$ is obvious. For the proof of the reverse inclusion, let $z \in \bigcap_{i \in I} X_i Y$. For every $i \in I$ there are $x_i \in X_i$ and $y_i \in Y$ such that $z = x_i y_i$. Let y be the shortest suffix of z such that $y = y_{i_0}$ for some $i_0 \in I$ and let $x \in \Lambda^*$ such that $z = xy$. We claim that $x \in \bigcap_{i \in I} X_i$. Indeed, let $j \in I$. We have $z = x_j y_j$ and $z = x_{i_0} y_{i_0}$. By the minimality of y_{i_0}, we have $x_j \leq x_{i_0} = x$, hence $x \in X_j$ since X_j is a final segment of Λ^*. This proves our claim. Since $z = xy$, $z \in (\bigcap_{i \in I} X_i)Y$, as required. $\Box$

We refer to [86] for the standard terminology on transition systems. A *transition system* on the *alphabet* Λ is a pair $\mathbf{M} := (Q, T)$, where $T \subseteq Q \times \Lambda \times Q$. The elements of Q are called *states* and those of T are referred to as *transitions*. Let $\mathbf{M} := (Q, T)$ and $\mathbf{M}' := (Q', T')$ be two transition systems on the alphabet Λ. A map $f : Q \longrightarrow Q'$ is a *morphism* of transition systems if for every transition $(p, \alpha, q) \in T$, we have $(f(p), \alpha, f(q)) \in T'$. When

f is bijective and f^{-1} is a morphism from $\mathbf{M}'$ to $\mathbf{M}$, we say that f is an *isomorphism*.

An *automaton* $\mathscr{A}$ on the alphabet Λ consists of a transition system $\mathbf{M} := (Q,T)$ together with two subsets I, F of Q called the sets of *initial* and *final states*, respectively. We denote the automaton as a triple $(\mathbf{M},I,F)$. A *path* in the automaton $\mathscr{A} := (\mathbf{M},I,F)$ is a sequence $c := (e_i)_{i<n}$ of consecutive transitions, that is, of transitions $e_i := (q_i,a_i,q_{i+1})$. The word $\alpha := a_0 \cdots a_{n-1}$ is the *label* of the path, the state q_0 is its *origin* and the state q_n is its *end*. For each state q in Q, we define a unique null path of length 0 with origin and end at q. Its label is the empty word $\Box$. A path is *successful* if its origin is in I and its end is in F. Finally, a word α on the alphabet Λ is *accepted* by the automaton $\mathscr{A}$ if it is the label of some successful path. The *language accepted* by the automaton $\mathscr{A}$, denoted by $L_{\mathscr{A}}$, is the set of all words accepted by $\mathscr{A}$. Let $\mathscr{A} := (\mathbf{M},I,F)$ and $\mathscr{A}' := (\mathbf{M}',I',F')$ be two automata. A *morphism* from $\mathscr{A}$ to $\mathscr{A}'$ is a map $f : Q \longrightarrow Q'$ satisfying the following two conditions:

1. f is a morphism from $\mathbf{M}$ to $\mathbf{M}'$;

2. $f(I) \subseteq I'$ and $f(F) \subseteq F'$.

If, moreover, f is bijective, $f(I) = I', f(F) = F'$ and f^{-1} is also a morphism from $\mathscr{A}'$ to $\mathscr{A}$, we say that f is an *isomorphism* and that the two automata $\mathscr{A}$ and $\mathscr{A}'$ are *isomorphic*.

To any metric space $\mathbf{E} := (E,d)$ over $\mathscr{H}_\Lambda := \mathbf{F}(\Lambda^*)$, we may associate the transition system $\mathbf{M} := (E,T)$ having E as set of states and $T := \{(x,a,y) : a \in d(x,y) \cap \Lambda\}$ as its set of transitions. Notice that such a transition system has the following properties: For all $x,y \in E$ and every $a,b \in \Lambda$ with $b \geq a$,
1) $(x,a,x) \in T$;
2) $(x,a,y) \in T$ implies $(y,\overline{a},x) \in T$;
3) $(x,a,y) \in T$ implies $(x,b,y) \in T$.
We say that a transition system satisfying the above properties is *reflexive* and *involutive* (cf. [85], [55]). Clearly if $\mathbf{M} := (Q,T)$ is such a transition system, the map $d_{\mathbf{M}} : Q \times Q \longrightarrow \mathscr{H}_\Lambda$, where $d_{\mathbf{M}}(x,y)$ is the language accepted by the automaton $(\mathbf{M}, \{x\}, \{y\})$, is a distance. We have the following:

Lemma 11.5. *Let $\mathbf{E} := (E,d)$ be a metric space over $\mathscr{H}_\Lambda := \mathbf{F}(\Lambda^*)$. The following properties are equivalent:*

1. *The map d is of the form $d_{\mathbf{M}}$ for some reflexive and involutive transition system $\mathbf{M} := (E,T)$;*

2. *For all $\alpha,\beta \in \Lambda^*$ and x, $y \in E$, if $\alpha\beta \in d(x,y)$, then there is some $z \in E$ such that $\alpha \in d(x,z)$ and $\beta \in d(z,y)$.*

Lemma 11.6. *Let* $\mathbf{M}_i := (Q_i, T_i)\,(i = 1,2)$ *be two reflexive and involutive transition systems. A map* $f : Q_1 \longrightarrow Q_2$ *is a morphism from* $\mathbf{M}_1$ *to* $\mathbf{M}_2$ *if only if* f *is a nonexpansive map from* $(Q_1, d_{\mathbf{M}_1})$ *to* $(Q_2, d_{\mathbf{M}_2})$.

From Lemma 11.6, the category of reflexive and involutive transition systems with the morphisms defined above can be identified with a subcategory of the category whose objects are the metric spaces, with the nonexpansive maps as morphisms.

As is the case with directed graphs, Lemma 11.5 ensures that the various metric spaces mentioned in the introduction (injective, absolute retracts, etc.) come from transition systems. In particular, the distance $d_{\mathscr{H}_\Lambda}$ defined on $\mathscr{H}_\Lambda$ is the distance of some transition system, say $\mathbf{M}_{\mathscr{H}_\Lambda}$. According to Theorem 11.1, every reflexive, involutive transition system embeds isometrically into some power of $\mathbf{M}_{\mathscr{H}_\Lambda}$. As in the case of graphs, this transition system is countably infinite (for more, see [55, 56, 57]).

11.4 A categorical approach of generalized metric spaces

Let $\mathscr{C}$ be a category, with objects, say $\mathbf{P}$, $\mathbf{Q}$, ... and morphisms f, g,.... We say that the object $\mathbf{P}$ is a *retract* of the object $\mathbf{Q}$ and we write $\mathbf{P} \lhd \mathbf{Q}$ if there are morphisms $f : \mathbf{P} \longrightarrow \mathbf{Q}$ and $g : \mathbf{Q} \longrightarrow P$ such that $g \circ f = \mathrm{id}_{\mathbf{P}}$, where $\mathrm{id}_{\mathbf{P}}$ is the identity map on $\mathbf{P}$.

We illustrate this definition with two examples:

1. The objects of the category are the posets and the morphisms are the order-preserving maps (i.e. the maps f such that $x \leq y$ implies $f(x) \leq f(y)$).

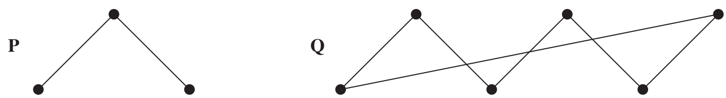


FIGURE 11.10: $\mathbf{P}$ is retract of $\mathbf{Q}$.

2. The objects of the category are all reflexive graphs (which are the undirected graphs with a loop at every vertex, or, equivalently, the reflexive and symmetric binary relations) and the morphisms are all edge-preserving maps (note that an edge joining two vertices can be mapped to a loop).

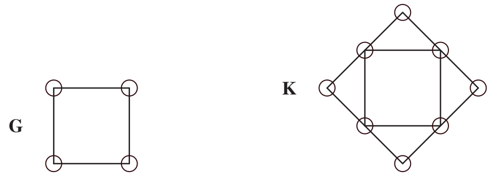


FIGURE 11.11: **G** is retract of **K**.

The central question about retractions is to decide, for two given objects **P** and **Q**, whether or not **P** is a retract of **Q**. A related question is to decide whether a given morphism $f : \mathbf{P} \longrightarrow \mathbf{Q}$ has a companion $g : \mathbf{Q} \longrightarrow \mathbf{P}$ such that $g \circ f = \mathrm{id}_{\mathbf{P}}$; if this is the case, f is said to be *coretraction* and its companion is a *retraction*. In fact, these questions are still largely unsolved, even for very simple categories like those of posets and graphs. Neverthlesss a fruitful approach to a solution of said problems is as follows:
Identify a general property, say (p), that the coretractions enjoy in the category considered; for example, in the above category of posets each coretraction is an order-embeding (that is a map f such that $x \leq y$ is equivalent to $f(x) \leq f(y)$). Now looking at (p) as an *approximation* of the coretractions, characterize the objects **P** for which this approximation is accurate, that is, the objects for which every morphism of source **P** and with property (p), is a coretraction. These objects are commonly called the *absolute retracts* (briefly *AR*) (a terminology not perfectly adequate, since these objects depend upon the approximation, but commonly used in the field), we will rather say *AR* with respect to the approximation (p). In the category of metric spaces with nonexpansive mappings these observations lead to the following definitions:

11.4.1 Retraction, coretraction, absolute retract

Let **E** and **F** be two metric spaces over a Heyting algebra $\mathscr{H}$. The space **E** is said to be a *retract* of **F**, denoted as $\mathbf{E} \lhd \mathbf{F}$, if there are nonexpansive maps $f : \mathbf{E} \to \mathbf{F}$ and $g : \mathbf{F} \to \mathbf{E}$ such that $g \circ f = \mathrm{id}_{\mathbf{E}}$. If this is the case, f is said to be *coretraction* and g a *retraction*. If **E** is a subspace of **F**, then **E** is a retract of **F** if there is a nonexpansive map from **F** to **E** such that $g(x) = x$ for all $x \in E$, where E is the domain of **E**. We can easily see that every coretraction is an isometry. A metric space is an *absolute retract* if it is a retract of every isometric extension.

11.4.1.1 Injectivity and extension property

A metric space $\mathbf{E}$ is said to be *injective* if for all spaces $\mathbf{F}$ and $\mathbf{E}'$, each nonexpansive mapping $f : \mathbf{F} \to \mathbf{E}$, and every isometry $g : \mathbf{F} \to \mathbf{E}'$ there is a nonexpansive mapping $h : \mathbf{E}' \to \mathbf{E}$ such that $h \circ g = f$.

A metric space $\mathbf{E}$ has the *one-point extension property* if for every space $\mathbf{E}' := (E', d')$ and every subset F of E', every nonexpansive map $f : \mathbf{E}'_{\restriction F} \to \mathbf{E}$ extends, for some $x' \in E' \setminus F$ (if any), to a nonexpansive map from $\mathbf{E}'_{\restriction F \cup \{x'\}}$ into $\mathbf{E}$.

Using Zorn's lemma one has immediately:

Lemma 11.7. *A metric space $\mathbf{E} := (E, d)$ over $\mathscr{H}$ is injective iff it has the one-point extension property.*

Proof. Trivially, injectivity implies the one-point extension property. For the converse, let $\mathbf{E}' := (E'd')$, $F \subseteq E'$ and $f : F \to E$ be a nonexpansive map from $\mathbf{E}'_{\restriction F}$ into E. Consider the collection of all nonexpansive maps $f' : F' \to E$ that extend f. This collection of maps is inductive. From Zorn's lemma, it has a maximal element g. The domain F'' of g is E', otherwise, pick $x \in E' \setminus E''$; since $\mathbf{E}$ has the one-point extension, g would extend to x, a contradiction. $\square$

As it will become apparent in Theorem 11.3, we may replace the phrase "for some x'" by "every x'" in the definition above.

11.4.1.2 Hyperconvexity

We say that a space $\mathbf{E}$ is *hyperconvex* if the intersection of every family of balls $(B_{\mathbf{E}}(x_i, r_i))_{i \in I}$ is non-empty whenever $d(x_i, x_j) \leq r_i \oplus \overline{r_j}$ for all $i, j \in I$.

Hyperconvexity is equivalent to the conjunction of the following conditions:

1) *Convexity* : for all $x, y \in E$ and $p, q \in \mathscr{H}$ such that $d(x, y) \leq p \oplus q$, there is $z \in E$ such that $d(x, z) \leq p$ and $d(z, y) \leq q$.

2) The *2-Helly property*, also called the *2-ball intersection property* : The intersection of every set (or, equivalently, of every family) of balls is non-empty, provided that their pairwise intersections are all non-empty.

11.4.2 A description of hyperconvex metric spaces

As it is easy to see, the collection of hyperconvex spaces over a Heyting algebra is stable under (non-empty) products and retracts. Thus, in the terminology of Duffus and Rival [30], this collection forms a variety. A less trivial property is this:

Theorem 11.2. *[47] The metric space* $\mathbf{H} := (\mathscr{H}, d_{\mathscr{H}})$ *is hyperconvex.*

Proof. We briefly sketch the proof; the interested reader might refer to [47] for details.

We show first that $\mathbf{H}$ is convex. Indeed, let $x, y \in \mathbf{H}$ and $p, q \in \mathscr{H}$ satisfy $d_{\mathscr{H}}(x,y) \leq p \oplus q$. Set $z := (x \oplus p) \vee (y \oplus \overline{q})$. It is easy to verify that $d_{\mathscr{H}}(x,z) \leq p$ and that $d_{\mathscr{H}}(z,y) \leq q$.

Next we tackle the fact that balls in $\mathbf{H}$ are intervals of $\mathscr{H}$. More precisely, any ball $B_{\mathbf{H}}(x,r)$ of $\mathbf{H}$ is the closed interval $[q,p] := \{y \in \mathscr{H} : q \leq y \leq r\}$ where $q := \bigwedge B_{\mathbf{H}}(x,r)$ and $p := \bigvee B_{\mathbf{H}}(x,r)$.

To conclude the proof, we observe that the closed intervals of a complete lattice have the 2-Helly property. $\square$

In what follows, we recall the notions of metric forms .

Let $\mathbf{E} := (E,d)$ be a metric space over a Heyting algebra $\mathscr{H}$. A *weak metric form* is any map $f : E \longrightarrow \mathscr{H}$ satisfying

$$d(x,y) \leq f(x) \oplus \overline{f(y)}, \tag{11.13}$$

for all $x, y \in E$.

This is a *metric form* if it is a weak metric form satisfying:

$$f(x) \leq d(x,y) \oplus f(y) \tag{11.14}$$

for all $x, y \in E$.

We denote by $\mathscr{C}(\mathbf{E})$, $(\mathscr{L}(\mathbf{E}))$, the set of weak metric forms, (metric forms) over $\mathbf{E}$. We equip these sets by the distance induced by the sup-distance on the power $\mathbf{H}^E$.

Lemma 11.8. *Let* $\mathbf{E} := (E,d)$ *be a metric space over* $\mathscr{H}$*, and* $f : E \to \mathscr{H}$*. The following properties are equivalent:*

(i) f *is a metric form;*

(ii) f *satisfies*

$$d_{\mathscr{H}}(d(x,y), f(x)) \leq f(y) \tag{11.15}$$

for all $x, y \in E$*;*

(iii) In the product space $\mathbf{H}^E$ *equipped with the "sup" distance,* $d(\bar{\delta}(y), f) = f(y)$ *for all* $y \in E$*;*

(iv) There is some isometric extension $\mathbf{E}' := (E', d')$ *of* $\mathbf{E}$ *and* $u \in E'$ *such that* $f(y) = d'(y,u)$ *for all* $y \in E$.

Proof. $(i) \Rightarrow (ii)$ According to the definition of the distance $d_{\mathscr{H}}$, conditions (11.13) and (11.14) amount to the inequality $d_{\mathscr{H}}(d(x,y), f(x)) \leq f(y)$, that is precisely condition (11.15).

$(ii) \Rightarrow (iii)$

According to formula (11.9):

$$d(\bar{\delta}(y), f) = \bigvee_{x \in E} d_{\mathscr{H}}(d(x,y), f(x)) \leq f(y).$$

Taking $x = y$, we get $d(x,y) = 0$, and $d_{\mathscr{H}}(0, f(y)) = f(y)$, thus the supremum in the inequality above is $f(y)$.

$(iii) \Rightarrow (iv)$ Since $\bar{\delta}$ is an isometric embedding from $\mathbf{E}$ into $\mathbf{H}^E$, it suffices to take $\mathbf{E}' := \mathbf{H}^E$ and $u := f$.

$(iv) \Rightarrow (i)$ This impliction is obvious from the triangle inequality. $\square$

Corollary 11.1. *The image of $\bar{\delta}$ is included into $\mathscr{L}(\mathbf{E})$, hence, $\bar{\delta}$ is an isometry from $\mathbf{E}$ into $\mathscr{L}(\mathbf{E})$.*

Proof. Let $u \in E$. We check that $\bar{\delta}(u)$ is a weak metric form for every $u \in E$. For that we show that inequality (11.15) holds with $f := \bar{\delta}(u)$. Indeed, we have $d_{\mathscr{H}}(d(x,y), \bar{\delta}(u)(x)) = d_{\mathscr{H}}(d(x,y), d(x,u)) \leq d(y,u) := \bar{\delta}(u)(y)$. $\square$

We recall Lemma II-4.4 of [47].

Lemma 11.9. *Let $\mathbf{E} := (E,d)$ be a metric space over $\mathscr{H}$. For every weak metric form f, the map $f_M : E \to \mathscr{H}$ defined by $f_M(x) := \bigwedge\{d(x,y) \oplus f(y) : y \in E\}$ is the largest metric form below f and $\bigcap\{B(x, f(x)) : x \in E\} = \bigcap\{B(x, f_M(x)) : x \in E\}$. Furthermore, the map $f \mapsto f_M$ is a retraction from $\mathscr{C}(\mathbf{E})$ onto $\mathscr{L}(\mathbf{E})$.*

Proof. The verification is straightforward (the difficulty lies in finding the precise formulation).

One proves first that if $g \in \mathscr{L}(\mathbf{E})$ and $g \leq f$ then $g \leq f_M$. Indeed, it follows from the fact that g is a metric form, that for every $x, y \in E$, one has $g(x) \leq d(x,y) \oplus g(y)$ and since $g \leq f$, one has $g(y) \leq f(y)$. Thus $g(x) \leq d(x,y) \oplus f(y)$, from which it follows that $g(x) \leq \bigwedge\{d(x,y) \oplus f(y) : y \in E\} =: f_M(x)$.

Next, one proves that f_M is a metric form, that is, that $d(x,y) \leq f_M(x) \oplus \overline{f_M(y)}$ and that $f_M(x) \leq d(x,y) \oplus f_M(y)$ for all $x, y \in E$. The right hand side of the first inequality $f_M(x) \oplus \overline{f_M(y)}$ is equal to $\bigwedge\{d(x,z) \oplus f(z) : z \in E\} \oplus \bigwedge\{\overline{f(t)} \oplus d(t,y) : t \in E\}$. Using the distributivity condition on $\mathscr{H}$, this yields $\bigwedge\{d(x,z) \oplus f(z) \oplus \overline{f(t)} \oplus d(t,y) : z, t \in E\}$. From the triangle inequality and the fact that $f(z) \oplus \overline{f(t)} \geq d(z,t)$, one gets $d(x,z) \oplus f(z) \oplus \overline{f(t)} \oplus d(t,y) \geq d(x,y)$, hence $f_M(x) \oplus \overline{f_M(y)} \geq d(x,y)$. For the second inequality, one observes that $d(x,z) \oplus f(z) \leq d(x,y) \oplus d(y,z) \oplus f(z)$ for all $z \in E$, hence

$f_M(x) := \bigwedge\{d(x,z) \oplus f(z) : z \in E\} \le \bigwedge\{d(x,y) \oplus d(y,z) \oplus f(z) : z \in E\} = d(x,y) \oplus \{\bigwedge d(y,z) \oplus f(z) : z \in E\} =: f_M(z)$.

These two facts yield that f_M is the largest metric form below f.

For the equality of the intersections of balls, note that the inclusion $\bigcap\{B(x, f_M(x)) : x \in E\} \subseteq \bigcap\{B(x, f(x)) : x \in E\}$ follows immediately from the fact that $f_M \le f$. For the reverse inclusion, pick $t \in \bigcap\{B(x, f(x)) : x \in E\}$, that is $\overline{\delta}(t)(x) = d(x,t) \le f(x)$, for every $x \in E$ or, equivalently, $\overline{\delta}(t) \le f$. Since $\overline{\delta}(t)$ is a metric form and f_M is the largest metric form below f, it is clear that $\overline{\delta}(t) \le f_M$, which yields $t \in \bigcap\{B(x, f_M(x)) : x \in E\}$.

Finally, one checks that the map $f \mapsto f_M$ is a retraction from $\mathscr{C}(\mathbf{E})$ onto $\mathscr{L}(\mathbf{E})$.

Since f_M is the largest metric form below f, this map fixes $\mathscr{L}(\mathbf{E})$ pointwise. To conclude, it suffices to prove that this map is nonexpansive, i.e., that $d(f_M, g_M) \le d(f,g)$, for all $f, g \in \mathscr{C}(\mathbf{E})$. Let $f, g \in \mathscr{C}(\mathbf{E})$. By definition of the distance on $\mathscr{C}(\mathbf{E})$, one has $f(y) \le g(y) \oplus \overline{d(f,g)}$, hence $d(x,y) \oplus f(y) \le d(x,y) \oplus g(y) \oplus \overline{d(f,g)}$, for all $x, y \in E$. This yields $f_M(x) := \bigwedge\{d(x,y) \oplus f(y) : y \in E\} \le \bigwedge\{d(x,y) \oplus g(y) \oplus \overline{d(f,g)} : y \in E\} = \bigwedge\{d(x,y) \oplus g(y) : y \in E\} \oplus \overline{d(f,g)} =: g_M(x) \oplus \overline{d(f,g)}$, that is $f_M(x) \le g_M(x) \oplus \overline{d(f,g)}$. The same argument shows that $g_M(x) \le f_M(x) \oplus \overline{d(f,g)}$. Consequently, $d_{\mathscr{H}}(f_M(x), g_M(x)) \le d(f,g)$. Since the latter holds for every $x \in E$, it follows that $d(f_M, g_M) \le d(f,g)$, as required. The proof of the lemma is then complete. □

Lemma 11.9 was obtained independently by Katětov [58]. It plays a key role in the description of hyperconvex spaces, of injective envelopes and of hole-preserving maps.

We next state the following hyperconvexity test.

Proposition 11.1. *Let* $\mathbf{E} := (E, d)$ *be a metric space over a Heyting algebra* $\mathscr{H}$. *The following properties are equivalent:*

(i) $\mathbf{E}$ *is hyperconvex;*

(ii) For every weak metric form $f : E \to \mathscr{H}$, *the intersection of balls* $B(x, f(x))$ *is non-empty;*

(iii) For every isometric extension $\mathbf{E}' := (E', d')$ *of* $\mathbf{E}$ *and every* $u \in E' \setminus E$, *there is a retraction of* $\mathbf{E}'_{\restriction E \cup \{u\}}$ *onto* $\mathbf{E}$.

Proof. $(iii) \Rightarrow (ii)$ Let $f : E \to \mathscr{H}$ be a weak metric form and f_M be the largest metric form below f given by Lemma 11.9. According to Corollary 11.1, $\bar{\delta}$ is an isometry of $\mathbf{E}$ into $\mathscr{L}(\mathbf{E})$. Thus, setting $\mathbf{E}' := \mathscr{L}(\mathbf{E})$, we may view $\mathbf{E}'$ as an isometric extension of $\mathbf{E}$. Since f_M is a metric form, Lemma 11.8 ensures that $d_{\mathbf{E}'}(\bar{\delta}(y), f_M) = f_M(y)$ for all $y \in E$. Thus

$f_M \in \bigcap_{x\in E} B_{\mathbf{E}'}(\bar{\delta}(x), f_M(x))$. Any retraction of $\mathbf{E}'_{\restriction \bar{\delta}(E)\cup\{u\}}$ onto $\mathbf{E}$ will send f_M into $\bigcap_{x\in E} B_{\mathbf{E}}(x, f_M(x))$. According to Lemma 11.9, this intersection is $\bigcap_{x\in E} B_{\mathbf{E}}(x, f(x))$.

$(ii) \Rightarrow (i)$ Let $\big(B(x_i, r_i)\big)_{i\in I}$ be a family of balls of $\mathbf{E}$ such that

$$d(x_i, x_j) \leq r_i \oplus \overline{r_j} \tag{11.16}$$

for all $i, j \in I$.

Define $f : E \to \mathscr{H}$ as follows: for each $x \in E$, set $f(x) = \bigwedge_{i\in I, x_i = x} r_i$. The distributivity condition on $\mathscr{H}$ ensures that

$$d(x,y) \leq f(x) \oplus \overline{f(y)},$$

for all $x, y \in E$. Hence f is a weak metric form. It follows that:

$$\emptyset \neq \bigcap_{x\in E} B\big(x, f(x)\big) \subseteq \bigcap_{i\in I} B(x_i, r_i).$$

$(i) \Rightarrow (iii)$ Let $\mathbf{E}' := (E', d')$ be an isometric extension of $\mathbf{E}$ and $u \in E' \setminus E$. For all $x, y \in E$, we have $d(x,y) = d'(x,y) \leq d'(x,u) \oplus d'(u,y)$. Since $\mathbf{E}$ is hyperconvex, the set $\bigcap_{x\in E} B\Big(x, d'(x,u)\Big)$ is non-empty. Let u' be an arbitrary element of this intersection. The map $g : E \cup \{u\} \to E$ defined by $g(x) = x$, for every $x \in E$ with $g(u) = u'$ is a retraction. $\square$

We conclude this paragraph with a characterization theorem:

Theorem 11.3. *[47] Let $\mathscr{H}$ be an Heyting algebra. Then, for a metric space $\mathbf{E} := (E, d)$ over $\mathscr{H}$, the following conditions are equivalent:*

(i) $\mathbf{E}$ *is an absolute retract;*

(ii) $\mathbf{E}$ *is injective;*

(iii) $\mathbf{E}$ *is hyperconvex;*

(iv) $\mathbf{E}$ *is a retract of a power of* $\mathscr{H}$.

Proof. We sketch the proof and refer the reader to [47] for the details.

$(i) \Rightarrow (iv)$ According to theorem 11.1, the space $\mathbf{E}$ isometrically embeds into a power of $\mathbf{H} := (\mathscr{H}, d_{\mathscr{H}})$; since it is an absolute retract, it must be a retract of such a power.

$(iv) \Rightarrow (iii)$ The space $\mathbf{H}$ is hyperconvex and the class of hyperconvex spaces is closed under products and retracts, i.e, in our terminology, it constitutes a variety.

$(iii) \Rightarrow (ii)$ We prove that the one-point extension holds. Let $\mathbf{E}' :=$

(E', d'), $A' \subseteq E'$, $x' \in E' \setminus A'$ and $f : A \to E$ be a nonexpansive map. Let $\mathscr{B} := (B_{\mathbf{E}}(f(a'), d'(a', x')))_{a' \in A'}$. Since f is nonexpansive, this family of balls satisfies the hyperconvexity condition, namely

$$d(f(a'), f(a'')) \leq d'(a', a'') \leq d'(a', x') \oplus \overline{d'(a'', x')}.$$

Hence, $\mathscr{B}$ has non-empty intersection. Pick an element x of this intersection and set $f(x') := x$.

$(ii) \Rightarrow (i)$ Trivial. □

11.4.3 Injective envelope

A nonexpansive map $f : \mathbf{E} \longrightarrow \mathbf{E}'$ is said to be *essential* if for every nonexpansive map $g : \mathbf{E}' \longrightarrow \mathbf{E}''$, the map $g \circ f$ is an isometry if and only if g is isometry (note that, in particular, f is an isometry). An essential nonexpansive map f from $\mathbf{E}$ into an injective metric space $\mathbf{E}'$ over $\mathscr{H}$ is called an *injective envelope* of $\mathbf{E}$. We will rather say that $\mathbf{E}'$ is an *injective envelope* of $\mathbf{E}$. Indeed, this says in substance that an injective envelope of a metric space $\mathbf{E}$ is a minimal injective metric space over $\mathscr{H}$, containing $\mathbf{E}$ isometrically.

The construction of injective envelopes is based upon the notion of *minimal metric form*, a notion borrowed to Isbell [45] that he calls *extremal*.

Let us recall that a (weak) metric form is *minimal* if there is no other (weak) metric form g satisfying $g \leq f$ (that is $g(x) \leq f(x)$ for all $x \in E$). Since from Lemma 11.9, every weak metric form majorizes a metric form, the two notions of minimality coincide. Due to the distributivity condition and to the completeness of $\mathscr{H}$, a straightforward application of Zorn's lemma yields the existence of a minimal metric form , below any given weak metric form.

As shown in [47], (cf. also theorem 2.2 of [55]):

Theorem 11.4. *Every generalized metric space* $\mathbf{E}$ *over a Heyting algebra* $\mathscr{H}$ *has an injective envelope, namely the space* $\mathscr{N}(\mathbf{E})$ *of minimal metric forms.*

Proof. Let $\mathbf{E}$ be a metric space over the Heyting algebra $\mathscr{H}$.
One proves first that the space $\mathscr{L}(\mathbf{E})$ of metric forms is an absolute retract. This means that every isometric extension $\mathbf{E}' := (E', d')$ can be retracted on $\mathscr{L}(\mathbf{E})$. This is almost immediate. For every $u \in E'$, let $\varphi_u : E \to \mathscr{H}$ be defined by setting $\varphi_u(x) := d'(\bar{\delta}(x), u)$. Since the map $\bar{\delta} : E \to \mathscr{H}$ is an isometry, φ_u is a metric form. To conclude, one proves that the map $\varphi : u \mapsto \varphi_u$ is a retraction of $\mathbf{E}'$ on $\mathscr{L}(\mathbf{E})$. First, φ is the identity on $\mathscr{L}(\mathbf{E})$. Indeed, if $f \in \mathscr{L}(\mathbf{E})$, then, according to (iii) of Lemma 11.8, $\varphi_f(x) = d(\bar{\delta}(x), f) = f(x)$ for every $x \in E$, hence $\varphi_f = f$. Next, φ is nonexpansive, that is, $d(\varphi_u, \varphi_v) \leq d'(u, v)$ for all $u, v \in \mathbf{E}'$. It follows from the

triangle inequality that

$$d'(\bar{\delta}(x),u) \leq d'(\bar{\delta}(x),v) \oplus d'(v,u) \tag{11.17}$$

and that

$$d'(\bar{\delta}(x),v) \leq d'(\bar{\delta}(x),u) \oplus d'(u,v), \tag{11.18}$$

for every $x \in E$.
These inequalities translate to $\varphi_u(x) \leq \varphi_v(x) \oplus \overline{d'(u,v)}$ and $\varphi_v(x) \leq \varphi_u(x) \oplus d'(u,v)$, that is, to $d_{\mathscr{H}}(\varphi_u(x), \varphi_v(x)) \leq d'(u,v)$. This yields $d(\varphi_u, \varphi_v) := \bigvee_{x \in E} d_{\mathscr{H}}(\varphi_u(x), \varphi_v(x)) \leq d'(u,v)$, as required.

Next, one proves that the space $\mathscr{N}(\mathbf{E})$ of minimal metric forms over $\mathbf{E}$ is hyperconvex. According to *(iii)* of Proposition 11.1, this amounts to proving that for every isometric extension $\mathbf{E}' := (E', d')$ of $\mathscr{N}(\mathbf{E})$ and every $u \in E' \setminus \mathscr{N}(\mathbf{E})$, there is a retraction of $\mathbf{E}'_{\restriction \mathscr{N}(\mathbf{E}) \cup \{u\}}$ onto $\mathscr{N}(\mathbf{E})$. This reduces to the fact that the intersection of balls $A := \bigcap_{f \in \mathscr{N}(\mathbf{E})} B_{\mathbf{E}'}(f, d'(f,u))$ contains some element $\tilde{u}$ of $\mathscr{N}(\mathbf{E})$. Let $\varphi_u : \mathbf{E} \to \mathscr{H}$ be defined by setting $\varphi_u(x) := d'(\bar{\delta}(x), u)$. As illustrated above, this is a metric form on $\mathbf{E}$. Let $\tilde{u}$ be a minimal metric form on $\mathbf{E}$ below u. Let $\phi : \bar{\delta}(E) \cup \{u\} \to \mathscr{L}(\mathbf{E})$ be the nonexpansive map sending u to $\tilde{u}$ and leaving every other element fixed. Since $\mathscr{L}(\mathbf{E})$ is an absolute retract, it is injective, hence ϕ extends to a nonexpansive map Φ from $\mathbf{E}'$ into $\mathscr{L}(\mathbf{E})$. This map is the identity on $\mathscr{N}(\mathbf{E})$. Indeed, let $f \in \mathscr{N}(\mathbf{E})$. Since Φ is nonexpansive, we have $d(\bar{\delta}(x), \Phi(f)) \leq d'(\bar{\delta}(x), f)$, for every $x \in E$; in other words, $\Phi(f)(x) \leq f(x)$. Since f is minimal, it necessarily follows that $\Phi(f) = f$. This yields that $d(f, \tilde{u}) = d(\Phi(f), \Phi(u) \leq d'(f,u)$, for every $f \in \mathscr{N}(\mathbf{E})$. This proves that $\tilde{u}$ belongs to A. Hence $\mathscr{N}(\mathbf{E})$ is hyperconvex. By virtue of Theorem 11.3 it is injective. If $\mathbf{E}'$ is an injective space between $\mathbf{E}$ and $\mathscr{N}(\mathbf{E})$, then the identity map id on $\mathbf{E}$ extends to a nonexpansive map Φ from $\mathscr{N}(\mathbf{E})$ into $\mathbf{E}'$. As above, for every $f \in \mathscr{N}(\mathbf{E})$, we have $\Phi(f) \leq f$, hence $\Phi(f) = f$, since f is minimal. It follows that $\mathbf{E}' = \mathscr{N}(\mathbf{E})$. This proves that $\mathscr{N}(\mathbf{E})$ is a minimal injective metric space containing $\mathbf{E}$. □

A particularly useful fact is the following:

Lemma 11.10. *If a nonexpansive map from an injective envelope of* $\mathbf{E} := (E, d)$ *into itself fixes E pointwise, then it is the identity map.*

Note that two injective envelopes of $\mathbf{E}$ are isomorphic via the identity over $\mathbf{E}$. This justifies the use of the expression "the" injective envelope of $\mathbf{E}$. A particular injective envelope of $\mathbf{E}$, as $\mathscr{N}(\mathbf{E})$, will be called a *representation* of the injective envelope.

We describe the injective envelopes of two-element metric spaces (see [55] for the proofs of the statements below). Let $\mathscr{H}$ be a Heyting algebra and $v \in \mathscr{H}$. Let $\mathbf{E} := (\{x,y\},d)$ be a two-element metric space over $\mathscr{H}$ such that $d(x,y) = v$. We denote the injective envelope of $\mathbf{E}$ by $\tilde{\mathscr{N}_v}$. We give two representations of it. Let $\mathscr{C}_v$ be the set of all pairs $(u_1,u_2) \in \mathscr{H}^2$ such that $v \leq u_1 \oplus \overline{u_2}$. Equip this set with the ordering induced by the product ordering on $\mathscr{H}^2$ and denote the set of its minimal elements by $\mathscr{N}_v$. Each element of $\mathscr{N}_v$ defines a minimal metric form. We equip $\mathscr{H}^2$ with the supremum distance, namely:

$$d_{\mathscr{H}^2}\left((u_1,u_2),(u_1',u_2')\right) := d_{\mathscr{H}}(u_1,u_1') \vee d_{\mathscr{H}}(u_2,u_2').$$

Let $v \in \mathscr{H}$ and set $\mathscr{S}_v := \{\lceil v - \beta \rceil : \beta \in \mathscr{H}\}$; equipped with the ordering induced by the ordering over $\mathscr{H}$ this is a complete lattice. According to lemma 2.5 of [55], $(x_1,x_2) \in \mathscr{N}_v$ iff $x_1 = \lceil v - x_2 \rceil$ and $\overline{x_2} = \lceil -x_1 \oplus v \rceil$. This yields a correspondence between $\mathscr{N}_v$ and $\mathscr{S}_v$.

Lemma 11.11. *(Lemma 2.3, Proposition 2.7 of [55]) The space $\mathscr{N}_v$ equipped with the supremum distance and the set $\mathscr{S}_v$ equipped with the distance induced by the distance over $\mathscr{H}$ are injective envelopes of the two-element metric spaces $\{(0,v),(v,0)\}$ and $\{0,v\}$ respectively. These spaces are isometric to the injective envelope of* $\mathbf{E} := (\{x,y\},d)$*, where $d(x,y) = v$.*

We refer the reader to[55] and in [57] for further details, in particular, for a presentation in terms of Galois correspondences. An illustration is given in Section 11.7.

11.4.4 Hole-preserving maps

In this Subsection, we introduce the notions of hole-preserving maps. A large part of our account is borrowed from subsection II-4 of [47].

Let $\mathbf{E}$ and $\mathbf{F}$ be two metric spaces over a Heyting algebra $\mathscr{H}$. If f is a nonexpansive map from $\mathbf{F}$ into $\mathbf{E}$ and h is a map from F into $\mathscr{H}$, the *image* of h is the map h_f from E into $\mathscr{H}$ defined by $h_f(x) : \bigwedge\{h(y) : f(y) = x\}$ (in particular $h_f(x) = 1$ where 1 is the largest element of $\mathscr{H}$ for every x not in the range of f). A *hole* of $\mathbf{F}$ is any map $h : F \to \mathscr{H}$ such that the intersection of balls $B(x,h(y))$ of F $(x \in F)$ is empty. If h is a hole of $\mathbf{F}$, the map f *preserves* h provided that h_f is a hole of $\mathbf{E}$. The map f is *hole-preserving* if the image of every hole is a hole.

As it is easy to see, coretractions preserve holes and hole-preserving maps are isometries. One may then use hole-preserving maps as approximations of coretractions

We recall the following result of [47].

Theorem 11.5. *On an involutive Heyting algebra $\mathscr{H}$, the absolute retracts and the injectives wit respect to hole-preserving maps coincide. The class of* *these objects is closed under products and retractions. Moreover, every metric space embeds into some member of this class by some hole-preserving map.*

The proof relies on the introduction of the replete space $\mathscr{H}(\mathbf{E})$ of a generalized metric space $\mathbf{E}$. The space $\mathbf{E}$ is an absolute retract (with respect to the hole-preserving maps), or not, depending on whether or not $\mathbf{E}$ is a retract of $\mathscr{H}(\mathbf{E})$. Furthermore, with the existence of the replete space one may prove the *transferability* of hole-preserving maps (Lemma II-4.6 of [47]), that is the fact that for every nonexpansive map $f : \mathbf{F} \to \mathbf{E}$, and every hole-preserving map $g : \mathbf{F} \to \mathbf{G}$, there are a hole-preserving map $g' : \mathbf{G} \to \mathbf{E}'$ and a nonexpansive map $f' : \mathbf{G} \to \mathbf{E}'$ such that $g' \circ f = f' \circ g$. Indeed, one may choose $\mathbf{E}' = \mathscr{H}(\mathbf{E})$. As it is well known among categorists, the transferability property implies that absolute retracts and injective objects coincide [63].

In the sequel we define the replete space and present the proof of the transferability property.

Proofs are borrowed from [47].

$$
\begin{tikzcd}
F \arrow[r, Rightarrow, "g"] \arrow[d, "f"'] & G \arrow[d, dashed, "f'"] \\
E \arrow[r, dashed, "g'"'] & E'
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
F \arrow[r, Rightarrow, "g"] \arrow[d, "f"'] & G \arrow[d, dashed, "\bar{f}"] \\
E \arrow[r, dashed, "\bar{\delta}"'] & \mathscr{H}(E)
\end{tikzcd}
$$

FIGURE 11.12: Transferability.

Let $\mathscr{H}(\mathbf{E})$ be the subset of $\mathscr{L}(E)$ consisting of metric forms h such that the intersection of balls $B(x, h(x))$ for $x \in E$ is nonempty. If $\mathscr{H}$ is a Heyting algebra, we may equip $\mathscr{H}(\mathbf{E})$ with the distance induced by the sup-distance on $\mathbf{H}^E$. We call it the *replete space*.

We recall the following two results [47].

Lemma 11.12. *(see Lemma II-4.3 p. 195) If $\mathbf{E} := (E, d)$ is a metric space over a Heyting algebra $\mathscr{H}$, then $\overline{\delta} : \mathbf{E} \to \mathscr{H}(\mathbf{E})$, defined by $\overline{\delta}(x)(y) := d(y, x)$ is a hole-preserving map from $\mathbf{E}$ into $\mathscr{H}(\mathbf{E})$. Furthermore $\mathscr{H}(\mathbf{E})$ is an absolute retract with respect to the hole-preserving maps (i.e., this is a retract of every extension by a hole-preserving map).*

Proof. The proof of this lemma is almost immediate. We just indicate that $\mathscr{H}(\mathbf{E})$ is an absolute retract. Let $\mathbf{E}' := (E', d')$ be a hole-preserving extension of $\mathscr{H}(\mathbf{E})$. For $u \in \mathbf{E}'$, set $\tilde{u} : E \to \mathscr{H}$, defined by setting $\tilde{u}(x) :=$

$d'(\overline{\delta}(x),u)$, for all $x \in E$. By construction, $\tilde{u}$ is a metric form; moreover, it belongs to $\mathscr{H}(\mathbf{E})$. To conclude, observe that the map $u \mapsto \tilde{u}$ is a retraction. $\square$

Lemma 11.13. *(see Lemma II-4.5 p. 196) If* $\mathbf{E} := (E,d)$ *and* $\mathbf{F}$ *are two metric spaces over a Heyting algebra* $\mathscr{H}$*, then every nonexpansive map* $f : \mathbf{F} \to \mathbf{E}$ *extends to a nonexpansive map* $\mathscr{H}_f : \mathscr{H}(\mathbf{F}) \to \mathscr{H}(\mathbf{E})$.

Proof. The proof relies on Lemma 11.9. We define $\tilde{f} : \mathbf{H}^F \to \mathbf{H}^E$ by setting $\tilde{f}(h) := h_f$, where h_f is the map from E into $\mathscr{H}$ defined by $h_f(x) := \bigwedge\{h(y) : f(y) = x\}$). One checks first that this map is nonexpansive and next that if $h \in \mathscr{C}(\mathbf{F})$, then $\tilde{f}(h) \in \mathscr{C}(\mathbf{E})$. For $k \in \mathscr{C}(E)$, let k_M be the largest metric form below k given by Lemma 11.9. Let r be the retraction from $\mathscr{C}(\mathbf{E})$ onto $\mathscr{L}(\mathbf{E})$ defined by setting $r(k) := k_M$, for all $k \in \mathscr{C}(E)$. The composition $r \circ \tilde{f} : \mathscr{C}(\mathbf{F}) \to \mathscr{L}(\mathbf{E})$ is nonexpansive, for it is a composition of nonexpansive maps. It extends f once F and E are identified with their images $\overline{\delta}(F))$ and $\overline{\delta}(E)$, that is $(r \circ \tilde{f})(\overline{\delta}(y)) = \overline{\delta}(f(y))$ for all $y \in F$. Indeed, observe first that $\tilde{f}(\overline{\delta}(y))(f(y)) = 0$. Next, since by definition of r, $r(\tilde{f}(\overline{\delta}(y))) \leq \tilde{f}(\overline{\delta}(y))$, one has $(r \circ \tilde{f})(\overline{\delta}(y)(f(y)) = 0$. Since $r \circ \tilde{f}(\overline{\delta}(y))$ is a metric form, it necessarily follows that $r \circ \tilde{f}(\overline{\delta}(y)) = \overline{\delta}(f(y))$ (indeed, $d(\overline{\delta}(y), r \circ \tilde{f}(\overline{\delta}(y))) = 0$). Finally, by Lemma 11.9, we have $\bigcap\{B(y,v(y)) : y \in F\} = \bigcap\{B(x,\tilde{f}(v)(x) : x \in E\} = \bigcap\{B(x,(r \circ \tilde{f})(v)(x) : x \in E\}$. Consequently, $r \circ \tilde{f}(v) \in \mathscr{H}(\mathbf{E})$ for every $v \in \mathscr{H}(\mathbf{F})$. The restriction $\mathscr{H}_f$ of $r \circ \tilde{f}$ to $\mathscr{H}(\mathbf{F})$ has the required property. $\square$

Lemma 11.14. *(see Lemma II-4.6 p. 197) The hole-preserving maps are transferable.*

Proof. Let $f : \mathbf{F} \to \mathbf{E}$ be a nonexpansive map and $g : \mathbf{F} \to \mathbf{G}$ be a hole-preserving map. As above, denote by $\overline{\delta}$ the map from $\mathbf{E}$ into $\mathscr{H}(\mathbf{E})$ defined by $\overline{\delta}(x) := d(z,x)$ for $z \in E$. We define $\hat{f} : \mathbf{G} \to \mathscr{H}(\mathbf{E})$ in such a way that $\hat{f} \circ g = \overline{\delta} \circ f$.

For this purpose, define a nonexpansive map $\mathscr{I}_g : \mathbf{G} \to \mathscr{H}$ as follows. For every $u \in G$, set $\hat{u} : \mathbf{F} \to \mathscr{H}$, given by $\hat{u}(y) := (d(g(y),u)$ and set $\mathscr{J}(u) := \hat{u}$. We check successively that the map $\hat{u}$ belongs to $\mathscr{H}(\mathbf{F})$ (indeed, $u \in \bigcap_{y \in F} B(g(y), d(g(y),u))$; since g is hole-preserving, $\bigcap_{y \in F} B(y, d(g(y),u)) = \bigcap_{y \in F} B(y, \hat{u}(y))$ is non empty), hence $\mathscr{I}_g$ maps $\mathbf{G}$ into $\mathscr{H}$. Next, $\mathscr{I}_g$ is nonexpansive and finally $\mathscr{I}_g(g(y)) = \overline{\delta}(y)$, for every $y \in F$ (since g is hole-preserving, it is an isometry, thus $\mathscr{I}_g(g(y))(z) = d(g(z), g(x)) = d(z,y) = \overline{\delta}(y)(z)$, for every $z \in F$). Set $\hat{f} := \mathscr{H}_f \circ \mathscr{J}_g$, where $\mathscr{H}_f$ is given by Lemma 11.13. Then $\hat{f} \circ g = \overline{\delta} \circ f$. This proves that f is transferable. $\square$

11.4.4.1 Hole-preserving maps and one-local retracts

In his study of the fixed point property, Khamsi [60] introduced the notion of one-local retracts. This notion, defined for ordinary metric spaces, extends to metric spaces over a Heyting algebra. In fact, it extends to metric spaces over an ordered monoid equipped with an involution and, more generally, to binary structures that are reflexive and involutive in the sense of [61]. One-local retracts play a crucial role in the fixed point theorem presented in the next Section. In the sequel, unless otherwise stated, we do not suppose that $\mathscr{H}$ satisfies the distributivity condition.

Let $\mathbf{E} := (E,d)$ be a metric space over $\mathscr{H}$ and A be a subset of E. We say that $\mathbf{E}_{\restriction A} := (A, d_{\restriction A})$ is a *one-local retract* of $\mathbf{E}$ if it is a retract of $\mathbf{E}_{\restriction A\cup\{x\}} := (A\cup\{x\}, d_{\restriction A\cup\{x\}})$ (via the identity map) for every $x \in E$.

Lemma 11.15. *Let $\mathbf{E} := (E,d)$ be a metric space over $\mathscr{H}$ and A be a subset of E. Then $\mathbf{E}_{\restriction A}$ is a one-local retract of $\mathbf{E}$ iff for every family of balls $\Big(B(x_i,r_i)\Big)_{i\in I}$, with $x_i \in A$, $r_i \in \mathscr{H}$ for any $i \in I$, such that $\bigcap\limits_{i\in I} B_{\mathbf{E}}(x_i,r_i)$ is not empty, the intersection $\bigcap\limits_{i\in I} B_{\mathbf{E}}(x_i,r_i)\cap A$ is not empty.*

Proof. Suppose that $\mathbf{E}_{\restriction A}$ is a one-local retract of $\mathbf{E}$. Let I be a set. Consider a family of balls $\Big(B_{\mathbf{E}}(x_i,r_i)\Big)_{i\in I}$, with $x_i \in A$, $r_i \in \mathscr{H}$ for any $i \in I$, such that $B = \bigcap\limits_{i\in I} B_{\mathbf{E}}(x_i,r_i)$ is not empty. Let $a \in B$ and let h be a retraction from $\mathbf{E}_{\restriction A\cup\{a\}}$ onto $\mathbf{E}_{\restriction A}$. Set $a' := h(a)$. Since h fixes A and retracts a onto a', $a' \in B_{\mathbf{E}}(x_i,r_i)$, hence $a' \in \bigcap\limits_{i\in I} B_{\mathbf{E}}(x_i,r_i)\cap A$. Conversely, ones proves that $\mathbf{E}_{\restriction A}$ is a one-local retract provided that the intersection property of balls is satisfied. Let $a \in E\setminus A$ and define

$$\mathscr{B} := \{B(u,r) :\ u \in A,\ a \in B(u,r) \text{ and } r \in \mathscr{H}\}.$$

Set $B := \bigcap \mathscr{B}$. Then $a \in B$, which implies $B \neq \emptyset$. According to the ball's property, $B\cap A \neq \emptyset$. Let $a' \in B\cap A$. The map $h : A\cap\{a\} \to A$, which is the identity on A and satisfies $h(a) = a'$, is a retraction of $\mathbf{E}_{\restriction A\cup\{a\}}$.

□

Lemma 11.16. *Let $\mathbf{E}$ and $\mathbf{E}'$ be two metric spaces over $\mathscr{H}$. A nonexpansive map f from $\mathbf{E}$ into $\mathbf{E}'$ is hole-preserving iff f is an isometry of $\mathbf{E}$ onto its image and this image is a one-local retract of $\mathbf{E}'$.*

The routine proof is based on Lemma 11.15. We omit it.

11.5 Fixed point property

A central result in the category of ordinary metric spaces endowed with nonexpansive maps is the Sine-Soardi's fixed point theorem [88, 89]. This theorem asserts that every nonexpansive map on a bounded, hyperconvex metric space has a fixed point.

This result was generalized in two directions. First, Penot [73] introduced the notion of space endowed with a compact normal structure, extending the notion of bounded hyperconvex space. With this notion, Kirk's theorem [62] amounted to the fact that every nonexpansive map on a space endowed with a compact normal structure has a fixed point. The existence of a common fixed point for a commuting set of nonexpansive maps was considered by several authors (see [20, 28, 65]). In 1986, Baillon [5], extending the theorem of Sine-Soardi, proved that every set of nonexpansive maps which commute on a bounded hyperconvex space, has a common fixed point. Khamsi [60] extended this result to metric spaces endowed with a compact and normal structure. In [47] the theorem of Sine-Soardi was extended to bounded hyperconvex spaces over a Heyting algebra, for an appropriate notion of boundedness. The possible extension to commuting sets of nonexpansive maps remained open (only the case of a countable set was settled). In [61] the notion of compact normal structure for metric spaces over Heyting algebras (and more generally for systems of binary relations) was introduced and Khamsi's theorem was extended to families of nonexpansive maps which commute on a space endowed with a compact and normal structure.

Here we present first the generalization of the theorem of Sine-Soardi to bounded hyperconvex spaces over a Heyting algebra. Next, we introduce the notion of compact and normal structure and provide a brief description of the result by Khamsi-Pouzet.

In the sequel we consider generalized metric spaces whose set of values $\mathscr{H}$ does not necessarily satisfy the distributivity condition. For these spaces, we define the notions of diameter, radius and of Chebyshev center.

Let $\mathbf{E} := (E,d)$ be a metric space over $\mathscr{H}$. We denote by $\mathscr{B}_{\mathbf{E}}$ the set of balls of $\mathbf{E}$. Let A be a nonempty subset of E and $r \in \mathscr{H}$. The r-center is the set $C_{\mathbf{E}}(A,r) := \{x \in E : A \subseteq B(x,r)\}$. Set $\mathrm{Cov}_{\mathbf{E}}(A) := \bigcap\{B \in \mathscr{B}_{\mathbf{E}} : A \subseteq B\}$. The *diameter* of A is $\bigvee\{d(x,y) : x,y \in A\}$. The *radius* $r(A)$ is $\bigwedge\{v \in \mathscr{H} : A \subseteq B(x,v)$ for some $x \in A\}$. A subset A of E is said to be *equally centered* if $\delta(A) = r(A)$.

11.5.1 The case of hyperconvex spaces

We suppose that $\mathscr{H}$ is a Heyting algebra. We define the notion of boundedness.

An element $v \in V$ is called *self-dual* if $\overline{v} = v$, it is said to be *accessible* if there is some $r \in V$ with $v \not\leq r$ and $v \leq r \oplus \overline{r}$ and *inaccessible* otherwise. Clearly, 0 is inaccessible; every inaccessible element v is self-dual (otherwise, $\overline{v}$ is incomparable to v and we may choose $r := \overline{v}$).

Definition 11.1. We say that a space (E,d) is *bounded* if 0 is the only inaccessible element below $\delta(E)$.

Lemma 11.17. *Let A be an intersection of balls of (E,d). If $\delta(A)$ is inacessible then A is equally centered; the converse holds if (E,d) is hyperconvex.*

Proof. Suppose that $v := \delta(A)$ is inaccessible. Let $r \in \mathscr{H}$ such that $A \subseteq B(x,r)$. This yields $d(a,b) \leq d(a,x) \oplus d(x,b) \leq \overline{r} \oplus r$ for every $a,b \in A$. Thus $v \leq \overline{r} \oplus r$. Since v is inacessible, $v \leq r$, hence $v \leq r(A)$. Thus $v = r(A)$. Suppose that A is equally centered. Let r be such that $v \leq r \oplus \overline{r}$. The balls $B(x,r)$ $(x \in A)$ intersect pairwise and intersect each of the balls whose intersection is A; since (E,d) is hyperconvex, these balls have a nonempty intersection. Any member a of this intersection is in A and satisfies $A \subseteq B(a,\overline{r})$. Since A is equally centered $r(A) = v$. Hence, $v \leq \overline{r}$. Since v is self-dual, $v \leq r$. Thus v is inaccessible. □

Lemma 11.18. *Let $\mathbf{E}$ be a non empty hyperconvex metric space over a Heyting algebra $\mathscr{H}$ and $f : \mathbf{E} \longrightarrow \mathbf{E}$ be a nonexpansive mapping. Then there is a non empty hyperconvex subspace $\mathbf{S}$ of $\mathbf{E}$ such that $f(S) \subseteq S$, and its diameter $\delta(S) = \vee\{d(x,y) : x,y \in S\}$ is inaccessible.*

For a proof see Lemma III-1.1 of [47]. The next Lemma follows immediately.

Lemma 11.19. *Let $\mathbf{E}$ be a non empty hyperconvex space. Then there is a nonempty, hyperconvex, invariant subspace $\mathbf{S}$, whose diameter is inaccessible.*

Theorem 11.6. *Let $\mathbf{E}$ be a nonempty, bounded, hyperconvex space. Then every nonexpansive map f has a fixed point. Moreover, the restriction of $\mathbf{E}$ to the set $Fix(f)$ of its fixed points is hyperconvex*

Proof. Since 0 is the unique inaccessible element below the diameter $\delta(E)$, the diameter of the non empty set S given by lemma 11.18 is 0, thus S must consist of a single element, fixed by f. Let $\{B_F(x_i,r_i) : i \in I\}$ be a family of balls of $Fix(f)$ with $d(x_i,x_j) \leq r_i + \overline{r}_j$ for all $i,j \in I$. Since $\mathbf{E}$ is hyperconvex, then $T = \cap\{B_F(x_i,r_i) : i \in I\} \neq \emptyset$ and, as any intersection of balls of an hyperconvex space, $\mathbf{E}_{\restriction T}$ is hyperconvex and clearly, bounded. Now, since f is nonexpansive and the x_i are fixed by f, we have $f(T) \subseteq T$. The f.p.p applied to T yields the existence of $x \in Fix(f) \cap T$. Thus, the above intersection is non empty and $Fix(f)$ is hyperconvex. □

Corollary 11.2. *Let* **E** *be a nonempty bounded hyperconvex space. Among the subspaces of* **E***, the retracts of E are the sets of fixed points of the nonexpansive maps from* **E** *into itself.*

Proof. If **A** is a retract of **E**, then $A = Fix(g)$ for every retraction. Conversely, it follows from the above result that the set $Fix(f)$ of fixed points of a map $f : E \longrightarrow E$ is hyperconvex. But hyperconvex set are absolute retracts, thus $Fix(f)$ is a retract. □

11.5.2 Compact and normal structure

We next extend the notion of compact and normal structure defined by Penot, [72, 73], for ordinary metric spaces. We consider generalized metric spaces over an involutive ordered algebra $\mathscr{H}$ which, unless otherwise stated, does not necessarily satisfy the distributivity condition. We say that a metric space **E** has a *compact structure* if the collection of balls of **E** has the finite intersection property (f.i.p.) and it has a *normal structure* if for every intersection of balls A, either $\delta(A) = 0$ or $r(A) < \delta(A)$. This condition amounts to the fact that the only equally centered intersections of balls are singletons.

Lemma 11.17in conjunction with the fact that the 2-Helly property implies that the collection of balls has the finite intersection property, yields:

Corollary 11.3. *If a generalized metric space* $\mathbf{E} := (E,d)$ *over a Heyting algebra is bounded and hyperconvex, then it has a compact normal structure.*

We denote by $\hat{\mathscr{B}}_{\mathbf{E}}$, the collection of intersections of balls and by $\hat{\mathscr{B}}^*_{\mathbf{E}}$ the set of the non empty ones.

Lemma 11.20. *Let* $\mathbf{E} := (E,d)$ *be metric space over* $\mathscr{H}$*. Let* f *be a nonexpansive map* **E***. If* **E** *has a compact structure, then every member of* $\hat{\mathscr{B}}^*_{\mathbf{E}}$ *preserved by* f *contains a minimal one. If* $A \in \hat{\mathscr{B}}^*_{\mathbf{E}}$ *is a minimal member preserved by* f*, then* $\mathrm{Cov}_{\mathbf{E}}(f(A)) = A$ *and* A *is equally centered.*

This lemma allows us to deduce Penot's formulation [72, 73] of Kirk's fixed point theorem [62] under our formulation.

Theorem 11.7. *Let* $\mathbf{E} := (E,d)$ *be a generalized metric space over* $\mathscr{H}$*. Assume that* **E** *has a compact normal structure. Then every nonexpansive map* f *on* **E** *has a fixed point.*

An easy consequence of Theorem 11.7 is the following beautiful structural result:

Proposition 11.2. *Let* $\mathbf{E} := (E,d)$ *be a a metric space over* $\mathscr{H}$ *with a compact normal structure. Let* f *be an endomorphism* $\mathbf{E}$*. Then the restriction* $\mathbf{E}_{\restriction Fix(f)}$*, to the set* $Fix(f)$ *of fixed points of* f*, has a compact normal structure.*

Proposition 11.2 will allow us to prove that a finite set of commuting endomorphism maps has a common fixed point and that the restriction of $\mathbf{E}$ to the set of common fixed points has a compact normal structure. Obviously one would like to know whether such a conclusion still holds for infinitely many maps. In order to settle this point, one has to carefully investigate the structure of the set of fixed points of an endomorphism. This will rely on the properties of one-local-retracts.

The next result is the most important one, as it shows that a one-local retract enjoys the same properties as those of the larger set.

Lemma 11.21. *Let* $\mathbf{E} := (E,d)$ *be a metric space over* $\mathscr{H}$*,* $X \subseteq E$ *be a nonempty subset. Assume that* $\mathbf{E}_{\restriction X}$ *is a one-local retract of* $\mathbf{E}$*. If* $\mathbf{E}$ *has both a compact and a normal structure, then* $\mathbf{E}_{\restriction X}$ *also has a normal compact structure.*

Proposition 11.3. *Let* $\mathbf{E} := (E,d)$ *be a metric space over* $\mathscr{H}$*. Assume that* $\mathbf{E}$ *has a compact normal structure. Then for every nonexpansive map* f *of* $\mathbf{E}$*, the set of fixed points* $Fix(f)$ *of* f *is a nonempty one-local retract of* $\mathbf{E}$*. Thus* $\mathbf{E}_{\restriction Fix(f)}$ *has a compact normal structure.*

Proof. Let I be a set. Consider a family of balls $\left(B_{\mathbf{E}}(x_i,r_i)\right)_{i\in I}$, with $x_i \in Fix(f)$ and $r_i \in \mathscr{E}$ for $i \in I$, such that $A = \bigcap_{i\in I} B_{\mathbf{E}}(x_i,r_i)$ is not empty. Since each x_i belongs to $Fix(f)$, f preserves A. Since A is an intersection of balls, Lemma 11.20 ensures that A contains an intersection of balls A' which is minimal, preserved by f, and equally centered. From the normality of $\mathbf{E}$, A' must consist of a single element, i.e., A' consists of a fixed point of f. Consequently, $A \cap Fix(f) \neq \emptyset$. According to Lemma 11.15, $Fix(f)$ is a one-local retract. □

In [61], Khamsi and Pouzet proved the following:

Theorem 11.8. *If a generalized metric space* $\mathbf{E} := (E,d)$ *has a compact normal structure, then every commuting family* $\mathscr{F}$ *of nonexpansive self maps has a common fixed point. Furthermore, the restriction of* $\mathbf{E}$ *to the set* $Fix(\mathscr{F})$ *of common fixed points of* $\mathscr{F}$*, is a one-local retract of* (E,d)*.*

The fact that a space has a compact structure is an infinistic property (any finite metric space enjoys it). A description of generalized metric spaces with a compact normal structure eludes us, even in the case of ordinary metric spaces.

From Theorem 11.8, we obtain:

Corollary 11.4. *If a generalized metric space* $\mathbf{E}$ *is bounded and hyperconvex, then every commuting family of nonexpansive self maps has a common fixed point.*

In order to prove the existence of a common fixed point for a family of nonexpansive mappings in the context of hyperconvex metric spaces, Baillon [5] discovered an intersection property satisfied by this class of metric spaces. In order to prove an analogue to Baillon's conclusion in our setting, we will need the following lemma.

Lemma 11.22. *Let* $\mathbf{E} := (E,d)$ *be a metric space over* $\mathscr{H}$*, endowed with a compact normal structure. Let* κ *be an infinite cardinal. For every ordinal* α*,* $\alpha < \kappa$*, let* B_α *and* E_α *be subsets of* E *such that:*

1. $B_\alpha \supseteq B_{\alpha+1}$ *and* $E_\alpha \supseteq E_{\alpha+1}$*, for every* $\alpha < \kappa$*;*
2. $\bigcap\limits_{\gamma<\alpha} B_\gamma = B_\alpha$ *and* $\bigcap\limits_{\gamma<\alpha} E_\gamma = E_\alpha$*, for every limit ordinal* $\alpha < \kappa$*;*
3. $\mathbf{E}_\alpha := \mathbf{E}_{\restriction E_\alpha}$ *is a one-local retract of* $\mathbf{E}$ *and* B_α *is a nonempty intersection of balls of* $\mathbf{E}_\alpha$*.*

Then $B_\kappa := \bigcap\limits_{\alpha<\kappa} B_\alpha \neq \emptyset$*.*

The proof, which is beyond the scope of this Chapter, can be found in [61].

From Lemma 11.22 it follows:

Theorem 11.9. *Let* $\mathbf{E} := (E,d)$ *be a generalized metric space. Assume that* $\mathbf{E}$ *has a compact normal structure. Then, the intersection of every down-directed family* $\mathscr{F}$ *of one-local retracts, is a nonempty one-local retract.*

Proof. Let $\mathbf{E} := (E,d)$ be a generalized metric space. Assume that $\mathbf{E}$ has a compact normal structure. Let P be the set, ordered by inclusion, of nonempty subsets A of E such that $\mathbf{E}_{\restriction A}$ is a one-local retract of $\mathbf{E}$. As is the case for any ordered set, every down-directed subset of P has an infimum iff every totally ordered subset of P has an infimum (see [23] Proposition 5.9 p 33). We claim that P is closed under the intersection of every chain of its members. Indeed, we argue by induction on the size of totally ordered families of one-local retracts of $\mathbf{E}$. First we may suppose that E has more than one element. Next, we may suppose that these families are dually well ordered by induction. Thus, given an infinite cardinal κ, let $(\mathbf{E}_{\restriction E_\alpha})_{\alpha<\kappa}$ be a descending sequence of one-local retracts of $\mathbf{E}$. From the induction hypothesis, we may suppose that the restriction of $\mathscr{E}$ to $E'_\alpha := \bigcap_{\gamma<\alpha} E_\gamma$ is a one-local retract of $\mathbf{E}$, for each limit ordinal $\alpha < \kappa$. Hence, we may suppose that $E_\alpha := \bigcap_{\gamma<\alpha} E_\gamma$, for each limit ordinal $\alpha < \kappa$. Since $\mathbf{E}_\alpha := \mathbf{E}_{\restriction E_\alpha}$ is a one-local retract of $\mathbf{E}$ and $\mathbf{E}$ has a normal structure, $\mathbf{E}_\alpha$ has a normal structure (Lemma 11.21). Hence, either E_α is a singleton, say x_α, or $r_{\mathbf{E}_\alpha}(E_\alpha) \setminus \delta_{\mathbf{E}_\alpha}(E_\alpha) \neq \emptyset$. In both cases, E_α is a ball of $\mathbf{E}_\alpha$. Hence in the first case, $E_\alpha = B_{\mathbf{E}_\alpha}(x_\alpha, r_{\restriction E_\alpha})$, whereas in second case, $E_\alpha \subseteq B_{\mathbf{E}_\alpha}(x, r)$, for some $x \in E_\alpha, r \in r_{\mathbf{E}_\alpha}(E_\alpha) \setminus \delta_{\mathbf{E}_\alpha}(E_\alpha)$. Thus, Lemma 11.22 applies with $B_\alpha = E_\alpha$ and yields that E_κ is nonempty. Let us prove that $\mathbf{E}_\kappa := \mathbf{E}_{\restriction E_\kappa}$ is a one-local retract of $\mathbf{E}$. We apply Lemma 11.15. Let $(B_{\mathbf{E}}(x_i, r_i))_{i\in I}, x_i \in E_\kappa$, $r_i \in \mathscr{H}$ be a family of balls with nonempty intersection. Since $\mathbf{E}_\alpha$ is a one-local retract of $\mathbf{E}$, the intersection $B_\alpha := E_\alpha \bigcap \bigcap_{i\in I} B_{\mathbf{E}}(x_i, r_i)$ is nonempty for every $\alpha < \kappa$. Now, Lemma 11.22 applied to the sequence $(E_\alpha, B_\alpha)_{\alpha<\kappa}$ yields the fact that $B_\kappa := E_\kappa \bigcap \bigcap_{i\in I} B_{\mathbf{E}}(x_i, r_i)$ is nonempty. According to Lemma 11.15, $\mathbf{E}_{\restriction B_\kappa}$ is a one-local retract of $\mathbf{E}$. □

The desired fixed point result follows from the preceding Theorem.

Proof of Theorem 11.8. For a subset $\mathscr{F}'$ of $\mathscr{F}$, let $Fix(\mathscr{F}')$ be the set of fixed points of $\mathscr{F}'$. Using Proposition 11.3, we conclude that $\mathbf{E}_{\restriction Fix(\mathscr{F}')}$ is a nonempty one-local retract of $\mathbf{E}$, for every finite subset $\mathscr{F}'$ of $\mathscr{F}$. We show this by induction on the number n of elements of $\mathscr{F}'$. If $n = 1$, the claim follows from Proposition 11.3. Let $n \geq 1$. Suppose that the property holds for every subset $\mathscr{F}''$ of $\mathscr{F}'$ such that $|\mathscr{F}''| < n$. Let $f \in \mathscr{F}'$ and $\mathscr{F}'' := \mathscr{F}' \setminus \{f\}$. From the inductive hypothesis, $\mathbf{E}_{\restriction Fix(\mathscr{F}'')}$ is a one-local retract of $\mathbf{E}$. Thus, according to lemma 11.21, $\mathbf{E}_{\restriction Fix(\mathscr{F}'')}$ has a compact normal structure. Now since f commutes with every member g of $\mathscr{F}''$, f preserves $Fix(\mathscr{F}'')$. Indeed, if $u \in Fix(\mathscr{F}'')$, we have $g(f(u)) = f(g(u)) = f(u)$, that is $f(u) \in Fix(\mathscr{F}'')$. Thus f induces an endomorphism f'' of $\mathbf{E}_{\restriction Fix(\mathscr{F}'')}$. According to Proposition 11.3, the restriction of $\mathbf{E}_{\restriction Fix(\mathscr{F}'')}$ to $Fix(f'')$, that is $\mathbf{E}_{\restriction Fix(\mathscr{F}')}$, is a nonempty one-local retract of $\mathbf{E}_{\restriction Fix(\mathscr{F}'')}$. Since the notion of one-local retract is transitive,s it follows that $\mathbf{E}_{\restriction Fix(\mathscr{F}')}$ is a nonempty one-local retract of $\mathbf{E}$. Let $\mathscr{P} := \{Fix(\mathscr{F}'') : |\mathscr{F}''| < \aleph_0\}$ and $P := \bigcap \mathscr{P}$. According to theorem 11.9, $\mathbf{E}_{\restriction P}$ is a one-local retract of $\mathbf{E}$. Since $P = Fix(\mathscr{F})$, the conclusion follows. □

11.6 Illustrations

11.6.1 The case of ordinary metric and ultrametric spaces

Let $\mathscr{H} := \mathbb{R}^+ \cup \{+\infty\}$. The inaccessible elements are 0 and $+\infty$. ence, if one deals with ordinary metric spaces, unbounded spaces in the above sense are those which are unbounded in the ordinary sense. If one deals with ordinary metric spaces, infinite products may yield spaces for which $+\infty$ is attained. On may replace powers with ℓ^∞-spaces. In this case, the notions of absolute retract, injective, hyperconvex and retract of some $\ell^\infty_{\mathbb{R}}(I)$-space coincide. This is the result of Aronzajn and Panitchpakdi [3]. The existence of an injective envelope was proved by Isbell [45]. The injective envelope of a 2-element, ordinary metric space is a closed interval $[0,v]$ of the real line with the distance given by the absolute value; injective envelopes of ordinary metric spaces consisting of at most five elements have been described [33]. For some applications, see [22, 33].

The existence of a fixed point for a nonexpansive map on a bounded hyperconvex space is the famous result of Sine and Soardi. When Theorem 11.8 is applied to a bounded, hyperconvex metric space, Baillon's fixed point theorem is obtained. The application of Theorem 11.8 to a metric space with a compact normal structure yields a result obtained by Khamsi [60]. For a survey about hyperconvex spaces we refer the reader to [32].

The results presented about injective spaces apply to ultrametric spaces over $\mathbb{R}^+ \cup \{+\infty\}$. A characterization similar to the one just presented was obtained in [12]; a description of the injective envelope is also given there. The paper [75] contains a study of ultrametric spaces over a complete lattice satisfying our distributivity condition, also called an *op-frame*. Metric spaces over op-frames are studied in [2]. Ultrametric spaces over a lattice and their connection with collections of equivalence relations are also studied in [19].

11.6.2 The case of ordered sets

Set $\mathscr{H} := \{-,0,1,+\}$ with $0 < -,+ < 1$. The retracts of powers of this lattice are all complete lattices. This is confirmed by the following fact.

Proposition 11.4. *A metric space $(\mathbf{E},d)$ over $\mathscr{H}$ is hyperconvex iff the corresponding poset is a complete lattice.*

Since 0 is the only inacessible element of $\mathscr{H}$, Theorem 11.4 applies:

Every commuting family of order-preserving maps on a complete lattice has a common fixed point. This is the full version of Tarski's theorem. Posets resulting from $\mathscr{H}$-metric spaces with a compact normal structure are a bit more general than complete lattices, hence Theorem 11.8 on compact normal structures could yield a bit more than Tarski's fixed point theorem. As will be seen below, in the case of one order-preserving map, this is no more than Abian-Brown's fixed-point theorem.

Let $\mathbf{P} := (E, \leq)$ be a poset. We observe first that the f.i.p. property of the collection of balls $\mathscr{B}_{\mathbf{P}} := \{\downarrow x : x \in E\} \cup \{\uparrow y : y \in E\}$ is an infinistic condition: it holds for every finite poset. In order to describe it we introduce the following notions. A pair of subsets (A, B) of E is called a *gap* of $\mathbf{P}$ if every element of A is dominated by every element of B but there is no element of E which dominates every element of A and is dominated by every element of B (cf. [30]). In other words: $(\bigcap_{x \in A} B(x, \leq)) \cap (\bigcap_{y \in B} B(y, \geq)) = \emptyset$ while $B(x, \leq) \cap B(y, \geq) \neq \emptyset$ for every $x \in A, y \in B$. A *subgap* of (A, B) is any pair (A', B') with $A' \subseteq A$, $B' \subseteq B$, which is a gap. The gap (A, B) is *finite* if A and B are finite, otherwise it is *infinite*. Say that an ordered set $\mathbf{Q}$ *preserves* a gap (A, B) of $\mathbf{P}$, if there is an order-preserving map g of $\mathbf{P}$ to $\mathbf{Q}$ such that $(g(A), g(B))$ is a gap from $\mathbf{Q}$. For further information on the preservation of gaps, see [69].

Lemma 11.23. *Let $\mathbf{P} := (E, \leq)$ be a poset. Then:*

1. *$\mathbf{P}$ is a complete lattice iff $\mathbf{P}$ contains no gap;*

2. *An order-preserving map $f : \mathbf{P} \to \mathbf{Q}$ preserves all gaps of $\mathbf{P}$ iff it preserves all holes of $\mathbf{P}$ with values in $\mathscr{H} \setminus \{0\}$, iff $\mathbf{Q}_{\restriction P}$ is a one-local retract of $\mathbf{Q}$;*

3. *$\mathscr{B}_{\mathbf{P}}$ satisfies the f.i.p. iff every gap of $\mathbf{P}$ contains a finite subgap, iff every hole is finite.*

Since the proof is straightforward, it will be omitted. We may note the similarity of (b) and Lemma 11.16.

From item (c) of Lemma 11.23 it follows that every nonempty chain in a poset $\mathbf{P}$ for which the collection of balls has the $f.i.p$, has a supremum and an infimum. Such a poset is said to be *chain-complete*.

Abian-Brown's theorem [1] asserts that *in a chain-complete poset with a least or largest element, every order-preserving map has a fixed point*. The fact that the collection of intersections of balls of $\mathbf{P}$ has a normal structure means that every nonempty intersection of balls of $\mathbf{P}$ has either a least or a largest element. Being the intersection of the empty family of balls, $\mathbf{P}$ has either a least element or a largest element. Consequently, if $\mathbf{P}$ has a compact normal structure, we may suppose without loss of generality that it has a least element. Since every nonempty chain has a supremum, it follows from Abian-Brown's theorem that every order preserving map has a fixed point.

On the other hand, a description of posets with a compact normal structure is still open. We just observe that retracts of powers of $\bigvee$ or retracts of powers of $\bigwedge$ have a compact normal structure.

Theorem 11.8 above yields a fixed point theorem for a commuting family of order-preserving maps on any retract of a power of $\bigvee$ or of a power of $\bigwedge$. But this result says nothing about retracts of products of $\bigvee$ and $\bigwedge$. These two posets fit into the category of fences. As we have seen in Subsection 11.3.4, every poset embeds isometrically (with respect to the fence distance) into a product of fences. Fences are hyperconvex, hence the following Theorem follows from Theorem 11.4.

Theorem 11.10. *[61, Theorem 4.18] If a poset* $\mathbf{Q}$ *is a retract of a product* $\mathbf{P}$ *of finite fences of bounded length, every commuting set of order-preserving maps on* $\mathbf{Q}$ *has a fixed point.*

Since every complete lattice is a retract of a power of the two-element chain, this result contains Tarski's fixed point theorem.

11.6.3 The case of graphs

Retracts of (undirected) graphs have been considered by various authors, for reflexive as well as for irreflexive graphs (see [7, 36, 38, 39, 40]. The existence of the injective envelope of an undirected graph (presented in [47]) is given in [74], a characterization of injective graphs is presented in [8].

To each directed graph we have associated its zigzag distance, yielding a metric with values in $\mathscr{H}_\Lambda := \mathbf{F}(\Lambda^*)$. Metric spaces over $\mathscr{H}_\Lambda$ whose distance is the zigzag distance associated with a reflexive directed graph, were characterized by Lemma 11.2. The condition stated there is a weak form of convexity, thus it holds for hyperconvex spaces. Let $\mathscr{D} := \mathscr{H}_\Lambda$ and $\mathscr{G}_\mathscr{D}$ be the class of graphs whose zigzag distance belongs to $\mathscr{D}$. With the homomorphisms of graphs, this class becomes a category. As a category, $\mathscr{G}_\mathscr{D}$ identifies with a full subcategory of $\mathscr{M}_\mathscr{D}$, the category of metric spaces over $\mathscr{D}$, with the nonexpansive maps as morphisms (see Lemma 11.1).

According to Theorem11.3, one has:

Theorem 11.11. *A member* $\mathbf{E} := (E,d)$ *of* $\mathscr{M}_\mathscr{D}$ *is an absolute retract iff the distance on* E *comes from a directed graph and this graph is an absolute retract in* $\mathscr{G}_\mathscr{D}$*, with respect to isometric embedding.*

These members of $\mathscr{G}_\mathscr{D}$ are described by the following result:

Theorem 11.12. *[47] For a reflexive directed graph* $\mathbf{G} = (E,\mathscr{E})$*, the following conditions are equivalent* *(i)* $\mathbf{G}$ *is an absolute retract with respect to isometries;*

(ii) $\mathbf{G}$ *is injective with respect to isometries;*

(iii) *G has the extension property;*

(iv) *The collection of balls* $B(x, \uparrow \alpha)_{x \in E, \alpha \in \Lambda^*}$ *has the 2-Helly property;*

(v) $\mathbf{G}$ *is a retract of a power of* $\mathbf{G}_{\mathscr{H}_\Lambda}$.

Every metric space $\mathbf{E}$ over $\mathscr{H}_\Lambda$ has an injective envelope; being injective, its metric comes from a graph. If $\mathbf{E}$ comes from a graph, the graph corresponding to the injective envelope of $\mathbf{E}$ is the injective envelope in $\mathscr{G}_{\mathscr{D}}$. For more recent facts about the injective envelope, see [57].

We just mention a simple example of hyperconvex graph.

Lemma 11.24. *The metric space associated to any directed zigzag* $\mathbf{Z}$ *has the extension property. In particular, every nonexpansive map sending two vertices of a reflexive directed graph* $\mathbf{G}$ *to the extremities of* $\mathbf{Z}$*, extends to a graph homorphism from* $\mathbf{G}$ *to* $\mathbf{Z}$*.*

Proof. Let $\mathbf{Z}$ be a directed zigzag (with loops). Its symmetric hull (obtained by deleting the orientation of arcs in $\mathbf{Z}$) is a path. The balls in $\mathbf{Z}$ are intervals of that path, and each of these intervals is either finite or the full path. Hence, every family of balls has the 2-Helly property. Since convexity holds trivially in this case, $\mathbf{Z}$, as a metric space over $\mathscr{H}_\Lambda$, is hyperconvex, hence according to Theorem 11.3, it satisfies the extension property. □

We refer the reader to [53, 55] for further discussions on this topic.

11.6.4 The case of oriented graphs

The situation of oriented graphs is different from that of the preceding Section. Oriented graphs cannot be modeled over a Heyting algebra (theorem IV-3.1 of [47] is erroneous), but the absolute retracts in this category can (this was proved by Bandelt, Saïdane and the second author and included in [85]; see also the forthcoming paper [11]). The appropriate Heyting algebra is the *MacNeille completion* of Λ^*, where $\Lambda := \{+, -\}$.

The MacNeille completion of Λ^* is in some sense the least complete lattice that extends Λ^*. The definition goes as follows. If X is a subset of Λ^* ordered by the subword ordering, then

$$\uparrow X := \{\beta \in \Lambda^* : \alpha \leq \beta \text{ for some } \alpha \in X\}$$

is the *final segment generated by X* and

$$\downarrow X := \{\alpha \in \Lambda^* : \alpha \leq \beta \text{ for some } \beta \in X\}$$

is the *initial segment generated by* X. For a singleton $X = \{\alpha\}$, we omit the set brackets and call $\uparrow\alpha$ and $\downarrow\alpha$ a *principal final segment* and a *principal initial segment*, respectively. We refer to

$$X^{\Delta} := \bigcap_{x \in X} \uparrow x$$

as the *upper cone* generated by X, and

$$X^{\nabla} := \bigcap_{x \in X} \downarrow x$$

is the *lower cone* generated by X. The pair (Δ, ∇) of mappings on the complete lattice of subsets of Λ^* constitutes a Galois connection. Thus, a set Y is an upper cone if and only if $Y = Y^{\nabla\Delta}$, while a set W is an lower cone if and only if $W = W^{\Delta\nabla}$. This Galois connection (Δ, ∇) yields the *MacNeille completion* of Λ^*. This completion is realized as the complete lattice $\{W^{\nabla} : W \subseteq \Lambda^*\}$, ordered by inclusion or alternatively, $\{Y^{\Delta} : Y \subseteq \Lambda^*\}$ ordered by the reverse inclusion. We choose as completion the set $\{Y^{\Delta} : Y \subseteq \Lambda^*\}$ ordered by the reverse inclusion and we denote it by $\mathbf{N}(\Lambda^*)$. This complete lattice is studied in detail in [10]. We recall the following characterization of members of the MacNeille completion of Λ^*.

Proposition 11.5. *[10] corollary 4.5. A member Z of $\mathbf{F}(\Lambda^*)$ belongs to $\mathbf{N}(\Lambda^*)$ if and only if it satisfies the following cancellation rule: if $u + v \in Z$ and $u - v \in Z$ then $uv \in Z$.*

The concatenation, order and involution defined on $\mathbf{F}(\Lambda^*)$ induce a Heyting algebra $\mathscr{N}_\Lambda$ on $\mathbf{N}(\Lambda^*)$ (see Proposition 2.2 of [10]). Being a Heyting algebra, $\mathscr{N}_\Lambda$ supports a distance $d_{\mathscr{N}_\Lambda}$ and this distance is the zigzag distance of a graph $\mathbf{G}_{\mathscr{N}_\Lambda}$. But, it is not true that every oriented graph embeds isometrically into a power of that graph. For example, an oriented cycle cannot be embedded. The following result characterizes graphs which can be isometrically embedded, via the zigzag distance, into products of reflexive and oriented zigzags. It is partially stated in Subsection IV-4 of [47], cf. Proposition IV-4.1.

Theorem 11.13. *For a directed graph $\mathbf{G} := (E, \mathscr{E})$ equipped with the zigzag distance, the following properties are equivalent:*

(i) $\mathbf{G}$ is isometrically embeddable into a product of reflexive and oriented zigzags;

(ii) $\mathbf{G}$ is isometrically embeddable into a power of $\mathbf{G}_{\mathscr{N}_\Lambda}$;

(iii) The values of the zigzag distance between vertices of $\mathbf{G}$ belong to $\mathscr{N}_\Lambda$.

The proof follows along the same lines as the proof of Proposition IV-5.1 p.212 of [47].

We may note that the product can be infinite even if the graph **G** is finite. Indeed, if **G** consists of two vertices x and y with no value on the pair $\{x,y\}$ (that is the underlying graph is disconnected), then we need infinitely many zigzags of arbitrarily long length.

Theorem 11.14. *An oriented graph* $\mathbf{G} := (V, \mathscr{E})$ *is said to be an absolute retract in the category of oriented graphs, if and only if it is a retract of a product of oriented zigzags.*

The proof will be sketched; see Chapter V of [85] and the forthcoming paper [11] for specific details. We proceed in three steps. Let **G** be an absolute retract. It will first be proved that **G** has no 3-element cycle. In the second stage it is shown that the zigzag distance between two vertices of **G** satisfies the cancelation rule. From Proposition 11.5, it belongs to $\mathbf{N}(\Lambda^*)$; from theorem 11.13, **G** isometrically embeds into a product of oriented zigzags. Since **G** is an absolute retract, it is a retract of that product. As illustrated by the results of Tarski and Sine-Soardi, absolute retracts are appropriate candidates for the fixed point property. Reflexive graphs with the fixed point property must be antisymmetric, i.e., oriented. Having described absolute retracts among oriented graphs, it is clear from Theorem 11.4 that the bounded ones have the fixed point property. We start with a characterization of accessible elements of $\mathscr{N}_\Lambda$. The proof is omitted.

Lemma 11.25. *Every element* v *of* $\mathscr{N}_\Lambda \setminus \{\Lambda^*, \emptyset\}$ *is accessible.*

Theorem 11.15. *If a graph* **G***, finite or not, is a retract of a product of reflexive and directed zigzags of bounded length, then every commuting set of endomorphisms has a common fixed point.*

Proof. We may suppose that **G** has more than one vertex. The diameter of **G** equipped with the zigzag distance belongs to $\mathscr{N}_\Lambda \setminus \{\Lambda^*, \emptyset\}$. According to Lemma 11.25, it is accessible, hence as a metric space, **G** is bounded. Being a retract of a product of hyperconvex metric spaces, it is hyperconvex. Theorem 11.4 then applies. □

The properties of reflexive and involutive transition systems extend almost verbatim the properties of directed graphs. They have been extended to non-necessarily reflexive transition systems ([75], [43, 44]. Instead of presenting these properties, we illustrate their use in the following section.

11.7 An illustration of the usefulness of the injective envelope

Using the notion of injective envelope, we prove that on an ordered alphabet Λ the monoid $\mathbf{F}^{\circ}(\Lambda^*) := \mathbf{F}(\Lambda^*) \setminus \{\emptyset\}$ is free. This result is presented in [56].

Theorem 11.16. $\mathbf{F}^{\circ}(\Lambda^*)$ *is a free monoid.*

We recall that a member F of $\mathbf{F}(\Lambda^*)$ is *irreducible* if it is distinct from Λ^* and is not the concatenation of two members of $\mathbf{F}(\Lambda^*)$, distinct of F (note that with this definition, the empty set is irreducible). The fact that $\mathbf{F}^{\circ}(\Lambda^*)$ is free amounts to the fact that each member decomposes in a unique way as a concatenation of irreducible elements. Both a synctactical proof and a geometrically-flavored proof are given in in [56]; only the last one will be presented here.

We suppose that Λ is equipped with an involution (this is not a restriction: we may choose the identity on Λ as our involution). Then, we consider metric spaces such that the values of their distances belong to $\mathbf{F}(\Lambda^*)$. The category of metric spaces over $\mathbf{F}(\Lambda^*)$, with the nonexpansive maps as morphisms, has enough injectives. Furthermore, for every final segment F of Λ^*, the 2-element metric space $\mathbf{E} := (\{x, y\}, d)$ such that $d(x, y) = F$, has an *injective envelope* $\mathscr{S}_F$.

We define the gluing of two metric spaces by a common vertex. Suppose that two metric spaces $\mathbf{E}_1 := (E_1, d_1)$ and $\mathbf{E}_2 := (E_2, d_2)$ have only one common vertex, say r. On the union $E_1 \cup E_2$ we may define a distance extending both d_1, d_2, setting $d(x, y) := d_i(x, r) \oplus d_j(r, y)$ for $x \in E_i, y \in E_j$, and $i \neq j$. If E_1 and E_2 are arbitrary, we may replace them by isometric copies with a common vertex. We apply this construction to the injective envelope of two-element metric spaces. Let v_1 and v_2 be two elements of a Heyting algebra and $\mathscr{S}_{v_1}, \mathscr{S}_{v_2}$ be their injective envelopes. Suppose that $\mathscr{S}_{v_1}$ is the injective envelope of $\{x_1, y_1\}$, with $x_1 := 0, y_1 := v_1$ and that $\mathscr{S}_{v_2}$ is the injective envelope of $\{x_2, y_2\}$ with $x_2 := v_1$ and has no other element in common with $\mathscr{S}_{v_1}$. Let $\mathscr{S}_{v_1} \oplus \mathscr{S}_{v_2}$ be their gluing. Since the distance from x_1 to y_2 is $v_1 \oplus v_2$, this space embeds isometrically into the injective envelope $\mathscr{S}_{v_1 \oplus v_2}$. For some Heyting algebras (and v_1, v_2 distinct from 1), these two spaces are isometric (see Figure 11.13 for a geometric interpretation). This is the case of the Heyting algebra $\mathbf{F}(\Lambda^*)$ (Corollary 4.9, p. 177 of [55]). In terms of this Heyting algebra, this yields (with self-explanatory notation):

$$\mathscr{S}_{F_1}\mathscr{S}_{F_2} \cong \mathscr{S}_{F_1 F_2} \text{ for all } F_1, F_2 \in \mathbf{F}^{\circ}(\Lambda^*). \tag{11.19}$$

Say that an injective which is not the gluing of two proper injectives is *irreducible*. From (11.19) it follows that an injective of the form $\mathscr{S}_F$ is irreducible iff F is irreducible.

In order to prove that the decomposition of a final segment F into a concatenation of irreducible final segments is unique, we consider the transition system $\mathscr{M}_F$ on the alphabet Λ, with transitions (p,a,q) if $a \in d(p,q)$, corresponding to the injective envelope $\mathscr{S}_F$. The automaton $\mathscr{A}_F := (\mathscr{M}_F, \{x\}, \{y\})$, with $x = \Lambda^*$ as initial state and $y = F$ as final state, accepts F. A transition system yields a directed graph whose arcs are the ordered pairs (x,y) linked by some transition. Since the transition system $\mathscr{M}_F$ is reflexive and involutive, the corresponding graph $\mathbf{G}_F$ is undirected and has a loop at every vertex. For an example, if $F = \Lambda^*$, $\mathscr{S}_F$ is the one-element metric space and $\mathbf{G}_F$ reduces to a loop. If $F = \emptyset$, $\mathscr{S}_F$ is the two-element metric space $E := (\{x,y\}, d)$ with $d(x,y) = \emptyset$ and $\mathbf{G}_F$ has no edge. The cut vertices of $\mathbf{G}_F$ (vertices whose deletion increases the number of connected components) allow to reconstruct the irreducible components of $\mathscr{S}_F$.

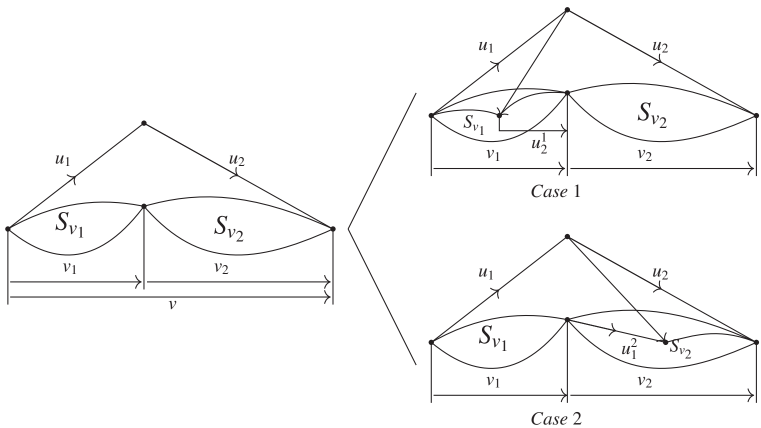


FIGURE 11.13: Interpretation of the convexity property of a pair (v_1, v_2).

With the notion of cut vertex and block borrowed from graph theory, we prove:

Theorem 11.17. *Let F be a final segment of Λ^*, distinct from Λ^*. Then F is irreducible if and only if $\mathscr{S}_F$ is irreducible, if and only if $\mathbf{G}_F$ has no cut vertex. If F is not irreducible, the blocks of $\mathbf{G}_F$ are the vertices of a finite path $C_0, \ldots, C_{n-1}$ with $n \geq 2$, whose end vertices C_0 and C_{n-1} contain respectively the initial state x and the final state y of the automaton $\mathscr{A}_F$ accepting F. Furthermore, F is the concatenation $F_0 \ldots F_i \ldots F_{n-1}$, the automaton $\mathscr{A}_{F_i}$*

accepting F_i being isomorphic to $(\mathscr{M}_F \restriction C_i, \{x_i\}, \{x_{i+1}\})$*, where* $x_0 := x$*,* $x_n := y$ *and* $\{x_{i+1}\} = C_i \cap C_{i+1}$ *for* $0 \leq i < n-1$.

From this result, the fact that $\mathbf{F}^{\circ}(\Lambda^*)$ is free follows easily.

This result does not yield a concrete test for irreducibility. The size of the injective envelope $\mathscr{S}_F$ in terms of the length of words generating F can be a double exponential (see Subsection 4.5 of [56]). But it suggests a similar result for the minimal automaton recognizing F. In [56], we prove

Theorem 11.18. *If $\mathscr{A}$ is the minimal deterministic automaton recognizing a final segment $F \in \mathbf{F}^{\circ}(\Lambda^*)$, then F is irreducible iff there is no vertex z distinct from the initial state x and the final state y, which lies on all directed paths going from x to y.*

11.8 Further developments

There are several interesting examples of generalized metric spaces for which the set of values is not a Heyting algebra.

This is the case for metric spaces over a Boolean algebra (except if the Boolean algebra is the power set of a set). If B is a Boolean algebra, not necessarily complete, or not satisfying the distributivity condition, residuation holds (i.e., for every $x, y \in B$, $y \setminus x$ is the least element r of B such that $x \leq y \vee r$); hence, one may define a distance d over B: the distance $d(p,q)$ between two elements p, q of B is the symmetric difference $p\Delta q := p \setminus q \cup q \setminus p$. If B is complete, Theorem 11.1 holds.

Another example is arithmetic in nature. The Chinese remainder theorem can be viewed as a property of balls in a metric space. Indeed, if a_i, r_i $(i \in I)$ is a family of pairs of integers we may view each congruence class of a_i modulo r_i in $\mathbb{Z}$ as a (closed) ball $B(a_i, r_i) := \{x \in \mathbb{Z} : d(a_i, x) \preceq r_i\}$, for a suitable distance d on $\mathbb{Z}$ and an order $\preceq$ on the set of values of the distance. The Chinese remainder theorem characterizes the situation when these balls have a nonempty intersection. As we have seen, the Helly property and convexity are the keywords to ensure a non-empty intersection of balls. In our case, $\mathbb{Z}$ has a structure of ultrametric space with values in $\mathbb{N}$ provided that $\mathbb{N}$ is ordered by the reverse of divisibility, setting $n \preceq m$ if n is a multiple of m. In this way $(\mathbb{N}, \preceq)$ becomes a distributive, complete lattice, with least element 0 and largest element 1; the join $n \vee m$ of n and m is their largest common divisor. Replacing the addition by the join and setting $d(a,b) := |a-b|$ for any two elements $a, b \in \mathbb{Z}$, we have $d(a,b) = 0$ iff $a = b$; $d(a,b) = d(b,a)$ and $d(a,b) \preceq d(a,c) \vee d(c,b)$ for all $a, b, c \in \mathbb{Z}$.

With this definition, closed balls in $\mathbb{Z}$ are congruence classes of the additive group $(\mathbb{Z},+)$. In an ordinary metric space $\mathscr{V} := (V,\preceq)$, the necessary condition for the non-emptiness of the intersection of two balls $B(a_i,r_i)$ and $B(a_j,r_j)$ is the *convexity property*, namely the condition that the distance between centers is at most the sum of the radii. In this case, the above condition translates into $d(a_i,a_j) \preceq r_i \vee r_j$, i.e. a_i and a_j are congruent modulo $lcd(r_i,r_j)$. The Chinese remainder theorem expresses that the intersection of *finitely many* balls $B(a_i,r_i)$ is non-empty iff this family of balls $B(a_i,r_i)$ satisfies the convexity property and the finite 2-Helly property. This property does not extend to infinite families: the space $\mathbb{Z}$ is not hyperconvex (and $\mathbb{N}$ equipped with the join as a monoid operation is not a Heyting algebra); we may say that it is *finitely hyperconvex*.

Metric spaces over $(\mathbb{N},\preceq)$, like $\mathbb{Z}$, are examples of metric spaces over a join-semilattice $\mathscr{V} := (V,\preceq)$ with a 0. They fit into the category of ultrametric spaces. If $\mathbf{E} := (E,d)$ is such a metric space, set $\equiv_r := \{(x,y) \in E : d(x,y) \preceq r\}$ for each r; this defines an equivalence relation on E. Let $\mathrm{Eqv}(E)$ be the set of equivalence relations on E and set $\mathrm{Eqv}_d(E) := \{\equiv_r : r \in V\}$. Then, it is easy to see that any two members of $\mathrm{Eqv}_d(E)$ commute and $\equiv_r \circ \equiv_s = \equiv_s \circ \equiv_r = \equiv_{r\vee s}$, for every $r,s \in V$, iff (E,d) is convex. If the meet of every non-empty subset of V exists, then $\mathrm{Eqv}_d(E)$ is an intersection-closed subset of $\mathrm{Eqv}(E)$. Furthermore, (E,d) is hyperconvex iff $\mathrm{Eqv}_d(E)$ is a completely meet-distributive lattice of $Eq(E)$ (Proposition 3.12 of [76]). A sublattice L of the lattice $\mathrm{Eqv}(E)$ of equivalence relations is *arithmetical* (see [50]) if it is distributive and pairs of members of L commute with respect to composition. As is well known (see [50]), arithmetic lattices can be characterized in terms of the Chinese remainder conditions (expressed as in the theorem mentionned above). This property amounts to finite hyperconvexity, and it yields the one-extension property for maps with finite domains and the fact that if E is countable, then every partial nonexpansive map from a finite subset of E, extends to E [49]. The study of maps preserving congruences, nonexpansive maps in our setting, is a very basic subject of universal algebra (for a beautiful recent result, see [21]). Some results about metric spaces over meet-distributive lattices and their nonexpansive maps were obtained in [75, 77]. The relation with universal algebra (and arithmetic) suggests the consideration of possible extensions.

Bibliography

[1] S. Abian, A. Brown, *A theorem on partially ordered sets, with applications to fixed point theorems*, Canad. J. Math. 13 (1961) 78- 82.

[2] N. L. Ackerman, *Completeness in generalized ultrametric spaces,* p-Adic Numbers Ultrametric Anal. Appl. 5 (2013), no. 2, 89-105.

[3] N. Aronszajn, P. Panitchpakdi, *Extension of uniformly continuous transformations and hyperconvex metric spaces*, Pacific J.Math. 6 (1956), 405-439.

[4] K. Baclawski, A. Björner, *Fixed points in partially ordered sets*, Advances in Math. 31,(1979) 263-287.

[5] J. B. Baillon, *Non expansive mapping and hyperconvex spaces*, Fixed point theory and its applications (Berkeley, CA, 1986), 11-19, Contemp. Math., 72, Amer. Math. Soc., Providence, RI, (1988).

[6] B. Banaschewski, G. Bruns, *Categorical characterization of the MacNeille completion*, Arch. Math. (Basel) 18 (1967) 369–377.

[7] H-J. Bandelt, M. Farber, P. Hell, *Absolute reflexive retracts and absolute bipartite retracts*, Discrete Applied Math., 44, 35 (1983), 9-20.

[8] H-J. Bandelt, E. Pesch, *Dismantling absolute retracts of reflexive graphs*, European J. Combin. 10 (1989), no. 3, 211–220.

[9] H-J. Bandelt, A. W. M. Dress, *A canonical decomposition theory for metrics on a finite set*, Adv. Math. 92 (1992), 47–105.

[10] H-J. Bandelt, M. Pouzet, *A syntactic approach to the MacNeille completion of* Λ^**, the free monoid over an ordered alphabet* Λ, Order, 16pp https://doi.org/10.1007/s11083-018-9462-7, published on line June 25, (2018).

[11] H-J. Bandelt, M. Pouzet, F. Saïdane, *Absolute retracts of reflexive oriented graphs*, circulating manuscrit, 20 pp. April 2018.

[12] J. M. Bayod, J. Martínez-Maurica, *Ultrametrically injective spaces,* Proc. Amer. Math. Soc. 101 (3), (1987) 571-576.

[13] G. Birkhoff, *Lattice theory*. Corrected reprint of the 1967 third edition. American Mathematical Society Colloquium Publications, 25. American Mathematical Society, Providence, R.I., (1979). vi+418 pp.

[14] L.M. Blumenthal, *Boolean geometry,* I. Rend. Circ. Mat. Palermo (2) 1 (1952), 343–360.

[15] L.M. Blumenthal, *Theory and applications of distance geometry,* Second edition Chelsea Publishing Co., New York (1970) xi+347 pp.

[16] L.M. Blumenthal, K. Menger, *Studies in geometry,* W. H. Freeman and Co., San Francisco, Calif. (1970) xiv+512 pp.

[17] T. S. Blyth, M.F. Janowitz, *Residuation Theory*, Volume 102 in International Series of Monographs on Pure and Applied Mathematics, Oxford, New-York, Pergamon Press, 382 pp.

[18] J. A. Bondy, U. S. R. Murty, *Graph theory*. Graduate Texts in Mathematics, 244. Springer, New York, (2008), xii+651 pp.

[19] S. Braunfeld, Λ-*ultrametric spaces and lattices of equivalence relations*, 8pp, arXiv:1811.11120v1.

[20] R. E. Bruck, *A common fixed point theorem for a commuting family of non-expansive mappings*, Pac. J. Math., 53 (1974), 59-71.

[21] P. Cégielski, S. Grigorieff, I. Guessarian, *Newton representation of functions over natural integers having integral difference ratios*, Int. J. Number Theory 11 (2015), no. 7, 2109–2139.

[22] V. Chepoï, *A T_X-approach to some results on cuts and metrics*, Adv. in Appl. Math. 19 (1997), no. 4, 453–470.

[23] P. M. Cohn, *Universal algebra*, Harper & Row, Publishers, New York-London (1965) xv+333 pp.

[24] G. Conant, *Distance structures for generalized metric spaces*, Annals of Pure and Applied Logic 168 (2017), no. 3, 622–650.

[25] C. Delhommé, C. Laflamme, M. Pouzet, N. Sauer, *Divisibility of countable metric spaces*, European Journal of Combinatorics **28** (2007) 1746-1769.

[26] C. Delhommé, C. Laflamme, M. Pouzet, N. Sauer, *Indivisible ultrametric spaces*, Topology and its Applications **155** (2008) 1462-1478.

[27] C. Delhommé, C. Laflamme, M. Pouzet, N. Sauer, *On homogeneous ultrametric spaces*, pdf, 23 pp; arXiv:1509.04346, (2015).

[28] R. DeMarr, *Common fixed point theorem for commuting contraction mappings*, Pac. J. of Math., 13 (1963), 1139-1141.

[29] M. Deza, E. Deza, *Encyclopedia of distances*, Fourth edition. Springer, Heidelberg, (2016) xxii+756 pp.

[30] D. Duffus, I. Rival : *A structure theory of ordered sets*, J. of Discrete Math. 35 (1981), 53-118.

[31] P. Eklund, J. Gutiérrez García, U. Höhle, J. Kortelainen, *Semigroups in Complete Lattices: Quantales, Modules and Related Topics*, Developments in Mathematics, Springer (2018).

[32] R. Espínola, M. A. Khamsi, *Introduction to hyperconvex spaces*, In Handbook of metric fixed point theory, 391–435, Kluwer Acad. Publ., Dordrecht, (2001).

[33] A. W. N. Dress : *Trees, tight extensions of metric spaces, and the cohomological dimension of certain groups, a note on combinatorial properties of metric spaces*, Adv. in Math. 53 (3) (1984), 321-402.

[34] R. Fraïssé, Sur l'extension aux relations de quelques propriétés des ordres. *Annales scientifiques de l'École Normale Supérieure*, Sér. 3, **71** no. 4 (1954), p. 363-388.

[35] R. Fraïssé, *Theory of Relations*, Revised Edition, in : *Studies in Logic and the Foundations of Mathematics*, **145**, North Holland 2000.

[36] P. Hell, *Absolute retracts of graphs*, Lecture notes 406 (1974), PP. 291-301.

[37] P. Hell, *Graph retractions*, Atti dei conveigni lincei 17, teorie combinatorie (1976), PP. 263-268.

[38] P. Hell, *Rétractions de graphes*, PhD. Université de Montréal, Juin 1972, 148 pages.

[39] P. Hell, *Subdirect products of bipartite graphs*, in Infinite and Finite Sets (V.T.Sos et al., eds), Colloq. Math. Soc. J.Bolyai 10 (1975) 857-866.

[40] P. Hell, I. Rival, *Absolute retracts and varieties of reflexive graphs*, Canadian J. of Math. 39 (1987) 544-567.

[41] G. Higman, *Ordering by divisibility in abstract algebra*, Proc. London Math. Soc (3) 2 (1952), 326-336.

[42] J. Hubička, M. Konečný, J. Nešetřil, *Conant's generalised metric spaces are Ramsey*, available at arXiv:1710.04690 [math.CO] (2017), 22 pages.

[43] A. Hudry, *Rétractions, corétractions et envelope injective d'une algèbre de transitions,* Discrete Math. 247 (2002), no. 1-3, 117–134.

[44] A. Hudry, *Injective envelope and parallel decomposition of a transition system,* Discrete Math. 289 (2004), no. 1-3, 45-61.

[45] J. Isbell : *Six theorems about injective metric spaces*, Comment. Math. Helv. 39 (1964), 65-76.

[46] E. Jawhari, *Les rétractions dans les graphes. Applications et généralisations*. Thèse de 3ième cycle, Université Claude-Bernard, 7 juillet 1983, Lyon.

[47] E. Jawhari, D. Misane, M. Pouzet, *Retracts : graphs and ordered sets from the metric point of view*, (I.Rival, ed) Contemporary Mathematics, Vol 57, (1986), 175-226.

[48] B. Jónsson, *Homogeneous universal relational systems*, Mathematica Scandinavica, **8** (1960), 137-142.

[49] K. Kaarli, *A new characterization of arithmeticity*, Conference on Lattices and Universal Algebra (Szeged, 1998). Algebra Universalis **45** (2001), no. 2-3, 345–347.

[50] K. Kaarli, A.F. Pixley, *Polynomial completeness in algebraic systems*, Chapman and Hall/CRC, Boca Raton, FL, 2001. xvi+358 pp.

[51] K. Kaarli, S. Radeleczki, *Representation of integral quantales by tolerances*, Algebra Universalis 79 (5) (2018), https://doi.org/10.1007/s00012-018-0484-1

[52] M. Kabil : *Enveloppe injective de graphes et de systèmes de transitions et idéaux de mots*, Thèse de doctorat, Lyon; Juin 1992, n^o 149-92.

[53] M. Kabil, M. Pouzet, *Indécomposabilité et irréductibilité dans la variété des rétractes absolus des graphes réflexifs,* C.R.Acad.Sci. Paris Série A 321 (1995), 499-504.

[54] M. Kabil, *Une approche métrique de l'indécomposabilité et de l'irréductibilité dans la variété des rétractes absolus des graphes et des systèmes de transitions*, Thèse de doctorat d'État, Université Hassan II Aïn Chock, Casablanca, 19 Décembre 1996.

[55] M. Kabil, M. Pouzet, *Injective envelope of graphs and transition systems,* Discrete Math.192 (1998), 145-186.

[56] M. Kabil, M. Pouzet, I. G. Rosenberg, *Free monoids and metric spaces*, To the memory of Michel Deza, Europ. J. of Combinatorics, Online publication complete: April 2, (2018), https://doi.org/10.1016/j.ejc.2018.02.008. arXiv: 1705.09750v1, 27 May 2017.

[57] M. Kabil, M. Pouzet, *Injective envelopes of transition systems and Ferrers languages*, 23pp. July 4, (2019), arXiv:1907.02231

[58] M. Katětov, *On universal metric spaces*, In: General topology and its relations to modern analysis and algebra, VI (Prague, 1986), Res. Exp. Math., 16, Heldermann, Berlin, (1988).

[59] A. S. Kechris, V. G. Pestov, S. Todorcevic, *Fraïssé limits, Ramsey theory, and topological dynamics of automorphism groups*, Geom. Funct. Anal., **15** (2005)106-189.

[60] M. A. Khamsi, *One-local retract and common fixed point for commuting mappings in metric spaces*, Nonlinear Anal. 27 (1996), no. 11, 1307–1313.

[61] M. A. Khamsi, M. Pouzet, *A fixed point theorem for commuting families of relational homomorphisms. Applications to metric spaces, oriented graphs and ordered sets*, 24 pp. March 2018, arXiv:1805.02594; accepted for publication at the special volume of Topology Appl. dedicated to Ales Pultr.

[62] W. A. Kirk, *A fixed point for mappings which do not increase distances*, Amer. Math. Monthly, 72 (1965), 1004-1006.

[63] E .W. Kiss, L. Márki, P. Pröhle, W. Tholen, *Algebraic properties. A compendium on amalgamation, congruence extension, epimorphisms, residual smallness, and injectivity* Studia Sci. Math. Hungar. 18 (1982), no. 1, 79–140.

[64] F. W. Lawvere, *Metric spaces, generalized logic, and closed categories*, Rend. Sem. Mat. Fis. Milano 43 (1973), 135–166 (1974).

[65] T. C. Lim, *A fixed point theorem for families of non-expansive mappings*, Pacific J. of Math., 53 (1974), 487-493.

[66] D. Misane, *Rétracts absolus d'ensembles ordonnées et de graphes. Propriétés du point fixe*, Thèse de 3ème cycle, Université Claude-Bernard, 14 septembre 1984.

[67] C. J. Mulvey, *Second topology conference (Taormina, 1984)*. Rend. Circ. Mat. Palermo (2) Suppl. No. 12 (1986), 99-104.

[68] P. Nevermann, I. Rival, *Holes in ordered sets*, Graphs Combin. 1 (1985), no. 4, 339-350.

[69] P. Nevermann, R. Wille, *The strong selection property and ordered sets of finite length*, Algebra Universalis, 18 (1984), 18-28.

[70] L. Nguyen Van Thé, *Structural Ramsey theory of metric spaces and topological dynamics of isometry groups*, Mem. Amer. Math. Soc. 206 (2010), no. 968, x+140 pp. ISBN: 978-0-8218-4711-4

[71] R. Nowakowski, I. Rival, *The smallest graph variety containing all paths*, J. of Discrete Math. 43 (1983), 185-198.

[72] J. P. Penot, *Fixed point theorems without convexity*, in Analyse non convexe (1977, PAU), Bull. Soc. Math. France, Mémoire 60, (1979), 129-152.

[73] J. P. Penot, *Une vue simplifiée de la théorie de la complexité*, Gaz. Math. No. 34 (1987), 61-77.

[74] E. Pesch, *Retracts of Graphs*, Athenaeum, Frankfurt, 1988.

[75] M. Pouzet, I.G. Rosenberg, *General metrics and contracting operations*, J. of Discrete Math. 130 (1994), 103-169.

[76] M. Pouzet, *Une approche métrique de la rétraction dans les ensembles ordonnés et les graphes*, Proceedings of the conference on infinitistic mathematics (Lyon, 1984), 59–89, Publ. Dép. Math. Nouvelle Sér. B, 85-2, Univ. Claude-Bernard, Lyon, 1985.

[77] M. Pouzet, *Operations preserving equivalence relations*, May 30, (2017), 14pp. arXiv:1705.10890v1.

[78] S. Priess-Crampe, P. Ribenboim, *Equivalence relations and spherically complete ultrametric spaces*, Comptes rendus de l'Académie des sciences. Série 1, Mathématique 320 (1995), no. 10, 1187)–1192.

[79] S. Priess-Crampe, P. Ribenboim, *Generalized ultrametric spaces* I, Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg, (1996), pp. 55.

[80] A. Quilliot, *Homomorphismes, points fixes, rétractions et jeux de poursuite dans les graphes, les ensembles ordonnés et les espaces métriques*, Thèse de doctorat d'Etat, Univ Paris VI (1983).

[81] A. Quilliot, *An application of the Helly property to the partially ordered sets*, J. Combin. Theory, serie *A*, 35 (1983), 185-198.

[82] I. Rival, *A fixed point theorem for finite partially ordered sets*, J. of Comb. theory (1976), 309–318.

[83] I. Rival, R. Wille, *The smallest order variety containing all chains*, Discrete Math. 35 (1981), 203–212.

[84] K.I. Rosenthal, *Quantales and their applications,* Pitman Research Notes in Mathematics Series, 234. Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1990. x+165 pp.

[85] F. Saïdane, *Graphes et langages : une approche métrique*, Thèse de doctorat, Lyon : Novembre 1991, *n*^{o} 206-91.

[86] J. Sakarovitch, *Elements of automata theory*, Translated from the 2003 French original by Reuben Thomas. Cambridge University Press, Cambridge, (2009).

[87] N. Sauer, *Distance sets of Urysohn metric spaces*, Canad. J. Math. 65 (2013), no. 1, 222-240.

[88] R. C. Sine, *On nonlinear contractions in sup norm spaces*, Nonlinear Analysis, 3 (1979), 885-890.

[89] P. Soardi, *Existence of fixed points of non-expansive mappings in certain Banach lattices*, Proc. A.M.S., 73 (1979), 25-29.

[90] A. Tarski, *A lattice theoretical fixed point theorem and its applications*, Pac. J. of Math. 5 (1955), 285-309.

[91] P. Urysohn, *Sur un espace métrique universel*, Bull. Sci. Math. (2) 51 (1927), 43-64.

[92] W. Ward, R. P. Dilworth, *Residuated lattices*, Trans. Amer. Math. Soc. 45 (1939), no. 3, 335-354.

Index