Springer Proceedings in Mathematics & Statistics

Cheikh Thiecoumba Gueye
Papa Ngom
Idy Diop *Editors*

Mathematics of Computer Science, Cybersecurity and Artificial Intelligence

5th Scientific Days of Doctoral School of Mathematics and Computer Sciences (S2DSMCS), Dakar, Senegal, December 20–22, 2023

Springer

Springer Proceedings in Mathematics & Statistics

Volume 468

This book series features volumes composed of selected contributions from workshops and conferences in all areas of current research in mathematics and statistics, including data science, operations research and optimization. In addition to an overall evaluation of the interest, scientific quality, and timeliness of each proposal at the hands of the publisher, individual contributions are all refereed to the high quality standards of leading journals in the field. Thus, this series provides the research community with well-edited, authoritative reports on developments in the most exciting areas of mathematical and statistical research today.

Cheikh Thiecoumba Gueye • Papa Ngom •
Idy Diop
Editors

Mathematics of Computer Science, Cybersecurity and Artificial Intelligence

5th Scientific Days of Doctoral School of Mathematics and Computer Sciences (S2DSMCS), Dakar, Senegal, December 20–22, 2023

Springer

Editors
Cheikh Thiecoumba Gueye
Departement de Mathématiques et Informatique
Université Cheikh Anta Diop
Dakar, Senegal

Papa Ngom
Mathématiques et Informatique
Université Cheikh Anta Diop
Dakar, Senegal

Idy Diop
Ecole Supérieure Polytechnique
Cheikh Anta Diop University
Dakar, Senegal

ISSN 2194-1009 ISSN 2194-1017 (electronic)
Springer Proceedings in Mathematics & Statistics
ISBN 978-3-031-66221-8 ISBN 978-3-031-66222-5 (eBook)
https://doi.org/10.1007/978-3-031-66222-5

Mathematics Subject Classification: 13P10, 47H05, 17A40, 62H30, 68T50

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The 5th Scientific Days of School Doctoral Mathematics and Computer Science Took place December 20–22, 2023, in Cheikh Anta Diop University of Dakar. They were organized by the School Doctoral Mathematics and Computer Science of Cheikh Anta Diop University of Dakar. The 5th edition received around 30 submissions, and all were reviewed by the program committee. Each paper was assigned at least to two reviewers. After highly interactive discussions, a careful deliberation, the program committee selected around 20 for presentation. The authors of accepted paper were given a week to prepare final version for these proceedings. We would like to note that these scientific days offer a prestigious and conducive forum for researchers in the fields of Mathematics, Computer Science and Telecommunications to share their most recent findings, and engage with internationally renowned colleagues and experts, and in so doing contribute to the advancement of science. The purpose of this scientific event is to stimulate research values, explore emerging challenges in science, and bring to the fore the concrete applications of mathematics and computer science on modern society. We are deeply grateful to the program committee for their hard work, enthusiasm, and conscientious efforts to ensure that each paper received a thorough and fair review. We also would like to thank Ousmane Ndiaye for writing Springer to an accelerated schedule for writing the proceedings. We also wish to heartily thank Laila Mesmoudi for her useful help, as well as sponsors of the event: Centre d'Excellence Africain en Mathématiques, Informatique et TIC (CEA-MITIC), Centre d'Excellence Africain ≪ Environnement, Santé, Sociétés ≫ (CEA-AGIR), Direction Générale du Chiffre et de la Sécurité des Systèmes d'Information (DCSSI) at Presidency of the republic of Senegal and Humboldt Chair at African Institute for Mathematical Sciences (AIMS) Senegal. Last but not least, we give thanks to all those who contributed to this event.

Dakar, Senegal
May 2024

Cheikh Thiécoumba Gueye

Scientific Days of Doctoral School of Mathematics and Computer Sciences December 20, 21, 22 , 2023—UCAD, Dakar (Senegal)

- General Chairs:
  - Cheikh Thiecoumba Gueye, Université Cheikh Anta Diop de Dakar (Senegal)
- Program co-Chairs:
  - Cheikh Thiécoumba Gueye, Université Cheikh Anta Diop de Dakar (Senegal)
  - Papa Ngom, Université Cheikh Anta Diop de Dakar (Senegal)
  - Idy Diop, Université Cheikh Anta Diop de Dakar (Senegal)

Program Committee

- Zekhnini Abdelkader, Mohammed Premier University, Oujda, Morocco
- Takoua Abdellatif, Polytechnic School of Tunisia, Tunisia
- Rabah Attia, Polytechnic School of Tunisia, Tunisia
- Abdel Malek Azizi, Mohammed I University, Morocco
- Ismaila Diouf, Université Cheikh Anta Diop de Dakar, Senegal
- Souidi El Mamoun, Mohammed V University, Morocco
- Nouzha El Yacoubi, Mohammed V University, Morocco
- Ibrahima Fall, Université Cheikh Anta Diop de Dakar, Senegal
- Cheikh Ahmadou Bamba Gueye, Université Cheikh Anta Diop de Dakar, Senegal
- Cheikh Thiecoumba Gueye, Université Cheikh Anta Diop de Dakar, Senegal
- Abdelfattah Haily, Chouaib Doukkali University, El Jadida, Morocco
- Jerrari Idriss, Mohammed Premier University, Oujda, Morocco
- Mohamed Ben Maaouia, Université Gaston Berger de Saint-Louis, Senegal
- Abdenacer Makhlouf, Université de Haute Alsace, France
- Candido Martin Gonzalez, University of Malaga, Spain
- Edgar Martinez Moro, Universidad de Valladolid, Spain
- Abdoulaye Mbaye, Université Cheikh Anta Diop de Dakar, Senegal
- Laila Mesmoudi, Université Cheikh Anta Diop de Dakar, Senegal

- Boudi Nadia, Mohammed V University, Morocco
- Babacar Mbaye Ndiaye, Université Cheikh Anta Diop de Dakar, Senegal
- Ousmane Ndiaye, Université Cheikh Anta Diop de Dakar, Senegal
- Daouda Ndoye, Université Cheikh Anta Diop de Dakar, Senegal
- Papa Ngom, Université Cheikh Anta Diop de Dakar, Senegal
- Abdellatif Rochdi, University of Hassan II of Casablanca, Morocco
- Idrissa Sarr, Université Cheikh Anta Diop de Dakar, Senegal
- Djiby Sow, Université Cheikh Anta Diop de Dakar, Senegal
- Amadou Tall, Université Cheikh Anta Diop de Dakar, Senegal
- Ioannis CHALKIADAKIS, institut des sciences complexes Paris Île De France (ISCPIF), France
- Salma Mesmoudi, Université Paris1 panthéon-Sorbonne, France
- Charles Tijus, université Paris 8, France
- Paul Bourgine Complex Systems Digital Campus, France
- Waldo Cancino Sao Paulo University, Brasil

Organization Committee

- Idy Diop, Cheikh Anta Diop University, Senegal
- Ibra Diuom, Cheikh Anta Diop University, Senegal
- James Tamgno Kouawa, Multinational Higher School of Telecommunications, Senegal
- Cheikh LO, Cheikh Anta Diop University, Senegal
- Moustapha Mbaye, Cheikh Anta Diop University, Senegal
- Ousmane Ndiaye, Cheikh Anta Diop University, Senegal
- Demba Sow, Cheikh Anta Diop University, Senegal
- Marc Momar Tall, Cheikh Anta Diop University, Senegal

Contents

Part I
Invited Talks

Chapter 1
Mathematics and Computer Science in the Information Revolution

Diaraf Seck

Abstract This text is a presentation given during the doctoral sessions of the Ecole Doctorale Mathématiques et Informatique at Cheikh Anta Diop University. Our aim was simply to arouse the curiosity of those present, especially doctoral students in Mathematics, Computer Science, and Telecommunications. We set ourselves the goal of demonstrating the close link between Mathematics and Computer Science. After an overview of the background, we went on to look at some of the research topics that have interested us over the last 10 years, and which have the particularity of having a close link with Computer Science, with Mathematics as the main tool at our disposal.

Keywords Linear Programming · Optimal Transport · Graphs · Optimization · Machine Learning · Deep Learning · Information theory · Information coding · Cryptography · Computer Science · Quantum model

Introduction

Boolean algebra, or Boolean calculus, is the part of mathematics concerned with an algebraic approach to logic, seen in terms of variables, operators, and functions on logical variables, which allows one to use algebraic techniques for dealing with two-valued expressions in the calculus of propositions. It was launched in 1854 by the British mathematician George Boole. Boolean algebra has many applications in computer science and in the design of electronic circuits.

It was first used for telephone switching circuits by Claude Shannon.

D. Seck (✉)
Laboratoire des Mathématiques de la Décision et d'Analyse Numérique, Dakar Fann, Sénegal

Ecole Doctorale de Mathématiques et Informatique U.C.A.D. Dakar, Dakar, Sénegal
e-mail: diaraf.seck@ucad.edu.sn

C. T. Gueye et al. (eds.), *Mathematics of Computer Science, Cybersecurity and Artificial Intelligence*, Springer Proceedings in Mathematics & Statistics 468,
https://doi.org/10.1007/978-3-031-66222-5_1

Claude Elwood Shannon (April 30, 1916 in Michigan–February 24, 2001 in Massachusetts) was an American electrical engineer and mathematician. He is one of the fathers, if not the founding father, of Information theory

Information theory, unspecified, is the common name for Shannon's information theory, which is a theory using probabilities to quantify the average information content of a set of messages, whose computer coding satisfies a statistical distribution that we think we know. This field has its scientific origins with Claude Shannon who was its founder with his article *A Mathematical Theory of Communication* published in 1948. Before ending this brief introduction, let us mention some meanings on information coding, computer science, and cryptography:

- Information coding concerns the means of formalizing information in order to be able to manipulate it, store it, or transmit it. It is not interested in the content, but only in the form and size of the information to be encoded.
- Computer science is a field of scientific, technical, and industrial activity concerning the automatic processing of digital information by the execution of computer programs hosted by electrical–electronic devices: embedded systems, computers, robots, automata, etc.
- Cryptography is one of the disciplines of cryptology focused on protecting messages (ensuring confidentiality, authenticity, and integrity) often using secrets or keys. It is distinguished from steganography which causes a message to pass unnoticed within another message, while cryptography makes a message supposedly unintelligible to anyone other than the appropriate person. It has been used since ancient times, but some of its most modern methods, such as asymmetric cryptography, date from the late twentieth century.

  Cryptography is a writing technique where an encrypted message is written using secret codes or encryption keys. Cryptography is mainly used to protect a message considered confidential.

The chapter is organized as follows: In the next section, we are going to do an overview about the birth of theoretical computer science. Section "Dialogue Between Mathematics and Computer Science, Some Works, and Illustrations" is devoted to a dialogue between Mathematics and Computer Science. And some illustrations coming from our own experiences will be introduced. And in the last section, we will discuss about the information revolution.

Birth of Theoretical Computer Sciences [8]

The Turing machine is an abstract machine model introduced in 1936 by the English researcher Alan Turing in a seminal article entitled *On Computable Numbers, with an Application to the Entscheidungsproblem*, in which he proposed an answer to a question posed 8 years earlier by the famous mathematician David Hilbert. This is the problem of decidability (in German "Entscheidungsproblem"), in essence: Is there an algorithm that decides whether a proposition stated in a logical system is valid or not?

Why has the calculation model proposed by Alan Turing become an essential tool in fundamental computer science and mathematics? His machine is in fact widely used in computability theory, complexity theory, or approximation theory. The main reason is its extreme simplicity which made it possible to establish results that are undoubtedly much more difficult to demonstrate with less rudimentary models. This machine is in fact the simplest model that can be designed and which satisfies the informal but universal criteria which characterize an algorithm (determinism, discretion, finitude, generality, etc.)

It should be noted that this model, like so many others, was created at a time when computers as we know them today did not exist, so it is above all an abstract tool. The generalization of computers in the 1950s and 1960s gave birth to the Register Addressable Memory (ram) model, closer to their physical and logical architecture. What is remarkable is that all these models were equivalent, and what can be calculated with model A can be calculated with model B and vice versa.

It is commonly accepted that any abstract model of calculation respecting the informal conditions on which scientists agree to speak of an algorithm (except for quantum calculation that defines a new paradigm) results in a model equivalent to the previous ones.

It should be kept in mind that the Turing machine is a universal model of calculation and that it can calculate anything that any physical computer can calculate (no matter how powerful it is). Conversely, what it cannot calculate cannot be calculated by a computer either. It therefore summarizes in a striking manner the concept of computer and constitutes an ideal support for reasoning around the notion of calculation or demonstration algorithm.

Complexity theory is based on methods for solving decision problems that are unambiguously described or generated in an algorithm. An algorithm is of polynomial complexity if there exists a polynomial P such that the number of elementary instructions carried out during its execution on any data of size n is at most $P(n)$.

Definition 1.1 (Decision Problem) A decision problem is a problem whose solution is formulated in yes/no terms.

Example 1.1 Given a graph $G = (X, E)$, does there exist a path of length $\leq L$?

Definition 1.2 (Class P) A problem is of polynomial complexity if there exists a polynomial complexity algorithm solving it. That is, if there exists an integer k for which the resolution time is $\mathcal{O}(n^k)$ with n the size of the instance.

Definition 1.3 (NP-Hard) A problem that does not admit an algorithm for its solution in polynomial time is called an NP-hard problem.

Definition 1.4 (NP Class) This is the abbreviation for Nondeterministic Polynomial time. This class contains all the decision problems of which we can associate with each of them a set of potential solutions (from cardinal to worst exponential) such that we can check in polynomial time if a potential solution satisfies the question asked.

Definition 1.5 (NP-Complete) A problem is said to be NP-complete if it is in NP.

In complexity theory, an NP-complete problem or NPC problem (i.e., a complete problem for the NP class) is a decision problem verifying the following properties:

- It is possible to verify a solution efficiently (in polynomial time); the class of problems verifying this property is denoted NP.
- All the problems of the NP class are reduced to this one via a polynomial reduction; this means that the problem is at least as difficult as all other problems in the NP class.
- An NP-hard problem is a problem that satisfies the second condition and therefore may be in a larger problem class and therefore more difficult than the NP class.

Remark 1.1

- Although we can quickly verify any proposed solution of an NP-complete problem, we do not know how to find one efficiently.
- This is the case, for example, of the traveling salesman problem or the knapsack problem.

Remark 1.2

- All known algorithms for solving NP-complete problems have an exponential execution time depending on the size of the input data in the worst case and are therefore unusable in practice even for instances of moderate size.
- The second property of the definition implies that if there exists a polynomial algorithm to solve any NP-complete problem, then all problems of the NP class can be solved in polynomial time.
- Finding a polynomial algorithm for an NP-complete problem or proving that one does not exist would answer whether $P = NP$ or $P \neq NP$, an open question that is among the most important unsolved problems in mathematics to date.

Dialogue Between Mathematics and Computer Science, Some Works, and Illustrations

We are going to introduce very quickly three of eighteen Smale's problems:

Problem 3 $P = NP$?
Smale said "I sometimes think of this problem as a gift from computer scientists to mathematicians. It may help to put it in a form that looks more like traditional math."

Problem 9: On the Linear Programming Problem
Is there a polynomial time algorithm on real numbers that decides whether the linear system of inequalities $Ax \geq b$ has a solution?

Problem 18: On the Limits of Intelligence
What are the limits of intelligence, both artificial and human?

For more details about these questionings, we invite the reader to see [7].

Social Network Problems [1, 2]

A social network is longtime used to represent the interactions between the individuals (or the organizations) in different contexts. The network permits to represent the individuals (or organizations) and ties among them (the individuals or the organizations).

Formally, a social network is modeled by an undirected or a directed graph $G(V, E)$, with $V = (v_1, ..., v_n)$ the set of nodes and $E \subseteq (V \times V)$ the set of edges or links. The nodes represent the individuals (or organizations) and the edges the relations between the individuals (or organizations).

Influence maximization is to find a subset k-nodes (i.e., seeds set) in a social network that could maximize the influence spread. Mathematically we can define the problem as follows: Find S_k^* such as

$$S_k^* = argmax_{S \subseteq V, |S|=k} \sigma(S),$$

where $\sigma(S)$ is an activation function which gives an integer. D. Kempe, J. Kleinberg, and E. Tardos in 2003 showed that this problem is NP- hard. It is very difficult to choose a small k-nodes (seeds set) which maximize the influence spread.

Remark 1.3 The first works in the influence maximization problem are proposed by Domingos and Richardson in 2001 as an algorithmic problem. They modeled the problem using Markov random fields and proposed heuristic solutions.

Spanning Graph and Geodesics One can propose a new approach to determine the seeds set. In the influence maximization problem in the social networks, the idea is to influence the maximum of nodes. In the propagation models, an active node never becomes inactive again. So the feedback in the active node is not necessary. Thus, the idea is to delete the cycles when we determine the seeds. So to determine the seeds, if a node u is chosen, then we eliminate all communication between them and all nodes $v \notin neighbor(u)$. So all transitivity are eliminated in the graph before determining the seeds. Thus one determines a particular acyclic spanning graph from the initial graph. Yet the closeness centrality measure of the node u gives the distance between it and the other nodes. The node that has the smallest closeness centrality can be considered as the central node of the network. The construction of the acyclic spanning graph begin with this node. And it is therefore possible to present two algorithms to find the acyclic spanning graph from a graph according to its connectivity. The input (of the algorithms) is a graph and the output an acyclic spanning graph which will maximize the influence spread.

The centrality is a fundamental concept in network analysis. The centrality measure in a graph gives a strict indication of how connected a node in the network, this apply to social networks, information networks, biological networks, etc. Several centrality measures have been proposed, and they do not have the same importance in the network. One can cite the degree centrality, the closeness centrality, the betweenness centrality, the degree discount centrality, Diffusion Degree, etc. For the construction of the acyclic spanning graph to maximize the influence spread, one can use the centrality closeness measure which is a concept that can be naturally defined in the metric space where the notion of distance of an element and the space is defined. In graph theory, the closeness centrality for a node v is the sum of geodesic distances to all other nodes of v accessible from the latter.

To calculate this measure, one can follow the next steps:

- Calculate the shortest path, which may also be known as the "geodesic distance," between the node u and all other node different of u.
- Calculate the sum of all the geodesic distances.

$$Cc(u) = \sum_{v \in V, v \neq u} d(u, v)$$

with $d(u, v)$ the geodesics distance between u and v.

Optimal Mass Transport for Activities Location Problem [3]

- The activities location problem is often formulated as a Quadratic Assignment Problem (QAP) by Koopmans and Beckmann (in 1957), which assigns n activities to n locations while minimizing the total cost location.
- The QAP is known to be NP-complete
- The question would be as follows:

  What is the optimal way to locate activities in the transportation network? How are the locations of clinics within a hospital decided? How to locate optimally administrative services?
- To address these challenges, there is the most challenging combinatorial optimization problem.
- The main idea is as follows:

  Locations k and l are separated by a distances of d_{kl}. On the other hand, entities i and j must exchange quantities of a given product f_{ij}. The cost of assigning i to k is c_{ik}, but an assignment also induces a product routing cost which is assumed to be proportional to the quantities of product to be exchanged and to the distance that separates the entities, i.e., $f_{ij}d_{kl}$.

The mathematical formulation of the Activities Location Problem ($\mathcal{ALP}$) is given as follows. Let:

- $X = \{1, ..., n\}$ the set of activities and $Y = \{1, ..., n\}$ the set of potential sites for new activities
- $F = (f_{ij})_{n\times n}$ the matrix of flows from activity i to activity j
- $D = (d_{kl})_{n\times n}$ the matrix of distances from site k to site l
- $C = (c_{ik})_{n\times n}$ the cost of assigning activity i to site k, independent of other locations
- P_n the set of all permutations of $\{1, ..., n\}$ in $\{1, ..., n\}$
- π_{ik} the assignment of activity i to site k

The activity location problem can be modeled as a QAP, which is to find the minimum cost assignment (location) of n activities to n locations. For the $\mathcal{ALP}$, the quadratic assignment formulation is shown in Eqs. (1.1)–(1.4)

$$(\mathcal{ALP}): \quad \min_{\pi \in P_n} \quad \sum_{i=1}^{n}\sum_{k=1}^{n} c_{ik}\pi_{ik} + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} \Gamma_{ijkl}\pi_{ik}\pi_{jl} \tag{1.1}$$

$$s.t.$$

$$\sum_{i=1}^{n} \pi_{ik} = 1, \qquad \forall k \in \{1, \ldots, n\} \tag{1.2}$$

$$\sum_{k=1}^{n} \pi_{ik} = 1, \qquad \forall i \in \{1, \ldots, n\} \tag{1.3}$$

$$\pi_{ik} \in \{0, 1\}, \qquad \forall\, i, k \in \{1, \ldots, n\}, \tag{1.4}$$

where

$$\pi_{ik} = \begin{cases} 1, & \text{if activity } i \text{ is located at zone } k, \\ 0, & \text{otherwise.} \end{cases} \tag{1.5}$$

$\Gamma_{ikjl} = f_{ij}d_{kl}$ is the cost of locating activity i at location k and activity j at location l. Γ_{ikjl} in Eq. (1.1) is a cost variable representing the combination of quantitative and qualitative measures in $\mathcal{ALP}$ models. Equation (1.2) ensures that each location is assigned to only one activity. Equation (1.3) ensures that each activity is assigned to only one physical location.

Discrete Optimal Transport Formulation of the Above Problem
Assume we are given discrete measured metric spaces $X = \{x_1, \ldots, x_n\}$ and $X = \{y_1, \ldots, y_n\}$ with metrics d_X and d_Y, respectively, and probability measures μ_X and μ_Y, respectively.

In the case where the marginals are given by convex combinations of Dirac measures supported in atoms $\{x_i\}_{i=1,\ldots,n} \subset X$ and $\{y_k\}_{k=1,\ldots,n} \subset Y$, respectively, i.e.,

$$\mu_X = \sum_{i=1}^{n} \mu_X(x_i)\delta_{x_i}, \quad \mu_Y = \sum_{k=1}^{n} \mu_Y(y_k)\delta_{y_k},$$

where δ_{x_i} and δ_{x_k} represents the Dirac mass at the points x_i and x_k, respectively. The *transport plan* between μ_X and μ_Y is described by matrix $(\pi(x_i, y_k))_{1\le i,k\le n}$ satisfying the following set of constraints:

$$\Pi(\mu_X, \mu_Y) = \left\{ \begin{array}{l} \pi \in \mathbb{R}^{n\times n} : \displaystyle\sum_{i=1}^{n} \pi(x_i, y_k) = \mu_Y(y_k), \text{ for all } 1 \le k \le n \\ \displaystyle\sum_{k=1}^{n} \pi(x_i, y_k) = \mu_X(x_i), \text{ for all } 1 \le i \le n \\ 0 \le \pi(x_i, y_k) \le 1. \end{array} \right\}. \tag{1.6}$$

In constraints (1.6), the first equation ensures that each location is assigned one and only one activity and the second equation ensures that each activity is assigned to one and only one physical location.

The objective function of the problem is thus given by

$$\left\{ \begin{array}{l} \displaystyle\inf_{\pi\in\Pi(\mu_X,\mu_Y)} L(\pi), \\ L(\pi) = \displaystyle\sum_{i=1}^{n}\sum_{k=1}^{n} c(x_i, y_k)\pi(x_i, y_k) + \\ \displaystyle\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} \Gamma_{X,Y}\pi(x_i, y_k)\pi(x_j, y_l), \end{array} \right. \tag{1.7}$$

where $c(x_i, y_k)$ is the cost to transporting a unit mass from $x_i \in X$ to $y_k \in Y$ independently and the cost function $\Gamma_{X,Y}$ is constructed from the assumption that if a map pairs $x_i \to y_k$ and $x_j \to y_l$, then the distance between x_i and x_j on X should match the distance between y_k and y_l on Y, where

$$\Gamma_{X,Y}(x_i, y_k, x_j, y_l) := |d_X(x_i, x_j) - d_Y(y_k, y_l)|^p; \; p \in [1, \infty)$$

is the distortion.

Internet: Another Complex Graph

How to evaluate the cost of transferring a mass distribution to an other mass distribution?

- Let the graph $G = (V, E)$ be a discrete extension of the Ricci curvature which uses the concept of optimal transport on the graph.
- $d(x, y)$ defines the distance between each of couple vertices $(x, y) \in V \times V$.

- This distance can be the number of minimal jumps, a weighting minimum on the distance, or any distance matrix.
- Consider a mass distribution $\mu(x)$, $\nu(y)$ over the nodes of a graph, and we want to transfer the mass that initially exists to another distribution.
- This distribution can be seen as data to be moved between a source and one of the destinations in a network

$$\sum_{x\in V}\mu(x)=1,\quad \sum_{y\in V}\nu(y)=1,$$

$$\left\{\begin{array}{l} \omega^*(\mu,\nu)=argmin_\omega \sum_{x,y\in V}\omega(x,y)d(x,y) \\ \sum_{y\in V}\omega(x,y)=\mu(x),\ \text{for all}\ \ x\in V \\ \sum_{x\in V}\omega(x,y)=\nu(y),\ \text{for all}\ \ y\in V \end{array}\right\}.$$

$\omega(x,y)$ is the mass to be transported from x to y, and $d(x,y)$ is the transport cost of one unity from x to y.

Definition 1.6 (Y. Ollivier, [6]) Let (X,d) be a metric space with a random walk m. Let $x,y\in X$ be two distinct points. The Ricci curvature of (X,d,m) in the direction (x,y) is

$$\kappa(x,y):=1-\frac{\mathcal{C}(m_x,m_y)}{d(x,y)},$$

$$\mathcal{C}(\nu_1,\nu_2):=\inf_{\xi\in\Pi(\nu_1,\nu_2)}\int_{X\times X}d(x,y)d\xi(x,y),$$

where $\mathcal{C}(\nu_1,\nu_2)$ is the L^1 transportation between ν_1 and ν_2, and $\Pi(\nu_1,\nu_2)$ is the set of measures on $X\times X$ projecting to ν_1 and ν_2. Intuitively $d\xi(x,y)$ represents the mass that is sent from x to y, hence the constraint on the projection of ξ, ensuring that the initial measure is ν_1 and the final measure is ν_2.

$\kappa(x,y)=1-\frac{\mathcal{C}(\mu(x),\nu(y))}{d(x,y)}, \mathcal{C}(\mu,\nu)=\sum_{x,y\in V}\omega^*(\mu,\nu)d(x,y).$

Numerical Approach of Network Problems in Optimal Mass Transportation [5]

$$\min\Big\{\int_\Omega dist(x,\Sigma)\,\mu(x)\,dx:\ \Sigma\subset\Omega\Big\}.$$

Σ is an admissible network with a finite length equal to $L>0$.

A possible discrete formulation of the problem is to introduce permutations σ.

Let us take a permutation σ defined on $\mathcal{S}_m := \{1, \cdots, m\}$ such that $\sigma(k) \neq k$.

Let $\sigma : \mathcal{S}_m \to \mathcal{S}_m$ $x_k \in \mathbb{R}^n$, $k \in \mathcal{S}_m$, $T(x_k) := y_{\sigma(k)}$, x_k is a data, $T : \mathbb{R}^n \to \mathbb{R}^n$ is the map transport $\mu(T^{-1}(B)) = \nu(B)$, $\forall B \subset \mathbb{R}^n$, bounded, μ and ν are measures defined on $\mathbb{R}^n$ with equal total mass, and $\|.\|$ is the Euclidean norm.

Using the approximation formulae $\int_{\overline{\Omega}} dist(x, \Sigma)\, dx \sim \sum_{k=1}^{m} d(x_k, y_k)$, the aim is to find all the points y_k minimizing the following optimization problem:

$$\min \sum_{k=1}^{m} \|x_k - y_k\|^2 \tag{1.8}$$

and such that

$$\inf_{\sigma} \sum_{k=1}^{m-1} \|y_{\sigma(k)} - y_{\sigma(k+1)}\|^2 \leq L.$$

This problem is equivalent to looking for the points y_k minimizing the following one:

$$\min_{\sigma} \sum_{k=1}^{m} \|x_k - y_{\sigma(k)}\|^2$$

under the constraints

$$\sum_{k=1}^{m-1} \|y_k - y_{k+1}\| \leq L.$$

For a scenario in $\mathbb{R}^n$, $n \geq 2$, if we consider m points, the number of programs to be solved becomes m^m. We leave the reader to verify that for:

$m = 3$ points, we solve 27 programs.
$m = 4$ points, we solve 256 programs.
...

How is it possible to overcome this issue?

A New Way to Study Linear Programming [4]

We call *linear programming problem* any problem which can be stated as

$$\max \sum_{j=1}^{n} c_j . x_j \quad c_j \in \mathbb{R}; n \in \mathcal{N}^*$$

$$\text{Constraints} \begin{cases} \sum_{j=1}^{j=n} a_{i,j}.x_j \leq b_i \\ \quad x_j \quad \geq 0 \end{cases} \tag{1.9}$$

$$a_{i,j}\,,\,b_i \in \mathbb{R};\, i = \{1\ldots m\} \subset \mathcal{N},$$
$$j = \{1\ldots n\} \subset \mathcal{N}.$$

Recall that convex polyhedron on E is a set P:

$$P = \{X \in \mathbb{R}^n \ : \ A.X \ \leq \ B\}, \tag{1.10}$$

where $A : E \longrightarrow \mathbb{R}$ is a linear application ($m \in \mathcal{N}$; if $m = 0$, $P = E$), $B \in \mathbb{R}^m$, and the inequality $A.X \leq b$ must be considered step by step (row by row) in $\mathbb{R}^m$ $(A.X)_i \leq b_i \quad \forall i \in \{1, \cdots, m\}$.

In case of equalities as $C.X = d$, it is always possible to come back to the form (1.10) if we replace an equality by two opposite inequalities $C.X \leq d$ and $-C.X \leq -d$. We call *polytop* a convex and bounded polyhedron.

Remark 1.4 The Hann–Banach theorems play a fundamental role in this work.

The Sketch of the Algorithm

In a maximization problem (a problem in which the purpose is to maximize the objective function, as it is the case in problem of the form (1.9)).

Suppose $H : f(X) = \alpha$ and K is the initial polytope defined by the set of constraints (1.9). If $H \cap K = \emptyset$, then we state $\alpha := \dfrac{\alpha}{2}$. That is to say that we split the gap between 0 and α into two parts. It is the main idea of the algorithm.

Now suppose that we have two values of α, both α_0 and α_1 such that $H : f(X) = \alpha_0$ and $K \cap H \neq \emptyset$ and $H : f(X) = \alpha_1$ and $K \cap H = \emptyset$. Then we take a new value of $\alpha = \dfrac{\alpha_0 + \alpha_1}{2}$.

Then, two cases are possible:

1. $H : f(X) = \alpha$ and $K \cap H = \emptyset$, then put $\alpha_1 := \alpha$.
2. $H : f(X) = \alpha$ and $K \cap H \neq \emptyset$, then put $\alpha_0 := \alpha$.

And so on.

Remark 1.5 The notation $i := i + 1$ which has no sense in mathematical field is a notation of programming, meaning that the value of the memory cell i, at which the value 1 is added, is affected to the memory cell i.

In this algorithm the chosen hyperplane is defined by

$$H = \{x \in \mathbb{R}^n \ / \ \sum_{i=1}^{n} c_i x_i = \alpha\}, \tag{1.11}$$

where $\sum_{i=1}^{n} c_i x_i$ is the objective function (the one to maximize in our purpose) of the LP.

Remark 1.6 From a geometric point of view, at each step of the algorithm, the new hyperplane is obtained from the previous one by a translation. In effect, we search the vectors which produce the space H defined by (1.11), and then we search an unitary vector $\overrightarrow{v}$ as vectorial step.

A Key Subproblem
In this algorithm there is a major difficulty in the test for vacuity—or not—of the intersection between the hyperplane and the constraint's polytope :

$$H \bigcap K = \emptyset \quad ? \tag{1.12}$$

The Number of General Steps
Suppose that the second part starts with $f(x) = \alpha, \alpha \in \mathcal{N}^*$, and the number of steps in order to reach a number in a unitary interval which contains solution will be $\log_2(\alpha)$ due to the dichotomic choices.

Lemma 1.1 *The number of steps in order to reach a number in a unitary interval which contains solution is* $\lceil \log_2(\alpha) \rceil$.

In each step the considered space is split into two parts. In order to compute one iteration of the second part of the algorithm, it is necessary to compute also one of the first parts in order to test if the actual constraint polytope is both avoid or not.

The Number of Steps Due to the Accuracy
Suppose now the calculus to be in the unitary interval and an accuracy gap of $10^{-\beta}, \beta \in \mathcal{N}^*$. We can consider the unitary interval with 10^{β} subintervals, and, in the worst case, it is necessary to compute $a = \lceil \log_2 10^{\beta} \rceil$ steps in order to reach this subinterval.

Lemma 1.2 *The maximal number of additional steps due to an accuracy of* $10^{-\beta}$ *is*

$$a = \lceil \log_2(10^{\beta}) \rceil, \ \beta \in \mathcal{N}^*.$$

Then, we have the following theorem:

Theorem 1.1 *The worst-case steps complexity of this algorithm is*

$$\mathcal{O}\left((a + \lceil \log_2 \alpha \rceil).n.m\right), \quad a, \alpha \in \mathbb{R}^*. \tag{1.13}$$

As seen previously in the classical algorithms, it is necessary to consider the bit complexity of our algorithm. This complexity depends on the size of numbers of the problem:

Lemma 1.3 *The size of the numbers expressed in bits in the problem is majored by*

$$L = \lceil \log_2 (\max_{i,j} \{|a_{i,j}|, |b_i|, |c_j|\}) \rceil,$$

$$i \in \{1, \ldots, m\},\ j \in \{1, \ldots, n\}.$$

But in each iteration of the algorithm, it is possible to use this complexity. From both Theorem 1.1 and Lemma 1.3, we deduce the following result:

Theorem 1.2 *The total bit-complexity of our algorithm is increased by*

$$\mathcal{O}\left((a + \lceil \log_2 \alpha \rceil).L.n.m\right), \quad a, \alpha \in \mathcal{R}^*. \tag{1.14}$$

Remark 1.7 The value of a depends on the tolerance gap.[1]

Remark 1.8 The same result can be obtained by a direct reasoning. If $\alpha \in \mathcal{N}^*$ and the wanted accuracy is 10^{β}, the maximal number of steps in a dichotomic way will be $\lceil \log_2 (\alpha.\beta) \rceil$. Effectively, in this case, it is sufficient to consider that the initial interval $[0, \alpha]$ is split into $(\alpha.\beta)$ cells. It is also the meaning of the coefficient $(a + \lceil \log_2 \alpha \rceil)$ in Eq. (1.13) with $a = \lceil \log_2 10^{\beta} \rceil$.

New codes and additional numerical tests need to be realized in order to be able to conclude if this new way brings a complete solution to this problem or not.

Information Revolution

Artificial Intelligence, Machine Learning, and Deep Learning

Artificial Intelligence

Artificial intelligence (AI) is a process of imitating human intelligence that relies on the creation and application of algorithms executed in a dynamic computing environment. Its goal is to enable computers to think and act like human beings. To achieve this, three components are necessary:

- Computer systems
- Data with management systems
- Advanced AI algorithms (codes)

To get as close to human behavior as possible, artificial intelligence needs a high amount of data and processing capacity.

[1] Which will be stated by the user of the algorithm. And as in the industrial world more you want accuracy, more you pay.

Today, AI is a thriving field with many practical applications and active research topics. We look to intelligent software to automate routine labor, understand speech or images, make diagnoses in medicine, and support basic scientific research.

Right from the start, the pioneering researchers working on artificial intelligence set themselves the goal of solving problems that were difficult for human beings but relatively simple for computers, problems that could be described by a list of formal mathematical rules.

The true challenge to artificial intelligence consists in solving the tasks that are easy for people to perform but hard to describe formally, meaning problems that we can solve intuitively and that we can sense such as speech and facial recognition.

Machine Learning

A machine learning algorithm is an algorithm that is able to learn from data. But what do we mean by learning? Mitchell (in 1997) provides the definition “A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.” One can imagine a very wide variety of experiences E, tasks T, and performance measures P.

We can learn many areas from mathematics, but here we want to focus on the mathematics of Machine Learning, i.e., the mathematics behind Machine Learning algorithms: Linear Algebra, Vectorial and Matrices Calculus, Probability and Statistics, Optimization, Analytic Geometry, etc. Let us note that: *Analytic geometry is an approach to geometry in which objects are described by equations or inequalities using a coordinate system. It is fundamental for physics and computer graphics. In analytical geometry, the choice of a reference is essential. All objects will be described relative to this referential (benchmark).*

Deep Learning

The simple machine learning algorithms work very well on a wide variety of important problems. However, they have not succeeded in solving the central problems in AI, such as recognizing speech or recognizing objects. The development of deep learning was motivated in part by the failure of traditional algorithms to generalize well on such AI tasks.

Many machine learning problems become exceedingly difficult when the number of dimensions in the data is high. This phenomenon is known as the curse of dimensionality. Of particular concern is that the number of possible distinct configurations of a set of variables increases exponentially as the number of variables increases.

- Local Constancy and Smoothness Regularization: In order to generalize well, machine learning algorithms need to be guided by prior beliefs about what kind of function they should learn. Previously, we have seen these priors incorporated as explicit beliefs in the form of probability distributions over parameters of the model. More informally, we may also discuss prior beliefs as directly influencing the function itself and only indirectly acting on the parameters via their effect on the function.

  Additionally, we informally discuss prior beliefs as being expressed implicitly, by choosing algorithms that are biased toward choosing some class of functions over another, even though these biases may not be expressed (or even possible to express) in terms of a probability distribution representing our degree of belief in various functions.

  Among the most widely used of these implicit "priors" is the smoothness prior or local constancy prior. This prior states that the function we learn should not change very much within a small region.

  Many simpler algorithms rely exclusively on this prior to generalize well, and as a result they fail to scale to the statistical challenges involved in solving AI-level tasks.

Manifold learning: An important concept underlying many ideas in machine learning is that of a manifold. A manifold is a connected region. Mathematically, it is a set of points, associated with a neighborhood around each point. From any given point, the manifold locally appears to be a Euclidean space. In everyday life, we experience the surface of the world as a 2D plane, but it is in fact a spherical manifold in 3D space.

There are many other topics that involve in Deep Learning. Let us quote some of them below:

- Deep Feedforward Networks, Regularization for Deep Learning, Optimization for Training Deep Models, and Convolutional Networks
- Convolutional Networks and Sequence Modeling: Recurrent and Recursive Nets and Linear Factor Models
- Autoencoders and Representation Learning
- Monte Carlo Methods, Confronting the Partition Function, Approximate Inference, Deep Generative Models, etc.
- Information Geometry
- Stochastic Analysis
- Dynamical Systems
- Partial Differential Equations

Quantum Model

The existence of the laser and the transistor along with the possibility of producing images of the interior of our body thanks to magnetic resonance imaging (MRI) or even geolocating to the nearest meter on the surface of the Earth using a GPS..., we owe all these scientific discoveries to quantum physics and Albert Einstein. The latter helped advance the discipline in the first part of the twentieth century. This is when quantum physics began to emerge, when scientists realized that physics alone could not explain what happens at the microscopic level.

To understand how the world works at the atomic scale, a new approach was needed: quantum physics. This allowed great advances, from the 1930s to the end of the 1980s. And then research took a further step by taking on the challenge of creating quantum computers, that is to say machines with increased computing power thanks to new technology. We can date this back to 1982. That year, American physicist Richard Feynman theorized "quantum simulators," a new type of computer.

In a classic computer, operation is done using "bits" which can take only two values: 0 or 1. This can be represented by a switch which is either on or off, a door open or closed. Qubits can be either 0, 1, or a combination of both at the same time. And this small detail of superposition of states is not one of them: It allows the quantum computer to carry out calculations much more quickly than traditional machines.

Quantum computing is therefore capable of solving difficult problems that would be impossible to solve with computers currently sold commercially. For example, imagine you have a locked box and you need to find the right key to open it. If we try one key at a time, it would take a long time. But a quantum computer can try all the keys at once, making the search much faster.

But all this is still very much in the realm of theory.

References

1. I. Gaye, G. Mendy, S.Ouya, D. Seck, Spanning graph for maximizing the influence spread in Social Networks, 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining
2. I. Gaye, G. Mendy, S.Ouya, D. Seck, New centrality measure in Social Networks based on Independent Cascade (IC) model, 2015 3rd International Conference on Future Internet of Things and Cloud
3. Mamadou Koné, Babacar M. Ndiaye and Diaraf Seck, Optimal mass transport for activities location problem accepted for publication in Nonlinear Analysis, Geometry and Applications - Proceedings of the First NLAGA-BIRS Symposium, Dakar, Senegal, June 24–28, 2019 edited by Birkhauser in 2020.
4. I. Lavallée, B. M. Ndiaye, D. Seck, A new way in linear programming in Proceedings of International on Logistics, Informatics and Service Science, Beijing, China vol. 2 pp 220–227 2011
5. L. Ndiaye, B. M Ndiaye, P. Mendy, D. Seck, Numerical Approach of Network Problems in Optimal Mass Transportation, in Applied Mathematics in 2012
6. Y. Ollivier, Ricci curvature of Markov chains on metric spaces, J. Funct. Anal. (2009)

7. Smale's problems are a list of eighteen unsolved problems in mathematics proposed by Steve Smale in 1998 and republished in 1999: Smale, Steve (1998). "Mathematical Problems for the Next Century". Mathematical Intelligencer. 20 (2): 7–15. CiteSeerX 10.1.1.35.4101. https://doi.org/10.1007/bf03025291.S2CID1331144. Smale, Steve (1999). "Mathematical problems for the next century". In Arnold, V. I.; Atiyah, M.; Lax, P.; Mazur, B. (eds.). Mathematics: frontiers and perspectives. American Mathematical Society. pp. 271–294. ISBN 978-0-8218-2070-4.
8. A. Turing, On computable numbers, with an application to the Entscheidungsproblem (Proc. Lond. Math. Soc., series 2 vol. 42 (1937), pp. 230–265) - A correction (ibid. vol. 43 (1937), p. 544–546)

Chapter 2
NLP and Some Research Results in Senegal

Samba Ndiaye

Abstract This chapter is an attempt to popularize science. It relies heavily on the INRIA seminar (https://fidle.cnrs.fr/w3/). Its objectives are on one hand to present NLP to PhD students and other researchers and on the other hand to take stock of the research done in Senegal in the field of NLP. We first introduce the basic definitions and concepts of NLP such as bags of words, TF-IDF, etc., before presenting new products such as LLMs and multimodal models. Finally, we have published publications obtained over the past five years by our research group in the field of NLP, PhDs, and master's thesis defended (Kandé et al., icetas.etssm.org; Kandé et al., FWLSA-score: French and Wolof lexicon-based for sentiment analysis, 2019; Kandé et al., Vector space model of text classification based on inertia contribution of document, 2019; Kandé et al., A novel term weighting scheme model, 2018; Kandji and Ndiaye, Design and realization of an NLP application for the massive processing of large volumes of resumes, 2022; Samb et al., Improved bilingual sentiment analysis lexicon using word-level trigram, 2019).

Keywords Natural Language Processing · Bag of Words · TF-IDF · Transformers · BERT · GPT · Large Language Models · Keyword Extraction

NLP: Definition, Issues, etc.

Definition

Natural Language Processing (NLP) relies on machine learning models to process natural language. Before, we said Textmining.

S. Ndiaye (✉)
Département de Mathématiques et Informatique, Faculté des Sciences et Techniques, Université Cheikh Anta Diop, Dakar, Sénégal
e-mail: samba.ndiaye@ucad.edu.sn

C. T. Gueye et al. (eds.), *Mathematics of Computer Science, Cybersecurity and Artificial Intelligence*, Springer Proceedings in Mathematics & Statistics 468,
https://doi.org/10.1007/978-3-031-66222-5_2

Challenges

Most of the information in organizations is not structured, as for the relational databases that we can deal with SQL queries. This information is the multiple documents received or sent, messages, emails, reports, etc. In short, textual documents. To provide solutions to these challenges, text mining as opposed to data mining has been proposed for the extraction of knowledge from texts with machine learning tools.

The Tasks of NLP

We can note several tasks related to NLP such as:

- Classification of whole sentences: Analyze the sentiment of a review, and detect if an email is spam.
- Classification of each word in a sentence: Identify the entities named (person, place, organization).
- Text generation: Complete the start of a text with automatically generated text, and replace missing or hidden words in a text.
- Extraction of an answer from a text: Given a question and context, extract the answer to the question based on the information provided by the context.
- Generation of new sentences from a text: Translate a text in another language, and summarize a text.
- Speech recognition and computer vision such as generating a transcription from an audio sample or description of an image.

Representation Models/Word Embedding/Text Vectorization

Word embeddings are a type of word representation that allows words to be represented as vectors in a continuous vector space $(x_1, x_2, ..., x_n)$. These vectors capture the semantic meanings of the words, such that words with similar meanings are located in close proximity to one another in the vector space. There are several types of word embedding techniques, each with its own methodology and applications. Here are some of the most widely recognized types:

- Bag-of-Words: The Bag-of-Words (BoW) model is a simple yet powerful approach used in natural language processing (NLP) and information retrieval (IR). It represents text data, such as sentences or documents, in a way that the structure or the order of words is disregarded, focusing instead on the occurrence of words within the document.

- TF-IDF: A popular variant of Bag-of-Words that not only counts occurrences but also weights each word's frequency by its inverse frequency across all documents. This helps in emphasizing words that are unique to a particular document.
- Word2Vec [6]: Developed by a team at Google led by Tomas Mikolov, Word2Vec is one of the most popular techniques. It comes in two flavors: continuous bag-of-words (CBOW) and Skip-Gram. CBOW predicts a word based on its context, whereas Skip-Gram does the opposite, predicting the context given a word.
- GloVe (Global Vectors for Word Representation) [7]: Developed by Stanford University researchers, GloVe is an unsupervised learning algorithm for obtaining vector representations for words by aggregating global word–word co-occurrence statistics from a corpus. The resulting embeddings showcase interesting linear substructures of the word vector space.
- FastText [1]: Created by Facebook's AI Research (FAIR) lab, FastText extends Word2Vec by not only considering whole words but also taking into account subword units (n-grams). This allows the model to capture the meaning of shorter words and understand suffixes and prefixes, making it better at handling rare words, misspellings, and morphologically rich languages.

One can download pretrained models and use the word vectors directly. There are several types of pretrained integration models available online. For example, the model:

- Glove-twitter-50 is a GloVe model of size 50 trained on a Twitter-previous dataset.
- Glove-wiki-gigaword-100 is a size 100 GloVe model trained on Wikipedia 2014 + Gigaword.
- word2vec-google-news-300 is a word2vec model of size 300 trained on Google News (100 billion words).

Second Revolution in NLP: Transformer (Google 2017)

New neural network architecture (DEEP LEARNING) allows calculations to be parallelized during Language Model training. The Transformer architecture was presented in June 2017 [13]. Initially, the research focused on the translation task. This was followed by the introduction of several influential models, including:

- June 2018: GPT [10], the first pretrained and fine-tuned transformer on different NLP tasks and having obtained state-of-the-art results.
- October 2018: BERT [3], another large pretrained model having been built to produce better text summaries.
- February 2019: GPT-2 [11], an improved (and larger) version of GPT which was not directly made public due to ethical reasons.

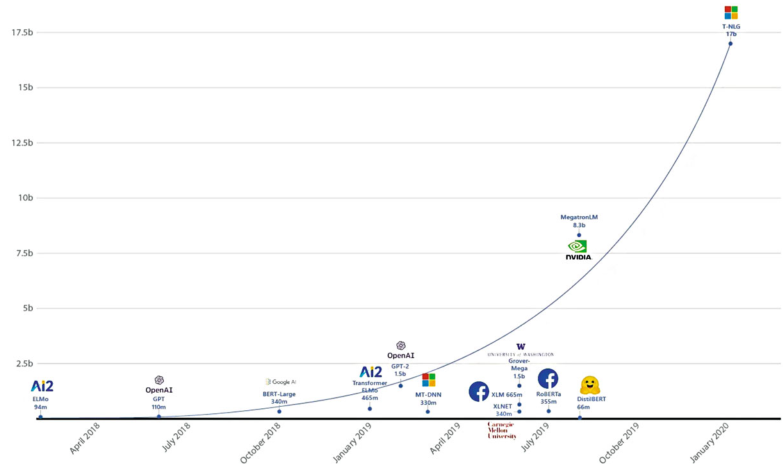


Fig. 2.1 Evolution of large language models

- May 2020: GPT-3 [2], an even larger version than GPT-2 with very good performance on a variety of tasks not requiring fine-tuning (called zero-shot learning).
- All mentioned transformers (GPT, BERT, etc.) were trained as language models. They were trained on a large quantity of raw texts in a self-supervised manner.
- Self-supervised learning is a type of training in which the goal is automatically calculated from the model inputs. This means humans are not needed to label the data!
- This type of model develops a statistical understanding of the language it was trained on, but it is not very useful for specific practical tasks.
- For this reason, the pretrained model then goes through a process called transfer learning. During this process, the model is fine-tuned in a supervised manner (i.e., using human-annotated labels) for a given task.
- An example task is to predict the next word in a sentence after reading the previous n words.
- The general strategy to obtain better performance consists of increasing the size of the models as well as the quantity of data used for training them as shown in this Fig. 2.1.

Unfortunately, training a model, and particularly a very large model, requires a significant amount of data. This becomes very expensive in terms of time and computational resources. In this case we speak of **large language model** or **llm**.

LLMs are used for a variety of tasks, such as text generation, machine translation, classification text, answering questions, and code autocompletion.

Generative AI

Stable Diffusion XL

Stable Diffusion XL (SDXL) [8] enhances the text-to-image generation capabilities of its predecessors through three significant advancements:

- The model features a UNet architecture that is three times larger (see Fig. 2.2), and it integrates an additional text encoder (OpenCLIP ViT-bigG/14) alongside the original encoder, substantially expanding its parameter count.
- It incorporates techniques for size and crop conditioning, allowing for the preservation of training data that would otherwise be discarded, and providing enhanced control over the cropping of generated images.
- SDXL introduces a dual-stage generation process. Initially, the base model produces an image, which can also function independently. This image then serves as the input for the refinement model, which enhances the image with additional, high-quality details.

Emu Video

Emu Video [5] introduces a streamlined text-to-video generation technique utilizing our Emu model, a diffusion-based framework. This innovative architecture facilitates video generation from diverse inputs, including text, images, or a combination of both. The process is divided into two distinct phases: initially creating images from text prompts, followed by crafting videos that incorporate both the initial text and the generated images. This segmented approach enhances the efficiency of training video generation models. Our method demonstrates that a singular diffusion model can accomplish this segmented, or "factorized," video generation process (see Fig. 2.3). We outline essential design strategies, such as optimizing noise schedules tailored for video diffusion and implementing a multi-phase training regimen. These strategies enable the direct creation of videos in higher resolutions.

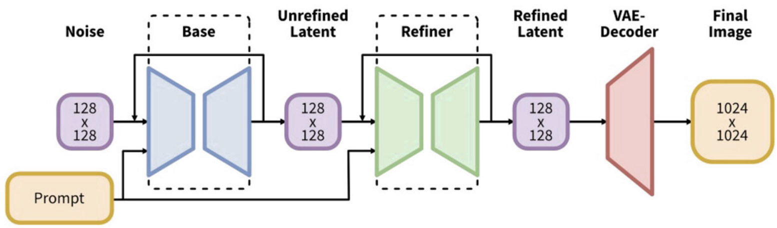


Fig. 2.2 SDXL architecture

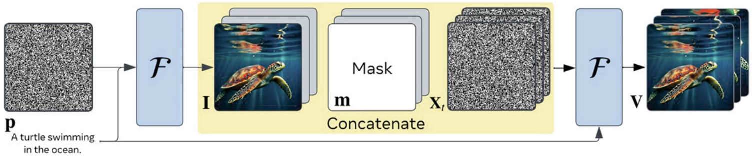


Fig. 2.3 Factorized text-to-video generation

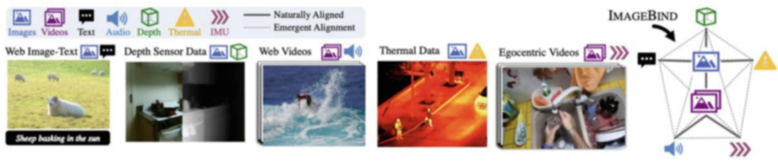


Fig. 2.4 ImageBind architecture

ImageBind

ImageBind [4] emerges as a part of Meta's expanding suite of open-source AI tools, aligning with cutting-edge computer vision technologies such as DINOv2, which introduces a method for training high-performance vision models without the need for fine-tuning, and Segment Anything (SAM), a versatile segmentation model capable of identifying any object in any image in response to any user prompt. ImageBind enhances this collection by focusing on multimodal representation learning, aiming to unify various modalities, including images and videos, within a single coherent feature space. Looking ahead, ImageBind is positioned to harness the advanced visual features of DINOv2, potentially boosting its multimodal learning capabilities (see Fig. 2.4).

Review of Open-Source Models

LLAMA 2

Llama 2 [12] is a set of transformer-based autoregressive causal language models designed to predict the next word(s) in a sequence given previous words. These models are trained through a self-supervised process on a large corpus of unlabeled data, totaling 2 trillion tokens from public sources, to minimize the difference between their predictions and the actual subsequent words. This training enables them to mimic linguistic and logical patterns found in the data, even though the specific data sources used are not disclosed in the research. Initially, these models do not directly answer prompts but generate text that is grammatically coherent with the given prompt. Specialized training methods like supervised learning and reinforcement learning are necessary to fine-tune these foundational models for specific tasks such as dialogue generation or creative writing.

Llama 2 models, building on the legacy of the original LLaMa, have become the basis for several significant open-source language models, offering a versatile foundation for developing models tailored to specific purposes.

BLOOM

BLOOM [14] stands out as a groundbreaking open-access multilingual language model boasting 176 billion parameters, developed over 3.5 months using 384 A100–80GB GPUs. It requires 330 GB of storage for a checkpoint. This model is the result of a monumental collaboration involving over 1000 scientists alongside the formidable Hugging Face team, making it a significant achievement that such an expansive multilingual model is freely accessible to all.

Operating as a causal language model, BLOOM is engineered to predict the subsequent token in a sequence based on the preceding ones. This approach, seemingly straightforward, enables the model to exhibit a form of reasoning capability. By predicting next tokens, BLOOM can weave together various concepts within sentences, tackling complex tasks like arithmetic, translation, and coding with a notable level of precision.

Architecturally, BLOOM is built on the Transformer framework, featuring an input embeddings layer, 70 Transformer blocks, and an output layer designed for language modeling. Each Transformer block is equipped with a self-attention mechanism and a multilayer perceptron, refined with norms both before and after the attention process, facilitating its advanced language processing capabilities.

Massively Multilingual Speech

The Massively Multilingual Speech (MMS) [9] initiative has successfully tackled numerous obstacles by integrating wav2vec 2.0, a groundbreaking self-supervised learning model, with an innovative dataset. This dataset includes labeled data for over 1100 languages and unlabeled data spanning nearly 4000 languages. Remarkably, some languages covered, like Tatuyo, have merely a few hundred speakers and lack any prior speech technology.

The findings from this project indicate that the MMS models surpass the performance of existing models, extending support to a scope of languages tenfold larger than before. Meta's commitment to multilingualism is evident not only in speech but also in text: The NLLB project has expanded multilingual translation to 200 languages, while the MMS project broadens the reach of speech technology to an even larger array of languages.

Universal Speech Model

The Universal Speech Model (USM) [15] represents a cutting-edge collection of speech models, boasting 2 billion parameters and trained across an extensive dataset comprising 12 million hours of speech and 28 billion sentences, covering more than 300 languages. Designed for applications such as YouTube's closed captioning, USM is capable of performing automatic speech recognition (ASR) not only for major languages like English and Mandarin but also for a diverse array of languages including Punjabi, Assamese, Santhali, Balinese, Shona, Malagasy, Luganda, Luo, Bambara, Soga, Maninka, Xhosa, Akan, Lingala, Chichewa, Nkore, Nzema, among others. Many of these languages have less than twenty million speakers, posing significant challenges in gathering sufficient training data.

The approach demonstrates the effectiveness of leveraging a vast, unlabeled multilingual dataset for the initial pretraining of the model's encoder, followed by fine-tuning with a smaller, labeled dataset. This strategy not only enables the recognition of these less commonly represented languages but also ensures the model's adaptability to new languages and datasets.

References

1. Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. Enriching word vectors with subword information. https://arxiv.org/abs/1607.04606, 2017.
2. Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. Language models are few-shot learners. https://arxiv.org/abs/2005.14165, 2020.
3. Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. https://arxiv.org/abs/1810.04805, 2019.
4. Rohit Girdhar, Alaaeldin El-Nouby, Zhuang Liu, Mannat Singh, Kalyan Vasudev Alwala, Armand Joulin, and Ishan Misra. ImageBind: One embedding space to bind them all. https://arxiv.org/abs/2305.05665, 2023.
5. Rohit Girdhar, Mannat Singh, Andrew Brown, Quentin Duval, Samaneh Azadi, Sai Saketh Rambhatla, Akbar Shah, Xi Yin, Devi Parikh, and Ishan Misra. Emu Video: Factorizing text-to-video generation by explicit image conditioning. https://arxiv.org/abs/2311.10709, 2023.
6. Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. https://arxiv.org/abs/1301.3781, 2013.
7. Jeffrey Pennington, Richard Socher, and Christopher Manning. GloVe: Global vectors for word representation. In Alessandro Moschitti, Bo Pang, and Walter Daelemans, editors, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543, Doha, Qatar, October 2014. Association for Computational Linguistics.
8. Dustin Podell, Zion English, Kyle Lacey, Andreas Blattmann, Tim Dockhorn, Jonas Müller, Joe Penna, and Robin Rombach. SDXL: Improving latent diffusion models for high-resolution image synthesis. https://arxiv.org/abs/2307.01952, 2023.

9. Vineel Pratap, Andros Tjandra, Bowen Shi, Paden Tomasello, Arun Babu, Sayani Kundu, Ali Elkahky, Zhaoheng Ni, Apoorv Vyas, Maryam Fazel-Zarandi, Alexei Baevski, Yossi Adi, Xiaohui Zhang, Wei-Ning Hsu, Alexis Conneau, and Michael Auli. Scaling speech technology to 1,000+ languages. https://arxiv.org/abs/2305.13516, 2023.
10. Alec Radford and Karthik Narasimhan. Improving language understanding by generative pre-training. https://s3-us-west-2.amazonaws.com/openai-assets/research-covers/language-unsupervised/language_understanding_paper.pdf, 2018.
11. Alec Radford, Jeff Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. https://d4mucfpksywv.cloudfront.net/better-language-models/language-models.pdf, 2019.
12. Hugo Touvron, Louis Martin, and al. Llama 2: Open foundation and fine-tuned chat models. https://arxiv.org/abs/2307.09288, 2023.
13. Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. https://arxiv.org/abs/1706.03762, 2023.
14. BigScience Workshop, :, Teven Le Scao, Angela Fan, Christopher Akiki, and al. BLOOM: A 176b-parameter open-access multilingual language model. https://arxiv.org/abs/2211.05100, 2023.
15. Yu Zhang, Wei Han, James Qin, Yongqiang Wang, Ankur Bapna, Zhehuai Chen, Nanxin Chen, Bo Li, Vera Axelrod, Gary Wang, Zhong Meng, Ke Hu, Andrew Rosenberg, Rohit Prabhavalkar, Daniel S. Park, Parisa Haghani, Jason Riesa, Ginger Perng, Hagen Soltau, Trevor Strohman, Bhuvana Ramabhadran, Tara Sainath, Pedro Moreno, Chung-Cheng Chiu, Johan Schalkwyk, Françoise Beaufays, and Yonghui Wu. Google USM: Scaling automatic speech recognition beyond 100 languages. https://arxiv.org/abs/2303.01037, 2023.

Part II
Contributed Talks: Mathematics and Modeling

Chapter 3
On Absolute-Valued Algebras with Nonzero Central Element

Alassane Diouf, Mbayang Amar, and Oumar Diankha

Abstract Let $\mathcal{A}$ be an absolute-valued algebra with nonzero element a such that a and a^2 are central, then $\mathcal{A}$ is pre-Hilbert space and admits an involution. We also show that if $\mathcal{A}$ is an absolute-valued algebra with nonzero central element satisfying $(x, x^2, x) = (x^2, x^2, x^2) = 0$, then $\mathcal{A}$ is finite-dimensional, flexible, and isomorphic to either $\mathbb{R}$, $\mathbb{C}$, $\overset{\star}{\mathbb{C}}$, $\mathbb{H}$, $\overset{\star}{\mathbb{H}}$, $\mathbb{O}$, or $\overset{\star}{\mathbb{O}}$.

Keywords Absolute-valued algebra · Division algebra · Central element · Involution

Mathematics Subject Classification 17A20, 17A30, 17A35, 17A60, 17A80

Introduction

The absolute-valued algebras are introduced by Ostrowski 1918 [16]. We recall that a (not necessarily associative) real algebra $\mathcal{A} \neq 0$ is said to be an absolute-valued algebra if its vector space is a normed space whose norm $\|.\|$ satisfies $\|xy\| = \|x\|\|y\|$ for all $x, y \in \mathcal{A}$. If, moreover, the norm $\|.\|$ comes from an inner product, then the algebra $\mathcal{A}$ is said to be a pre-Hilbert absolute-valued algebra.

Let $\mathbb{A}$ be one of the principal absolute-valued algebras $\mathbb{C}$, $\mathbb{H}$, $\mathbb{O}$. The mapping

$$\langle .,. \rangle : \mathbb{A} \times \mathbb{A} \to \mathbb{R} \quad (x, y) \mapsto \langle x, y \rangle = \frac{1}{2}(x\overline{y} + y\overline{x})$$

A. Diouf · M. Amar (✉) · O. Diankha
Département de Mathématiques et Informatique, Faculté des Sciences et Techniques, Université Cheikh Anta Diop, Dakar, Sénégal
e-mail: alassane1.diouf@ucad.edu.sn; mbayang1.amar@ucad.edu.sn; oumar.diankha@ucad.edu.sn

C. T. Gueye et al. (eds.), *Mathematics of Computer Science, Cybersecurity and Artificial Intelligence*, Springer Proceedings in Mathematics & Statistics 468,
https://doi.org/10.1007/978-3-031-66222-5_3

is an inner product [14, p. 208] which converts $\mathbb{A}$ into an Euclidian space $(\mathbb{A}, \|.\|)$: $\|x\| = \sqrt{x\overline{x}} = \sqrt{\overline{x}x}$ for all $x \in \mathbb{A}$, and we have: $\|xy\| = \|x\|\|y\|$. In addition, the equalities $\langle x(y+z), x(y+z)\rangle = \|x\|^2\langle y+z, y+z\rangle = \langle (y+z)x, (y+z)x\rangle$, valid for all $x, y, z \in \mathbb{A}$, give

$$\langle xy, xz\rangle = \langle yx, zx\rangle = \|x\|^2\langle y, z\rangle.$$

Let $\mathcal{A}$ be an absolute-valued algebra with unit, and then $\mathcal{A}$ is isomorphic to $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$ (Hamilton's quaternions), or $\mathbb{O}$ (Cayley's octonions) [22, Theoem 1]. Albert shows that every finite-dimensional absolute-valued algebra has dimension $n = 1, 2, 4$ or 8 and is isotopic to one of the classical absolute-valued algebras $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, or $\mathbb{O}$ [1].

Let $\mathbb{A} \in \{\mathbb{C}, \mathbb{H}, \mathbb{O}\}$. We recall that $^*\mathbb{A}$, $\mathbb{A}^*$, and $\overset{\star}{\mathbb{A}}$ are obtained by endowing the normed space $\mathbb{A}$ with the product $x \cdot y = \overline{x}y$, $x \cdot y = x\overline{y}$, and $x \cdot y = \overline{x}\,\overline{y}$, respectively, where $x \mapsto \overline{x}$ means the standard involution. We observe that $\overset{\star}{\mathbb{A}}$ contains a nonzero central idempotent.

An algebra $\mathcal{A}$ is called algebraic if $\mathcal{A}(x)$ is finite-dimensional for all $x \in \mathcal{A}$ and the bigger m such that $m = dim(\mathcal{A}(x))$ is called the degree of $\mathcal{A}$. Every absolute-valued algebraic algebra is finite-dimensional [15] and of degree 1, 2, 4, or 8. The algebra $\mathbb{R}$ is the unique absolute-valued algebra of degree 1. The absolute-valued algebras of degree 2 are $\mathbb{C}$, $^*\mathbb{C}$, $\overset{\star}{\mathbb{C}}$, $\mathbb{C}^*$, $\mathbb{H}$, $^*\mathbb{H}$, $\overset{\star}{\mathbb{H}}$, $\mathbb{H}^*$, $\mathbb{O}$, $^*\mathbb{O}$, $\overset{\star}{\mathbb{O}}$, $\mathbb{O}^*$, and $\mathbb{P}$ [18]. We precise that the algebras $\mathbb{R}$, $\mathbb{C}$, $\overset{\star}{\mathbb{C}}$, $\mathbb{H}$, $\overset{\star}{\mathbb{H}}$, $\mathbb{O}$, $\overset{\star}{\mathbb{O}}$, and $\mathbb{P}$ satisfy the identity $(x, x^2, x) = 0$.

In 1990, El-Mallah proves if $\mathcal{A}$ is an absolute-valued algebra with nonzero central idempotent, then $\mathcal{A}$ admits an involution [10].

In [17, Theorem 4.4], Rochdi and Rodriguez give a classification of finite-dimensional absolute-valued algebras containing a nonzero central idempotent.

In [5], the authors also gave a classification of finite-dimensional absolute-valued algebras containing a nonzero central idempotent.

Recently, the absolute-valued algebras containing a nonzero central idempotent and satisfying an identity $(x^p, x^q, x^r) = 0$ are fully described [3, 4, 8, 9, 11, 12]. They prove that $\mathcal{A}$ is finite-dimensional. More precisely, if $(x^p, x^q, x^r) = 0$ is symmetric, then $\mathcal{A}$ is equal to $\mathbb{R}$, $\mathbb{C}$, $\overset{\star}{\mathbb{C}}$, $\mathbb{H}$, $\overset{\star}{\mathbb{H}}$, $\mathbb{O}$, or $\overset{\star}{\mathbb{O}}$, and if $(x^p, x^q, x^r) = 0$ is asymmetric, then $\mathcal{A}$ is equal to $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, or $\mathbb{O}$.

In [6], the authors prove that if $\mathcal{A}$ is a pre-Hilbert absolute-valued algebra satisfying $(x, x^2, x) = (x^2, y, x^2) = 0$, then $\mathcal{A}$ is finite-dimensional, flexible, and isomorphic to either $\mathbb{R}$, $\mathbb{C}$, $\overset{\star}{\mathbb{C}}$, $\mathbb{H}$, $\overset{\star}{\mathbb{H}}$, $\mathbb{O}$, $\overset{\star}{\mathbb{O}}$, or $\mathbb{P}$.

In [7], the authors show that if $\mathcal{A}$ is an absolute-valued algebra satisfying $(x, x, x) = 0$ and containing a nonzero central element, then $\mathcal{A}$ is isomorphic to $\mathbb{R}$, $\mathbb{C}$, $\overset{\star}{\mathbb{C}}$, $\mathbb{H}$, $\overset{\star}{\mathbb{H}}$, $\mathbb{O}$, or $\overset{\star}{\mathbb{O}}$. They also prove that if $\mathcal{A}$ is an absolute-valued algebra satisfying $(x, x^2, x) = 0$ and containing an algebraic nonzero central element, then $\mathcal{A}$ is isomorphic to $\mathbb{R}$, $\mathbb{C}$, $\overset{\star}{\mathbb{C}}$, $\mathbb{H}$, $\overset{\star}{\mathbb{H}}$, $\mathbb{O}$, or $\overset{\star}{\mathbb{O}}$.

Motivated by the following results, we were interested in the study of the absolute-valued algebras having a nonzero central element a. We have made one contribution in the case where a and a^2 are central (Theorem 3.1) and another contribution in the case where $\mathcal{A}$ satisfies the identities $(x, x^2, x) = (x^2, x^2, x^2) = 0$ (Theorem 3.2). In the first case we have shown that $\mathcal{A}$ is pre-Hilbert space and admits an involution, and in the second case we have proved that $\mathcal{A}$ is finite-dimensional, flexible, and isomorphic to either $\mathbb{R}$, $\mathbb{C}$, $\overset{\star}{\mathbb{C}}$, $\mathbb{H}$, $\overset{\star}{\mathbb{H}}$, $\mathbb{O}$, or $\overset{\star}{\mathbb{O}}$. The second case directly implies that $\mathbb{R}$, $\mathbb{C}$, $\overset{\star}{\mathbb{C}}$, $\mathbb{H}$, $\overset{\star}{\mathbb{H}}$, $\mathbb{O}$, and $\overset{\star}{\mathbb{O}}$ are the only third power-associative absolute-valued algebras with nonzero central element.

Preliminary Notes

By an algebra we mean a vector space $\mathcal{A}$ over $\mathbb{R}$ endowed with a bilinear mapping $(x, y) \mapsto xy$ from $\mathcal{A} \times \mathcal{A}$ to $\mathcal{A}$ called the product of the algebra. Given elements a, b, c in any algebra, we set $(a, b, c) := (ab)c - a(bc)$ for the associator of a, b, and c; $[a, b] := ab - ba$ for the commutator of a and b. We recall that an element a is central if $[a, x] = 0$ for all x in $\mathcal{A}$. An algebra $\mathcal{A}$ is called flexible if $(x, y, x) = 0$ for all $x, y \in \mathcal{A}$. We denote by $\mathcal{A}(x)$, the subalgebra generated by every element $x \in \mathcal{A}$. An element $x \in \mathcal{A}$ is called algebraic if $\mathcal{A}(x)$ is finite-dimensional.

Let $x \in \mathcal{A}$. If $[x^2, x] = 0$, then we can set $x^3 := x^2x = xx^2$.

An involution on an absolute-valued algebra $\mathcal{A}$ is a mapping $x \mapsto x^\star$ from $\mathcal{A}$ to $\mathcal{A}$ satisfying:

(1) $(\alpha x + \beta y)^\star = \alpha x^\star + \beta y^\star$
(2) $x^{\star\star} = x$
(3) $xx^\star = x^\star x$
(4) $(xy)^\star = y^\star x^\star$
(5) $\|x^\star\| = \|x\|$

for all $x, y \in \mathcal{A}$ and $\alpha, \beta \in \mathbb{R}$. Axiom (5) follows immediately from the others [19]. Given an absolute-valued algebra $\mathcal{A}$ with involution, we denote by $\mathcal{A}_a$ the set of all self-adjoint elements of $\mathcal{A}$ and by $\mathcal{A}_s$ the set of all skew elements of $\mathcal{A}$. We have $\mathcal{A}_a := \{x \in \mathcal{A} : x^* = x\}$ and $\mathcal{A}_s := \{x \in \mathcal{A} : x^* = -x\}$. Obviously, we have $\mathcal{A} = \mathcal{A}_a \oplus \mathcal{A}_s$, as a direct sum of subspaces. We will assume that the involution of $\mathcal{A}$ is nontrivial if $\mathcal{A}_s \neq 0$. Clearly, $\mathcal{A}$ contains a unique nonzero self-adjoint idempotent [21]. If $\mathcal{B} := \mathbb{R}e \oplus \mathcal{A}_s$ is finite-dimensional, then $\mathcal{A} = \mathcal{B}$ and the idempotent e is central [9, Lemma 3.2].

Recall that an algebra $\mathcal{A}$ is said to be third power associative if it satisfies the identity $(x, x, x) = 0$, which can also be rewritten as $[x^2, x] = 0$. As an immediate consequence, a third power-associative algebra satisfies the identity $(x^2, x^2, x^2) = 0$. Every algebra satisfying the identity $(x, x, x) = 0$ also satisfies the identity $(x, x^2, x) = 0$ [2, Lemma 2.1].

Let us consider the identities of the form $(x^p, x^q, x^r) = 0$ with $p, q, r \in \{1, 2\}$. For our convenience, we will say that such an identity is symmetric when $p = r$ and asymmetric otherwise. Thus the symmetric identities are

$$(x, x, x) = 0, \quad (x, x^2, x) = 0, \quad (x^2, x, x^2) = 0 \quad \text{and} \quad (x^2, x^2, x^2) = 0,$$

whereas the asymmetric identities are

$$(x^2, x, x) = 0, \quad (x, x, x^2) = 0, \quad (x^2, x^2, x) = 0, \quad \text{and} \quad (x, x^2, x^2) = 0.$$

Main Results

Lemma 3.1 *Let $(\mathcal{A}, \|.\|)$ be an absolute-valued algebra containing a norm-one element a such that $[a^2, a] = [a^2, a^3] = [a^2, (a^2)^2] = 0$ and $a^3 \in \mathcal{L}_{in}\{a, a^2\}$. Then, we have $\mathcal{A}(a) = \mathcal{L}_{in}\{a, a^2\}$.*

Proof If a and a^2 are collinear, then the result is obvious. Suppose that a and a^2 are not collinear. As $\{a, a^2\}$ is commutative, so $\mathcal{L}_{in}\{a, a^2\}$ is an inner-product space [22, Lemma 1].

Now, let $a_0 \in \mathcal{L}_{in}\{a, a^2\}$ such that $\|a_0\| = 1$ and $a_0 \in a^{\perp}$. Then $a_0 = \alpha a + \beta a^2$ with $\alpha, \beta \in \mathbb{R}$ and $\beta \neq 0$. As a_0^2 commutes with a^2 and $\|a^2 - a_0^2\| = \|(a - a_0)(a + a_0)\| = \|a - a_0\|\|a + a_0\| = 2$, we have $a^2 + a_0^2 = 0$ [22, Lemma 3]. We get

$$\begin{aligned} 0 &= a^2 + a_0^2 \\ &= a^2 + (\alpha a + \beta a^2)^2 \\ &= a^2 + \alpha^2 a^2 + 2\alpha\beta a^2 a + \beta^2 (a^2)^2 \\ &= (1 + \alpha^2) a^2 + 2\alpha\beta a^3 + \beta^2 (a^2)^2, \end{aligned}$$

so $(a^2)^2 = -\frac{1}{\beta^2}[\,(1 + \alpha^2) a^2 + 2\alpha\beta a^3].$

As $a^3 \in \mathcal{L}_{in}\{a, a^2\}$, we have $a^3 = \gamma a + \delta a^2$ with $\gamma, \delta \in \mathbb{R}$, which implies that

$$\begin{aligned} (a^2)^2 &= -\frac{1}{\beta^2}[\,(1 + \alpha^2) a^2 + 2\alpha\beta(\gamma a + \delta a^2)] \\ &= -\frac{1}{\beta^2}[\,(1 + \alpha^2) a^2 + 2\alpha\beta\gamma a + 2\alpha\beta\delta a^2] \\ &= -\frac{1}{\beta^2}[\,2\alpha\beta\gamma a + (1 + \alpha^2 + 2\alpha\beta\delta) a^2], \end{aligned}$$

so $(a^2)^2 \in \mathcal{L}_{in}\{a, a^2\}$. We realize that $\mathcal{A}(a) = \mathcal{L}_{in}\{a, a^2\}$. □

Theorem 3.1 *Let* $(\mathcal{A}, \|.\|)$ *be an absolute-valued algebra containing a nonzero element* a *such that* a *and* a^2 *are central. Then* $\mathcal{A}$ *is pre-Hilbert space and admits an involution.*

Proof We can assume without loss of generality that $\|a\| = 1$. We will distinguish the following cases:

First case If a is collinear to a^2, then $a^2 = \gamma a$, where $\gamma \in \mathbb{R}\backslash\{0\}$. By putting, $a_0 = \gamma^{-1}a$, we have $a_0^2 = \gamma^{-1}a = a_0$. This implies that a_0 is a nonzero central idempotent of $\mathcal{A}$.

Second case If a is not collinear to a^2, as the set $\{a, a^2, a^3\}$ is commutative, so $\mathcal{L}_{in}\{a, a^2, a^3\}$ is an inner-product space [22, Lemma 1].

We assume that $dim(\mathcal{L}_{in}\{a, a^2, a^3\}) = 3$. There exist a norm one $a_0 \in \mathcal{L}_{in}\{a, a^2, a^3\}$ orthogonal to a and a^2. We get

$$\|a_0^2 - a^2\| = \|(a_0 - a)(a_0 + a)\| = \|a_0 - a\|\|a_0 + a\| = 2. \tag{3.1}$$

Since $a^2a_0^2 = a_0^2a^2$, then we have $a_0^2 + a^2 = 0$ [22, Lemma 3].

We have also that

$$\|a_0^2 - (a^2)^2\| = \|(a_0 - a^2)(a_0 + a^2)\| = \|a_0 - a^2\|\|a_0 + a^2\| = 2. \tag{3.2}$$

As $a_0^2 = -a^2$, so we have $(a^2)^2a_0^2 = a_0^2(a^2)^2$. We deduce that $a_0^2 + (a^2)^2 = 0$ [22, Lemma 3]. We realize that $(a^2)^2 - a^2 = (a^2 - a)(a^2 + a) = 0$. As $\mathcal{A}$ has no divisors of zero, we have $a^2 = a$ or $a^2 = -a$. Absurd, because a is not collinear to a^2. This implies that $dim(\mathcal{L}_{in}\{a, a^2, a^3\}) = 2$, so $a^3 \in \mathcal{L}_{in}\{a, a^2\}$.

Lemma 3.1 gives that $\mathcal{A}(a) = \mathcal{L}_{in}\{a, a^2\}$. As $\mathcal{A}(a)$ is two-dimensional division algebra, then $\mathcal{A}(a) = \mathcal{L}_{in}\{a, a^2\}$ contains a nonzero idempotent e_0 [20]. Then $e_0 = \eta a + \zeta a^2$, where $(\eta, \zeta) \in \mathbb{R}^2\backslash\{(0, 0)\}$ is a nonzero central idempotent of $\mathcal{A}$.

In all cases $\mathcal{A}$ contains a nonzero central idempotent, and consequently $\mathcal{A}$ is pre-Hilbert space and $\mathcal{A}$ admits an involution [10, Theorem 3.6 and Theorem 3.7]. □

The following result follows from Theorem 3.1 and [3, 4, 8, 9, 11, 12].

Corollary 3.1 *Let* $(\mathcal{A}, \|.\|)$ *be an absolute-valued algebra satisfying* $(x^p, x^q, x^r) = 0$*, and containing a nonzero element* a *such that* a *and* a^2 *are central. Then* $\mathcal{A}$ *is finite-dimensional. More precisely, we have:*

(1) If $(x^p, x^q, x^r) = 0$ *is asymmetric, then* $\mathcal{A}$ *is power associative, that is,* $\mathcal{A}$ *is equal to* $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, *or* $\mathbb{O}$.
(2) If $(x^p, x^q, x^r) = 0$ *is symmetric, then* $\mathcal{A}$ *is flexible, that is,* $\mathcal{A}$ *is equal to* $\mathbb{R}$, $\mathbb{C}$, $\overset{\star}{\mathbb{C}}$, $\mathbb{H}$, $\overset{\star}{\mathbb{H}}$, $\mathbb{O}$, *or* $\overset{\star}{\mathbb{O}}$.

Theorem 3.2 *Let* $(\mathcal{A}, \|.\|)$ *be an absolute-valued algebra with nonzero central element satisfying* $(x, x^2, x) = (x^2, x^2, x^2) = 0$*. Then* $\mathcal{A}$ *is finite-dimensional, flexible, and isomorphic to either* $\mathbb{R}$, $\mathbb{C}$, $\overset{\star}{\mathbb{C}}$, $\mathbb{H}$, $\overset{\star}{\mathbb{H}}$, $\mathbb{O}$, *or* $\overset{\star}{\mathbb{O}}$.

Proof We can assume without loss of generality that $\|a\| = 1$. We will distinguish the following cases:

First case If a is collinear to a^2, then $a^2 = \gamma a$, where $\gamma \in \mathbb{R}\backslash\{0\}$. This implies that $a_0 = \gamma^{-1}a$ is a nonzero central idempotent of $\mathcal{A}$, so $\mathcal{A}$ admits an involution [10, Theorem 3.7]. We deduce that $\mathcal{A}$ is isomorphic to $\mathbb{R}$, $\mathbb{C}$, $\overset{\star}{\mathbb{C}}$, $\mathbb{H}$, $\overset{\star}{\mathbb{H}}$, $\mathbb{O}$, or $\overset{\star}{\mathbb{O}}$ [4].

Second case If a is not collinear to a^2, we note that a^3 is defined because a is central. By linearizing $(x, x^2, x) = 0$, we obtain

$$(x, x^2, y) + (x, xy + yx, x) + (y, x^2, x) = 0 \quad (\mathbf{3.1}).$$

Taking $x = a$ and $y = a^2$ in **(3.1)**, we obtain

$$\begin{aligned} 0 &= (a, a^2, a^2) + (a, a^3 + a^3, a) + (a^2, a^2, a) \\ &= (a, a^2, a^2) + (a^2, a^2, a) \\ &= a^3a^2 - a(a^2)^2 + (a^2)^2a - a^2a^3 \\ &= a^3a^2 - a^2a^3 \\ &= [a^3, a^2]. \end{aligned}$$

As the set $\{a, a^2, a^3\}$ is commutative, so $\mathcal{L}_{in}\{a, a^2, a^3\}$ is an inner-product space [22, Lemma 1]. Suppose that $dim(\mathcal{L}_{in}\{a, a^2, a^3\}) = 3$. There exist a norm one $a_0 \in \mathcal{L}_{in}\{a, a^2, a^3\}$ orthogonal to a and a^2. We have

$$\|a_0^2 - a^2\| = \|(a_0 - a)(a_0 + a)\| = \|a_0 - a\|\|a_0 + a\| = 2,$$

and

$$\|a_0^2 - (a^2)^2\| = \|(a_0 - a^2)(a_0 + a^2)\| = \|a_0 - a^2\|\|a_0 + a^2\| = 2.$$

There are $\alpha, \beta, \gamma \in \mathbb{R}$ such that $a_0 = \alpha a + \beta a^2 + \gamma a^3$, so $a_0^2 = \alpha^2 a^2 + \beta^2(a^2)^2 + \gamma^2(a^3)^2 + 2\alpha\beta a^3 + 2\alpha\gamma aa^3 + 2\beta\gamma a^2a^3$.

We have $[(a^2)^2, a^2] = (a^2)^2a^2 - a^2(a^2)^2 = (a^2, a^2, a^2) = 0$. Linearizing **(3.1)**, we obtain

$$(x, xy + yx, y) + (x, y^2, x) + (y, x^2, y) + (y, xy + yx, x) = 0 \quad (\mathbf{3.2}).$$

Taking $x = a$ and $y = a^2$ in **(3.2)**, we have

$$\begin{aligned} 0 &= (a, a^3 + a^3, a^2) + (a, (a^2)^2, a) + (a^2, a^2, a^2) + (a^2, a^3 + a^3, a) \\ &= 2[(a, a^3, a^2) + (a^2, a^3, a)] \\ &= 2[(aa^3)a^2 - a(a^3a^2) + (a^2a^3)a - a^2(a^3a)] \end{aligned}$$

$$= 2[(aa^3)a^2 - a^2(a^3a)]$$
$$= 2[aa^3, a^2].$$

Putting $x = a$ and $y = (a^2)^2$ in **(3.1)**, we get

$$\begin{aligned} 0 &= (a, a^2, (a^2)^2) + (a, a(a^2)^2 + (a^2)^2a, a) + ((a^2)^2, a^2, a) \\ &= (a, a^2, (a^2)^2) + ((a^2)^2, a^2, a) \\ &= a^3(a^2)^2 - a(a^2(a^2)^2) + ((a^2)^2a^2)a - (a^2)^2a^3 \\ &= a^3(a^2)^2 - (a^2)^2a^3 \\ &= [a^3, (a^2)^2]. \end{aligned}$$

By linearizing $(x^2, x^2, x^2) = 0$, we obtain the identity

$$(x^2, x^2, xy + yx) + (x^2, xy + yx, x^2) + (xy + yx, x^2, x^2) = 0 \quad (\mathbf{3.3}).$$

Putting $x = a$ and $y = a^2$ in **(3.3)**, we derive that

$$\begin{aligned} 0 &= (a^2, a^2, a^3 + a^3) + (a^2, a^3 + a^3, a^2) + (a^3 + a^3, a^2, a^2) \\ &= 2[(a^2, a^2, a^3) + (a^2, a^3, a^2) + (a^3, a^2, a^2)] \\ &= 2[(a^2)^2a^3 - a^2(a^2a^3) + (a^2a^3)a^2 - a^2(a^3a^2) + (a^3a^2)a^2 - a^3(a^2)^2] \\ &= 4[(a^2a^3)a^2 - a^2(a^2a^3)] \quad \text{because} \quad [a^3, (a^2)^2] = 0 \\ &= 4[a^2a^3, a^2]. \end{aligned}$$

Taking $x = a$ and $y = a^2a^3$ in **(3.1)**, we have

$$\begin{aligned} 0 &= (a, a^2, a^2a^3) + (a, a(a^2a^3) + (a^2a^3)a, a) + (a^2a^3, a^2, a) \\ &= (a, a^2, a^2a^3) + (a^2a^3, a^2, a) \\ &= a^3(a^2a^3) - a(a^2(a^2a^3)) + ((a^2a^3)a^2)a - (a^2a^3)a^3 \\ &= a^3(a^2a^3) - (a^2a^3)a^3 \\ &= [a^3, a^2a^3]. \end{aligned}$$

Linearizing **(3.3)**, we obtain

$$\begin{aligned} &(x^2, x^2, y^2) + (x^2, xy + yx, xy + yx) + (x^2, y^2, x^2) \\ &+(xy + yx, x^2, xy + yx) + (xy + yx, xy + yx, x^2) + (y^2, x^2, x^2) = 0 \quad (\mathbf{3.4}). \end{aligned}$$

Putting $x = a$ and $y = a^2$ in **(3.4)**, we derive that

$$\begin{aligned}
0 &= (a^2, a^2, (a^2)^2) + (a^2, a^3 + a^3, a^3 + a^3) + (a^2, (a^2)^2, a^2) \\
&\quad +(a^3 + a^3, a^2, a^3 + a^3) \\
&\quad +(a^3 + a^3, a^3 + a^3, a^2) + ((a^2)^2, a^2, a^2) \\
&= (a^2, a^2, (a^2)^{2)}) + (a^2, a^3 + a^3, a^3 + a^3) + (a^3 + a^3, a^2, a^3 + a^3) \\
&\quad +(a^3 + a^3, a^3 + a^3, a^2) + ((a^2)^2, a^2, a^2) \\
&= (a^2)^2(a^2)^2 - a^2(a^2(a^2)^2) + ((a^2)^2a^2)a^2 - (a^2)^2(a^2)^2 \\
&\quad +4[(a^2, a^3, a^3) + (a^3, a^2, a^3) + (a^3, a^3, a^2)] \\
&= (a^2, (a^2)^2, a^2) + 4[(a^2a^3)a^3 - a^2(a^3)^2 + (a^3a^2)a^3 - a^3(a^2a^3) \\
&\quad +(a^3)^2a^2 - a^3(a^3a^2)] \\
&= 4[(a^2a^3)a^3 - a^2(a^3)^2 + (a^3a^2)a^3 - a^3(a^2a^3) + (a^3)^2a^2 - a^3(a^3a^2)] \\
&= 4[(a^3)^2a^2 - a^2(a^3)^2 + 2((a^2a^3)a^3 - a^3(a^3a^2))] \\
&= 4[(a^3)^2a^2 - a^2(a^3)^2] + 8[(a^2a^3)a^3 - a^3(a^3a^2)] \\
&= 4[(a^3)^2, a^2].
\end{aligned}$$

We realize that $[a^2, a^2] = [a^2, a^3] = [a^2, (a^2)^2] = [a^2, aa^3] = [a^2, a^2a^3] = [a^2, (a^3)^2] = 0$, so

$$\begin{aligned}
[a^2, a_0^2] &= [a^2, \alpha^2a^2 + \beta^2(a^2)^2 + \gamma^2(a^3)^2 + 2\alpha\beta a^3 + 2\alpha\gamma aa^3 + 2\beta\gamma a^2a^3] \\
&= \alpha^2[a^2, a^2] + \beta^2[a^2, (a^2)^2] + \gamma^2[a^2, (a^3)^2] + 2\alpha\beta[a^2, a^3] \\
&\quad +2\alpha\gamma[a^2, aa^3] + 2\beta\gamma[a^2, a^2a^3] \\
&= 0.
\end{aligned}$$

As $[a^2, a_0^2] = 0$ and $\|a_0^2 - a^2\| = 2$, we have $a_0^2 + a^2 = 0$ [22, Lemma 3]. The equality $a_0^2 = -a^2$ implies $(a^2)^2a_0^2 = a_0^2(a^2)^2$. Moreover as $\|a_0^2 - (a^2)^2\| = 2$, we realize that $a_0^2 + (a^2)^2 = 0$ [22, Lemma 3]. We have $(a^2)^2 - a^2 = (a^2 - a)(a^2 + a) = 0$. As $\mathcal{A}$ has no divisors of zero, we get $a^2 = a$ or $a^2 = -a$. Absurd, because a is not collinear to a^2. This implies that $dim(\mathcal{L}_{in}\{a, a^2, a^3\}) = 2$, so $a^3 \in \mathcal{L}_{in}\{a, a^2\}$.

Lemma 3.1 implies that $\mathcal{A}(a) = \mathcal{L}_{in}\{a, a^2\}$, so a is algebraic. This implies that $\mathcal{A}$ is isomorphic to either $\mathbb{R}$, $\mathbb{C}$, $\overset{*}{\mathbb{C}}$, $\mathbb{H}$, $\overset{*}{\mathbb{H}}$, $\mathbb{O}$, or $\overset{*}{\mathbb{O}}$ [7, Theorem 3]. □

We have the following consequences of Theorem 3.2.

Corollary 3.2 ([13, Theorem 2.1]) *Let* $(\mathcal{A}, \|.\|)$ *be a third power-associative absolute-valued algebra with nonzero central element. Then* $\mathcal{A}$ *is finite-dimensional and isotopic to* $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, *or* $\mathbb{O}$.

Corollary 3.3 ([7, Theorem 1.]) *Let $(\mathcal{A}, \|.\|)$ be a third power-associative absolute-valued algebra with nonzero central element. Then $\mathcal{A}$ is finite-dimensional and is isomorphic to either $\mathbb{R}$, $\mathbb{C}$, $\overset{\star}{\mathbb{C}}$, $\mathbb{H}$, $\overset{\star}{\mathbb{H}}$, $\mathbb{O}$, or $\overset{\star}{\mathbb{O}}$.*

Corollary 3.4 *Let $(\mathcal{A}, \|.\|)$ be an absolute-valued algebra with nonzero central element. Then the following statements are equivalent:*

(1) *$\mathcal{A}$ is third power associative.*
(2) *$\mathcal{A}$ satisfies $(x, x^2, x) = (x^2, x^2, x^2) = 0$.*
(3) *$\mathcal{A}$ is finite-dimensional and is isomorphic to either $\mathbb{R}$, $\mathbb{C}$, $\overset{\star}{\mathbb{C}}$, $\mathbb{H}$, $\overset{\star}{\mathbb{H}}$, $\mathbb{O}$, or $\overset{\star}{\mathbb{O}}$.*

By combining Theorem 3.2 and [7, Theorem 3], we also have the following corollary.

Corollary 3.5 *Let $(\mathcal{A}, \|.\|)$ be an absolute-valued algebra having nonzero central element a and satisfying $(x, x^2, x) = 0$. Then the following statements are equivalent:*

(1) *a is an algebraic element.*
(2) *The norm $\|.\|$ comes from an inner product.*
(3) *$\mathcal{A}$ is third power associative.*
(4) *$\mathcal{A}$ satisfies $(x^2, x^2, x^2) = 0$.*
(5) *$\mathcal{A}$ is finite-dimensional and is isomorphic to $\mathbb{R}$, $\mathbb{C}$, $\overset{\star}{\mathbb{C}}$, $\mathbb{H}$, $\overset{\star}{\mathbb{H}}$, $\mathbb{O}$, or $\overset{\star}{\mathbb{O}}$.*

Acknowledgments The authors would like to thank the referees for their careful reading of this paper and their useful suggestions.

References

1. A. A. Albert, Absolute valued real algebras. *Ann. of Math*. **48** (1947), 495–501.
2. M. Cabrera and A. Diouf. On power-associativity of algebras with no nonzero joint divisor of zero and containing a nonzero central idempotent. *Commun. Algebra* **52** (2024), 295–304.
3. A. Chandid and A. Rochdi, A survey on absolute valued algebras satisfying $(x^p, x^q, x^r) = 0$. *Int. J. Algebra* **2**, (2008), 837–852.
4. J. A. Cuenca, On composition and absolute valued algebras. *Proc. Roy. Soc. Edinburgh*, **136A** (2006), 717–731.
5. J. A. Cuenca, E. Darpö, and E. Dieterich, Classification of the finite dimensional absolute valued algebras having a non-zero central idempotent or a one-sided unity. *Bull. Sci. Math.* **134** (2010), 247–277.
6. K. Diaby, O. Diankha, A. Fall, and A. Rochdi, *On absolute-valued algebras satisfying* $(x^2, y, x^2) = 0$. *J. Algebra* **585** (2021), 484–500.
7. K. Diaby and A. Rochdi, A note on absolute valued algebras satisfying $(xx^2)x = x(x^2x)$. *J. Algebra. Appl*, (2022), 2250172, 8 pp.
8. M. L. El-Mallah, On finite dimensional absolute valued algebras satisfying $(x, x, x) = 0$. *Arch. Math.* **49** (1987), 16–22.
9. M. L. El-Mallah, Absolute valued algebras with an involution. *Arch. Math. (Basel)* **51** (1988), 39–49.

10. M. L. El-Mallah, Absolute valued algebras containing a central idempotent. *J. Algebra* **128** (1990), 180–187.
11. M. L. El-Mallah, Absolute valued algebras satisfying $(x, x, x^2) = 0$. *Arch. Math. (Basel)* **77** (2001), 378–382.
12. M. L. El-Mallah and M. El-Agawany, Absolute valued algebras satisfying $(x^2, x^2, x^2) = 0$. *Commun. Algebra* **32** (2004), 3537–3541.
13. M. L. El. Mallah and A. Micali, Sur les dimensions des algebres absolument valuees. *J. Algebra* **68** (2), (1981) 237–246.
14. F. Hirzebruch, M. Koecher and R. Remmert, *Numbers. Springer-Verlag*, (1991).
15. A. Kaidi, M. I. Ramirez and A. Rodriguez, Absolute valued algebraic algebras are finite dimensional. *J. Algebra* **195** (1997) 295–307.
16. A. Ostrowski, Über einige Lösungen der Funktionalgleichung $\Psi(x)\Psi(y) = \Psi(xy)$. *Acta Math*. **41**, (1918), 271–284.
17. A. Rochdi and Á. Rodríguez, Absolute valued algebras with involution. *Commun. Algebra* **37** (2009), 1151–1159.
18. A. Rodríguez, Absolute valued algebras of degree two. In: González, S., ed. Non-associative algebra and its applications. Dordrecht-Boston-london: *Kluwer Academic Publishers*, (1994), 350–356.
19. A. Rodríguez, On Urbanik's axioms for absolute valued algebras with involution. *Commun. Algebra* **36** (2008), 2588–2592.
20. B. Segre, La teoria delle algebre ed alcune questione di realta. *Univ. Roma, Ist. Naz. Alta. Mat., Rend. Mat. E App. Serie 5* **13** (1954) 157–188.
21. K. Urbanik, Absolute valued algebras with an involution. *Fundamenta Math* **49** (1961), 247–258.
22. K. Urbanik and F. B. Wright, Absolute valued algebras. *Proc. Amer. Math. Soc*. **11** (1960), 861–866.

Chapter 4
On Algebraic Algebras Without Divisors of Zero Satisfying $(x^p, x^q, x^r) = 0$

Mohamed Traoré and Alassane Diouf

Abstract Let $\mathcal{A}$ be an algebraic algebra without divisors of zero of degree $\neq 8$ with a nonzero idempotent e such that $[e, I(\mathcal{A})] = 0$ (resp., e is omnipresent). Then the following assertions are equivalent:

(1) $\mathcal{A}$ is quadratic with unit e.
(2) $\mathcal{A}$ is power associative.
(3) $\mathcal{A}$ satisfies $(x, x^q, x^r) = 0$ and e is a generalized left unit.
(4) $\mathcal{A}$ satisfies $(x^p, x^q, x) = 0$ and e is a generalized right unit.
(5) $\mathcal{A}$ satisfies $(x^p, x^q, x^r) = 0$ and e is a generalized unit.

Keywords Division algebra · Algebraic · Quadratic · Power commutative · Power associative · Third power associative · Generalized left unit

Mathematics Subject Classification 17A05, 17A30, 17A35

Introduction

In this chapter, the algebras are considered over $\mathbb{R}$ and $p, q, r \in \{1, 2\}$. The study of real division algebras was born since the construction of quaternions by Hamilton and octonions by Cayley in the middle of the 19 century. Despite its long history, the problem of classifying finite-dimensional real division algebras (FDRDA) remains open. Theorems (1, 2, 4, 8) state that the dimension of an FDRDA may be only 1, 2, 4, and 8 [5, 20, 21, 23]. The classification problem of FDRDA is solved completely in dimension one [26, 27] and two [2–4, 6, 17, 19, 22] and partially in four and eight dimensions. In dimension 4, the classification is effective for

M. Traoré · A. Diouf (✉)
Département de Mathématiques et Informatique, Faculté des Sciences et Techniques, Université Cheikh Anta Diop, Dakar, Sénégal
e-mail: mohamed5.traore@ucad.edu.sn; alassane1.diouf@ucad.edu.sn

C. T. Gueye et al. (eds.), *Mathematics of Computer Science, Cybersecurity and Artificial Intelligence*, Springer Proceedings in Mathematics & Statistics 468,
https://doi.org/10.1007/978-3-031-66222-5_4

power-commutative algebras [11], including quadratic algebras [15, 16, 27] and flexible algebras [4, 10]. The corollary of Hopf's theorem proves that every finite-dimensional real commutative division algebra with unit is isomorphic to $\mathbb{R}$ or $\mathbb{C}$ [21]. The Yang–Petro's theorem shows that every two-dimensional real division algebra with unit is isomorphic to $\mathbb{C}$ [29, 30]. Every real Cayley algebra (different from $\mathbb{R}$) has degree 2 [8].

A well-known Albert's result [1, Theorem 2] asserts that an algebra $\mathcal{A}$ is power associative if and only if $\mathcal{A}$ satisfies the identities $(x, x, x) = 0$ and $(x^2, x, x) = 0$. The four-dimensional third power-associative division algebras have been studied [13, 14].

By definition an absolute-valued algebra is an algebra $\mathcal{A}$ over $\mathbb{K}(\mathbb{K} = \mathbb{R}$ or $\mathbb{C})$ endowed with an absolute value, i.e., a norm $\| \cdot \|$ on the vector space of $\mathcal{A}$ satisfying $\|xy\| = \|x\|\|y\|$ for all $x, y \in \mathcal{A}$. Let $\mathbb{A} \in \{\mathbb{C}, \mathbb{H}, \mathbb{O}\}$. We recall that $^*\mathbb{A}$ is obtained by endowing the normed space $\mathbb{A}$ with the product $x \cdot y = \overline{x}y$, where $x \mapsto \overline{x}$ means the standard involution. In [9], the authors studied finite-dimensional absolute-valued algebras containing a generalized left unit. We also note that if $\mathcal{A}$ is a pre-Hilbert absolute-valued algebra satisfying $(x^2, y, x^2) = 0$ and containing a generalized left unit e which is an idempotent, then $\mathcal{A}$ is finite dimensional with left unit e and is isomorphic to $\mathbb{R}$, $\mathbb{C}$, $^*\mathbb{C}$, $\mathbb{H}$, $^*\mathbb{H}$, $\mathbb{O}$, or $^*\mathbb{O}$ [12, Theorem 3].

Let $\mathcal{A}$ be an algebra satisfying the identity $(x^p, x^q, x^r) = 0$. In [13], the authors prove that if $\mathcal{A}$ is a division algebra of degree ≤ 4 with slightly generalized unit, then $\mathcal{A}$ is power associative. In [24], the authors show that if $\mathcal{A}$ is unitary, then $\mathcal{A}$ is third power associative. They also prove that if $\mathcal{A}$ is an algebraic unital algebra without divisors of zero of degree $\neq 8$, then $\mathcal{A}$ is quadratic. If $(x^p, x^q, x^r) = 0$ is asymmetric and the algebra $\mathcal{A}$ with no nonzero joint divisor of zero, containing a nonzero central idempotent and satisfying an identity $(x^{p'}, x^{q'}, x^{r'}) = 0$, where $(p', q', r') \notin \{(p, q, r), (3 - r, 3 - q, 3 - p)\}$, then $\mathcal{A}$ is a unital power-associative algebra [7].

Motivated by these results, we became interested in the study of algebraic algebras without divisors of zero satisfying the identity $(x^p, x^q, x^r) = 0$. We have brought about some conditions leading to the associativity of the powers (Theorems 4.1 and 4.2).

Notations and Preliminary Results

Notations Let $\mathcal{A}$ be an algebra over a field of characteristic zero. We denote by:

(1) $\mathcal{A}(x)$, the subalgebra generated by every element $x \in \mathcal{A}$.
(2) $[x, y]$, the commutator $xy - yx$ of $x, y \in \mathcal{A}$.
(3) (x, y, z), the associator $(xy)z - x(yz)$ of $x, y, z \in \mathcal{A}$.

Definitions An algebra $\mathcal{A}$ is called:

(1) Third power associative if $(x, x, x) = 0$ for all $x \in \mathcal{A}$.

(2) Power commutative if $\mathcal{A}(x)$ is commutative for all $x \in \mathcal{A}$.
(3) Quadratic if it has a unit element e and x^2, x, e are linearly dependent.
(4) Power associative if $\mathcal{A}(x)$ is associative for all $x \in \mathcal{A}$.
(5) A division algebra if $L_x : \mathcal{A} \to \mathcal{A} \;\; a \mapsto xa$ and $R_x : \mathcal{A} \to \mathcal{A} \;\; a \mapsto ax$ are bijective for all x in $\mathcal{A}\backslash\{0\}$.
(6) Algebraic if $\mathcal{A}(x)$ is finite dimensional for all $x \in \mathcal{A}$ and the bigger m such that $m = dim(\mathcal{A}(x))$ is called the degree of $\mathcal{A}$. We note $deg(\mathcal{A}) = m$.

Let $\mathcal{A}$ be an algebra with two-dimensional subalgebras. A nonzero idempotent of $\mathcal{A}$ is said to be omnipresent if it belongs to every subalgebra of dimension two [11, page 1208].

A nonzero element e of an algebra $\mathcal{A}$ is called a generalized left unit if $e(xy) = x(ey)$ for all $x, y \in \mathcal{A}$. It is called a generalized unit if it satisfies both equalities $e(xy) = x(ey)$ and $(xy)e = (xe)y$ for all $x, y \in \mathcal{A}$ [9].

Let $a \in S(\mathbb{H})$ and $b \in S(\mathbb{H})\backslash\{-1, 1\}$, and we denote by $\mathbb{H}(a, b)$ and ${}^*\mathbb{H}(a, b)$ the algebras having $\mathbb{H}$ as underlying normed space and products $x \odot y$ given respectively by $axyb$ and $\overline{x}ayb$ [25]. We note that $\overline{a}$ is a generalized left unit of $\mathbb{H}(a, b)$ and a is a generalized left unit of ${}^*\mathbb{H}(a, b)$. However, none of these algebras contains a left unit.

Given an algebra $\mathcal{A}$ with product $(x, y) \mapsto xy$, the *opposite algebra* $\mathcal{A}^{(0)}$ of $\mathcal{A}$ is defined as the algebra consisting of the vector space of $\mathcal{A}$ and the product $(x, y) \mapsto yx$. Denoting by $(\cdot, \cdot, \cdot)^{(0)}$ the associator in the algebra $\mathcal{A}^{(0)}$ and $[\cdot, \cdot]^{(0)}$ the commutator in the algebra $\mathcal{A}^{(0)}$, it is clear that

$$(x, y, z)^{(0)} = -(z, y, x) \quad \text{for all} \quad x, y, z \in \mathcal{A}. \tag{4.1}$$

As a first consequence of (4.1), we deduce that

$$\mathcal{A} \quad \text{is associative} \quad \Leftrightarrow \quad \mathcal{A}^{(0)} \quad \text{is associative.} \tag{4.2}$$

As a second consequence of (4.1), we see that the following assertions are equivalent:

$$\begin{aligned} &\text{(i)} \;\; \mathcal{A} \text{ satisfies the identity } (x^p, x^q, x^r) = 0, \\ &\text{(ii)} \;\; \mathcal{A}^{(0)} \text{ satisfies the identity } (x^r, x^q, x^p) = 0. \end{aligned} \tag{4.3}$$

We note that $\mathcal{A}$ is power associative if and only if $\mathcal{A}^{(0)}$ is power associative [7], and $\mathcal{A}$ is quadratic if and only if $\mathcal{A}^{(0)}$ is quadratic.

Lemma 4.1 *Let $\mathcal{A}$ be a right division algebra having generalized left unit e which is an idempotent. If $\mathcal{A}$ satisfies the identity $(x, x^q, x^r) = 0$, then $\mathcal{A}$ is a unital third power-associative algebra.*

Proof For all x in $\mathcal{A}$, we have $e(xe) = xe$. We deduce that $L_e \circ R_e = R_e$, and consequently $L_e \circ R_e \circ R_e^{-1} = R_e \circ R_e^{-1}$. We realize that $L_e = I_{\mathcal{A}}$ and then e is a left unit. The result is a consequence of [13, Proposition 3]. □

Lemma 4.2 *Let $\mathcal{A}$ be a right division algebra having a generalized unit e which is an idempotent. If $\mathcal{A}$ satisfies the identity $(x^p, x^q, x^r) = 0$, then $\mathcal{A}$ is a unital third power-associative algebra.*

Proof It is clear that $e(xe) = xe$, and for the same arguments as before, we realize that e is a left-unit. As $(ex)e = ex$, we deduce that $xe = x$, and consequently e is the unit element of $\mathcal{A}$. This implies that $\mathcal{A}$ is third power associative [13, Proposition 2]. □

Lemma 4.3 *Let $\mathcal{A}$ be a right division algebra of degree $\neq 8$ having an idempotent e. Then $\mathcal{A}$ is power commutative in each one of the following cases:*

(1) *e is a generalized unit and $\mathcal{A}$ satisfies $(x^p, x^q, x^r) = 0$.*
(2) *e is a generalized left unit and $\mathcal{A}$ satisfies $(x, x^q, x^r) = 0$.*

Proof

(1) As e is a generalized unit, then e is a unit element and $\mathcal{A}$ is third power associative (Lemma 4.2). Let $x \in \mathcal{A}\backslash\{0\}$, and then $dim(\mathcal{A}(x)) \leq 4$. Also $\mathcal{A}(x)$ is a third power-associative division algebra, containing a central idempotent. We realize that $\mathcal{A}(x)$ is power commutative [14, Theorem 3], and consequently $\mathcal{A}(x)$ is commutative. Hence $\mathcal{A}$ is power commutative.
(2) As e is a generalized left unit, then e is a unit element and $\mathcal{A}$ is third power associative (Lemma 4.1). The same arguments as before conclude this case.

□

Main Results

Proposition 4.1 *Let $\mathcal{A}$ be a right division algebra of degree $\neq 8$, containing a nonzero idempotent e. Then the following conditions are equivalent:*

(1) *$\mathcal{A}$ satisfies $(x, x^q, x^r) = 0$ and e is a generalized left unit.*
(2) *$\mathcal{A}$ is quadratic.*

Proof The implication 2) $\Rightarrow$ 1) is obvious.

1) $\Rightarrow$ 2). Let $x \in \mathcal{A}\backslash\{0\}$. Lemmas (4.1 and 4.3) prove that $\mathcal{A}(x)$ is commutative division algebra with unit e. This implies that $\mathcal{A}(x)$ is isomorphic to $\mathbb{R}$ or $\mathbb{C}$ [20, 21]. Consequently, $\mathcal{A}$ is quadratic. □

Let $\mathbb{A}$ be one of the algebras $\mathbb{C}$, $\mathbb{H}$, or $\mathbb{O}$. The algebra $^*\mathbb{A}$, with multiplication $x \odot y = \overline{x}y$, contains a generalized left unit and satisfies all identities $(x^2, x^q, x^r) = 0$, but it is not a third power-associative algebra. So the hypothesis of e is a right-generalized unit in Proposition 4.2 can be removed. We have the following result.

Proposition 4.2 *Let $\mathcal{A}$ be a right division algebra of degree $\neq$ 8, containing a nonzero idempotent e. Then the following conditions are equivalent:*

(1) *$\mathcal{A}$ satisfies $(x^p, x^q, x^r) = 0$ and e is a generalized unit.*
(2) *$\mathcal{A}$ is quadratic.*

Proof The implication 2) $\Rightarrow$ 1) is clear.

1) $\Rightarrow$ 2). Let $x \in \mathcal{A}\backslash\{0\}$. Lemmas (4.2) and (4.3) prove that $\mathcal{A}(x)$ is commutative division algebra with unit e, so $\mathcal{A}(x)$ is isomorphic to $\mathbb{R}$ or $\mathbb{C}$ [20, 21]. We deduce that $\mathcal{A}$ is quadratic. □

Corollary 4.1 *Let $\mathcal{A}$ be a right division algebra of degree $\neq$ 8, containing a nonzero idempotent e. The following assertions are equivalent:*

(1) *$\mathcal{A}$ is quadratic.*
(2) *$\mathcal{A}$ satisfies $(x, x^q, x^r) = 0$ and e is a generalized left unit.*
(3) *$\mathcal{A}$ satisfies $(x^p, x^q, x^r) = 0$ and e is a generalized unit.*

Theorem 4.1 *Let $\mathcal{A}$ be an algebraic algebra without divisors of zero of degree $\neq$ 8 and containing a nonzero idempotent e such that $[e, f] = 0$ for all $f \in I(\mathcal{A})$. The following assertions are equivalent:*

(1) *$\mathcal{A}$ is quadratic.*
(2) *$\mathcal{A}$ is power associative.*
(3) *$\mathcal{A}$ satisfies $(x, x^q, x^r) = 0$ and e is a generalized left unit.*
(4) *$\mathcal{A}$ satisfies $(x^p, x^q, x) = 0$ and e is a generalized right unit.*
(5) *$\mathcal{A}$ satisfies $(x^p, x^q, x^r) = 0$ and e is a generalized unit.*

Proof The implications 1) $\Rightarrow$ 2), 1) $\Rightarrow$ 3), 1) $\Rightarrow$ 4), and 1) $\Rightarrow$ 5) are obvious.

2) $\Rightarrow$ 3). It is clear that $\mathcal{A}$ satisfies $(x, x^q, x^r) = 0$ for fixed $q, r \in \{1, 2\}$. Now let $x \in \mathcal{A}\backslash\{0\}$. The subalgebra $\mathcal{A}(x)$ is a finite-dimensional associative algebra, so $\mathcal{A}(x)$ is isomorphic to $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$ [18]. As $deg(\mathbb{H}) = 2$, we deduce that $\mathcal{A}(x)$ is isomorphic to $\mathbb{R}$ or $\mathbb{C}$. Let f the unit element of $\mathcal{A}(x)$. We assume that $e - f \neq 0$. Then $\mathcal{A}(e-f)$ is isomorphic to $\mathbb{R}$ or $\mathbb{C}$. There is $a \in \mathcal{A}(e-f)$ such that $(e-f)a = e-f = e^2 - f^2 = (e-f)(e+f)$. Then $a = e+f \in \mathcal{A}(e-f)$, and consequently $e, f \in \mathcal{A}(e-f)$. We realize that $\mathcal{A}(e-f)$ is unitary and contains at least two idempotents, absurd. We deduce that $e = f$, so $ex = xe = x$, for all $x \in \mathcal{A}$. We realize that e is a generalized left unit and $\mathcal{A}$ satisfies $(x, x^q, x^r) = 0$.

3) $\Rightarrow$ 1). Let $x \in \mathcal{A}\backslash\{0\}$. Then $\mathcal{A}(x)$ is finite-dimensional division algebra, so $\mathcal{A}$ contains a nonzero idempotent f [28]. As $fe = ef = ef^2 = f(ef)$, we have $ef = e = e^2$. This implies that $e = f$, consequently $e \in \mathcal{A}(x)$. Lemmas (4.1 and 4.3) prove that $\mathcal{A}(x)$ is power commutative with unit e, so $\mathcal{A}(x)$ is commutative, which implies that $\mathcal{A}(x)$ is isomorphic to $\mathbb{R}$ or $\mathbb{C}$ [20, 21]. This proves that $\mathcal{A}$ is quadratic.

4) $\Rightarrow$ 1). Suppose that $\mathcal{A}$ satisfies $(x^p, x^q, x) = 0$ and e is a generalized right unit. Then $\mathcal{A}^{(0)}$ satisfies $(x, x^q, x^p) = 0$, and e is a generalized left unit of $\mathcal{A}^{(0)}$. As e is an idempotent of $\mathcal{A}^{(0)}$ such that $[e, f]^{(0)} = 0$ for all $f \in I(\mathcal{A}^{(0)})$ and 3) $\Rightarrow$ 1), we deduce that $\mathcal{A}^{(0)}$ is quadratic, so $\mathcal{A}$ is quadratic.

5) $\Rightarrow$ 1). Let $x \in \mathcal{A}\backslash\{0\}$. The subalgebra $\mathcal{A}(x)$ is finite-dimensional division algebra, so $\mathcal{A}(x)$ contains a nonzero idempotent f [28]. For the same argument as before, we prove that $f = e \in \mathcal{A}(x)$. Then Lemmas (4.2) and (4.3) prove that $\mathcal{A}(x)$ is power commutative with unit e. We realize that $\mathcal{A}(x)$ is finite-dimensional commutative division algebra with unit. Hence $\mathcal{A}(x)$ is isomorphic to $\mathbb{R}$ or $\mathbb{C}$ [20, 21]. Consequently, $\mathcal{A}$ is quadratic. □

We have the following preliminary results.

Lemma 4.4 *Let $\mathcal{A}$ be an algebraic algebra without divisors of zero of degree $\neq 8$, containing a nonzero single idempotent e. Then the following conditions are equivalent:*

(1) *$\mathcal{A}$ is quadratic.*
(2) *$\mathcal{A}$ is power associative.*
(3) *$\mathcal{A}$ satisfies $(x, x^q, x^r) = 0$ and e is a generalized left unit.*
(4) *$\mathcal{A}$ satisfies $(x^p, x^q, x^r) = 0$ and e is a generalized unit.*

Proof The implications 1) $\Rightarrow$ 2), 1) $\Rightarrow$ 3), and 1) $\Rightarrow$ 4) are obvious.

2) $\Rightarrow$ 1). Let $x \in \mathcal{A}\backslash\{0\}$. The subalgebra $\mathcal{A}(x)$ is finite-dimensional associative division algebra. This implies that $\mathcal{A}(x)$ is isomorphic to $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$ [18]. As $deg(\mathbb{H}) = 2$, we deduce that $\mathcal{A}(x)$ is isomorphic to $\mathbb{R}$ or $\mathbb{C}$. We realize that e is the unit element of $\mathcal{A}(x)$, so $ex = xe = x$ for all $x \in \mathcal{A}$, and consequently, $\mathcal{A}$ is quadratic.

3) $\Rightarrow$ 1). Let $x \in \mathcal{A}\backslash\{0\}$, and then $\mathcal{A}(x)$ is finite-dimensional division algebra containing e [28]. Lemmas (4.1) and (4.3) prove that $\mathcal{A}(x)$ is power commutative with unit e (e is a generalized unit), and we realize that $\mathcal{A}(x)$ is commutative, which implies that $\mathcal{A}(x)$ is isomorphic to $\mathbb{R}$ or $\mathbb{C}$ [20, 21], so $\mathcal{A}$ is quadratic.

4) $\Rightarrow$ 1). Let $x \in \mathcal{A}\backslash\{0\}$. The subalgebra $\mathcal{A}(x)$ is finite-dimensional division algebra, so $\mathcal{A}(x)$ contains e [28]. Lemmas (4.2 and 4.3) prove that $\mathcal{A}(x)$ is power commutative with unit e. We realize that $\mathcal{A}(x)$ is finite-dimensional commutative division algebra with unit, so $\mathcal{A}(x)$ is isomorphic to $\mathbb{R}$ or $\mathbb{C}$ [20, 21]. This implies that $\mathcal{A}$ is quadratic. □

Lemma 4.5 *Let $\mathcal{A}$ be an algebra without divisors of zero such that $\mathcal{A}(x)$ is isomorphic to either $\mathbb{R}$ or $\mathbb{C}$ for all nonzero x in $\mathcal{A}$. Then $\mathcal{A}$ has only a single nonzero idempotent.*

Proof Let $e \in \mathcal{A}$ be an arbitrary nonzero idempotent, and assume that $\mathcal{A}$ contains a nonzero idempotent $f \neq e$ with $[f, e] = 0$. The operator $L_{e-f} : y \mapsto (e - f)y$ of the division algebra $\mathcal{A}(e - f)$ is bijective. As $e - f = (e - f)(e + f)$, the element $e + f = L_{e-f}^{-1}(e - f)$ belongs to $\mathcal{A}(e - f)$, and it must be its unit element that we note 1_{e-f}. Now,

$$e = \frac{(e + f) + (e - f)}{2}, \quad f = \frac{(e + f) - (e - f)}{2}$$

belong to $\mathcal{A}(e-f)$. This implies that the division unital algebra $\mathcal{A}(e-f)$ contains three distinct nonzero idempotents: $1_{e-f} = e+f$, e, f, absurd.

So two arbitrary nonzero commuting idempotents are equal.

Now suppose that $\mathcal{A}$ contains a nonzero idempotent $f \neq e$. We have $(e-f)^2 = 2(e+f) - (e+f)^2 \in \mathcal{A}(e-f) \cap \mathcal{A}(e+f)$. So $(e-f)^2$ commutes with both $e-f$ and $e+f$, and then $[(e-f)^2, e] = [(e-f)^2, f] = 0$. We distinguish the following two cases:

(1) If $\mathcal{A}(e-f) = \mathcal{A}(e+f)$, then e, f belong to the commutative algebra $\mathcal{A}(e-f)$, absurd.
(2) If $\mathcal{A}(e-f) \neq \mathcal{A}(e+f)$, then $\dim\big(\mathcal{A}(e-f) \cap \mathcal{A}(e+f)\big) = 1$, and there is a nonzero idempotent $g \in \mathcal{A}$ such that $(e-f)^2 = \lambda g$ with $\lambda \in \mathbb{R} \setminus \{0\}$. Now, we have $[g, e] = \lambda^{-1}[(e-f)^2, e] = 0$ and $[g, f] = \lambda^{-1}[(e-f)^2, f] = 0$ leading to the absurdity $e = g = f$.

Note that the single nonzero idempotent of $\mathcal{A}$ is the unit element of any subalgebra $\mathcal{A}(x)$, $x \in \mathcal{A} \setminus \{0\}$. □

Theorem 4.2 *Let $\mathcal{A}$ be an algebraic algebra without divisors of zero of degree $\neq 8$ and containing a nonzero omnipresent idempotent e. The following assertions are equivalent:*

(1) *$\mathcal{A}$ is quadratic.*
(2) *$\mathcal{A}$ is power associative.*
(3) *$\mathcal{A}$ satisfies $(x, x^q, x^r) = 0$ and e is a generalized left unit.*
(4) *$\mathcal{A}$ satisfies $(x^p, x^q, x^r) = 0$ and e is a generalized unit.*
(5) *$\mathcal{A}$ satisfies $(x^p, x^q, x) = 0$ and e is a generalized right unit.*

Proof The implications 1) ⇒ 2) and 1) ⇒ 5) are obvious.

2) ⇒ 3). Let $x \in \mathcal{A}\backslash\{0\}$. Then $\mathcal{A}(x)$ is a finite-dimensional associative division algebra. The Frobenius's theorem proves that $\mathcal{A}(x)$ is isomorphic to $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$ [18]. As $deg(\mathbb{H}) = 2$, we deduce that $\mathcal{A}(x)$ is isomorphic to $\mathbb{R}$ or $\mathbb{C}$. Lemma 4.5 proves that e is the single nonzero idempotent of $\mathcal{A}$. Lemma 4.4 concludes this case.

3) ⇒ 4). Let $x \in \mathcal{A}\backslash\{0\}$. Then $\mathcal{A}(x)$ is a division algebra such that $dim(\mathcal{A}(x)) \leq 4$. We will distinguish the following cases:

First case. If $dim(\mathcal{A}(x)) = 1$, it is clear that $\mathcal{A}(x)$ is isomorphic to $\mathbb{R}$ [26, 27].

Second case. If $dim(\mathcal{A}(x)) = 2$, then $e \in \mathcal{A}(x)$. As $\mathcal{A}(x)$ satisfies $(y, y^q, y^r) = 0$ for all $y \in \mathcal{A}(x)$, Lemma 4.1 proves that $\mathcal{A}(x)$ is unitary. We deduce that $\mathcal{A}(x)$ is isomorphic to $\mathbb{C}$ [29, 30].

Third case. If $dim(\mathcal{A}(x)) = 4$, we have $e \in \mathcal{A}(x)$ and $deg(\mathcal{A}(x)) = 4$. As $\mathcal{A}(x)$ satisfies $(y, y^q, y^r) = 0$ for all $y \in \mathcal{A}(x)$, Lemma 4.1 proves that $\mathcal{A}(x)$ is unitary and third power associative, absurd [13, corollary2] and [24, Proposition 5.1].

Consequently, $\mathcal{A}(x)$ is isomorphic to $\mathbb{R}$ or $\mathbb{C}$ for all $x \in \mathcal{A}\backslash\{0\}$. Lemma 4.5 implies that e is the single nonzero idempotent of $\mathcal{A}$. Finally, Lemma 4.4 concludes this case.

4) ⇒ 1). Let $x \in \mathcal{A}\backslash\{0\}$. Then $\mathcal{A}(x)$ is a division algebra such that $dim(\mathcal{A}(x)) \leq 4$. We have the following cases:

First case. If $dim(\mathcal{A}(x)) = 1$, it is obvious that $\mathcal{A}(x)$ is isomorphic to $\mathbb{R}$ [26, 27].

Second case. If $dim(\mathcal{A}(x)) = 2$, we deduce that $e \in \mathcal{A}(x)$. It is clear that $\mathcal{A}(x)$ satisfies $(y, y^q, y^r) = 0$ for all $y \in \mathcal{A}(x)$. Lemma 4.2 proves that $\mathcal{A}(x)$ is unitary, so $\mathcal{A}(x)$ is isomorphic to $\mathbb{C}$ [29, 30].

Third case. If $dim(\mathcal{A}(x)) = 4$, we have $e \in \mathcal{A}(x)$ and $deg(\mathcal{A}(x)) = 4$. As $\mathcal{A}(x)$ satisfies $(y, y^q, y^r) = 0$ for all $y \in \mathcal{A}(x)$, Lemma 4.2 implies that $\mathcal{A}(x)$ is third power associative with unit e, absurd [13, corollary2] and [24, Proposition 5.1].

We realize that $\mathcal{A}(x)$ is isomorphic to $\mathbb{R}$ or $\mathbb{C}$ for all $x \in \mathcal{A}\backslash\{0\}$. Lemma 4.5 proves that e is the single nonzero idempotent of $\mathcal{A}$. In last, Lemma 4.4 proves that $\mathcal{A}$ is quadratic.

5) $\Rightarrow$ 1). Suppose that $\mathcal{A}$ satisfies $(x^p, x^q, x) = 0$ and e is a generalized right unit. This implies that $\mathcal{A}^{(0)}$ satisfies $(x, x^q, x^p) = 0$ and e is an omnipresent generalized left unit of $\mathcal{A}^{(0)}$. As 3) $\Rightarrow$ 1), we deduce that $\mathcal{A}^{(0)}$ is quadratic, so $\mathcal{A}$ is quadratic. □

Corollary 4.2 *Let $\mathcal{A}$ be an algebraic algebra without divisors of zero of degree $\neq 8$ with left unit. The following assertions are equivalent:*

(1) *$\mathcal{A}$ is quadratic.*
(2) *$\mathcal{A}$ is power associative.*
(3) *$\mathcal{A}$ satisfies $(x, x^q, x^r) = 0$.*

Acknowledgments The authors would like to thank the referees for their careful reading of this paper and their useful suggestions.

References

1. A. A. Albert, *Power associative rings.* T. AMS **64**, (1948) 552–593.
2. S. C. Althoen and L. D. Kugler, *When is $\mathbb{R}^2$ a division algebra?* Amer. Math. Monthly **90**, (1983) 625–635.
3. G. M. Benkart, D. J. Britten, and J. M. Osborn, *On applications of isotopy to real division algebras*. Hadronic J. **4** (1981), 497–529.
4. G. M. Benkart, D. J. Britten, and J. M. Osborn, *Real flexible division algebras.* Canad. J. Math. **34** (1982), 550–588
5. R. Bott and J. Milnor, *On the parallelizability of the spheres.* Bull. Amer. Math. Soc. **64** (1958), 8789.
6. I. Burdujan, *Types of nonisomorphic two-dimensional real division algebras.* Proceedings of the national conference on algebra (Romanian), An. Stiint. Univ. Al. I. Cuza Iasi Sect. I a Mat. **31** (1985), suppl., 102–105.
7. M. Cabrera and A. Diouf, *On power-associativity of algebras with no nonzero joint divisor of zero and containing a nonzero central idempotent.* Commun. Algebra **52** (2024), 295–304
8. M. Cabrera and Á Rodríguez, *Non-associative Normed Algebras Volume 1: The Vidav-Palmer and Gelfand- Naimark Theorems*. Part of Encyclopedia of Mathematics and Its Applications, Cambridge University Press, 2014.
9. A. Chandid and A. Rochdi, *Mutations of absolute valued algebras.* Int. J. Algebra, **2**, (2008) 357–368.

10. J. A. Cuenca, R. De Los Santos Villodres, A. Kaidi, and A. Rochdi, *Real quadratic flexible division algebras.* Linear Algebra Appl. **290**, (1999) 1–22.
11. E. Darpö, and A. Rochdi, *Classification of the four-dimensional power-commutative real division algebras.* Proceedings of the Royal Society of Edinburgh, **141** A, (2011) 1207–1223.
12. K. Diaby, O. Diankha, A. Fall, and A. Rochdi, *On absolute-valued algebras satisfying* $(x^2, y, x^2) = 0$. J. Algebra **585** (2021), 484–500.
13. O. Diankha, A. Diouf, M. Traoré and A. Rochdi, *Division Algebras satisfying* $(x^p x^q)x^r = x^p(x^q x^r)$. Int. J. Algebra **7**(20), (2013) 959–972.
14. O. Diankha, M. Traoré, M. I. Ramirez and A. Rochdi, *Four-Dimensional Real third-power Associative Division Algebras* . Commun. Algebra **44**, (2016) 3397–3406.
15. E. Dieterich, *Real quadratic division algebras.* Commun. Algebra **28**, (2000) 941–947.
16. E. Dieterich, *Quadratic division algebras revisited.* Proceedings of the American Mathematical Society, **128** (11) (2000), 3159–3166.
17. E. Dieterich, *Classification, automorphism groups and categorical structure of the two-dimensional real division algebras.* J. Algebra Appl., **4** (2005), 517–538.
18. F. G. Frobenius, *Über lineare substitutionen und bilineare formen.* J. Reine Angew. Math. **84** (1878), 1–63.
19. E. Gottschling, *Die zweidimensionalen reellen Divisionsalgebren*, Seminarber. Fachb. Math. FernUniversitat-GHS in Hagen **63** (1998), 228–261.
20. F. Hirzebruch, M. Koecher and R. Remmert, *Numbers.* Springer-Verlag, (1991).
21. H. Hopf, *Ein topologischer beitrag zur reellen algebra.* Comment. Math. Helvet. (1940), 219–239.
22. M. Hübner and H. P. Petersson, *Two-dimensional real division algebras revisited.* Beiträge Algebra Geom., **45**, (2004) 29–36.
23. M. Kervaire, *Non-parallelizability of the n-sphere for* $n > 7$. Proc. Nat. Acad. Sci. USA **44** (1958), 280–283.
24. A. Moutassim, M. Benslimane, *Generalization of the Hopf commutative theorem* . Commun. Alg **45**, (2017) 883–888.
25. M. I. Ramírez, On four-dimensional absolute valued algebras, in Proceedings of the International Conference on Jordan (Málaga, 1997), University of Málaga, Málaga, 1999, pp. 169–173.
26. A. Rodríguez, *Absolute valued algebras, and absolute valuable Banach spaces.* Advanced courses of mathematical analysis I, 99–155, World Sci. Publ., Hackensack, NJ, (2004).
27. R. D. Schafer, *An introduction to nonassociative algebras.* Academic Press, New York (1966).
28. B. Segre, La teoria delle algebre ed alcune questione di realta. Univ. Roma, Ist. Naz. Alta. Mat., Rend. Mat. E App. Serie 5 **13** (1954) 157–188.
29. J. Petro, *Real division algebras of dimension* > 1 *contain* $\mathbb{C}$. Amer. Math. Monthly, **94** (1987), 445–449.
30. C. T. Yang, *Division algebras and fibrations of spheres by great spheres.* J. Differential Geometry, **16** (1981), 577–593.

Chapter 5
Computing Minimal Free Resolutions over Monomial Semirings with Coefficients in D-A Rings

Guy Mobouale Wamba, Soda Diop, and Djiby Sow

Abstract The study of Gröbner–Shirshov bases in a field, where the set of monomials is a semiring, was first extended to valuation rings by Yatma et al. In a subsequent development, S.Diop et al. further generalized these methods to the setting of D-A rings (divisible and annihilable rings), preserving the semiring structure for the set of monomials, whether commutative or not. Using the approach introduced by Diop et al. in the commutative case, we propose a new technique for computing a minimal free resolution for the ideal I as an R module. This extension of the method shows the variety and applicability of the Gröbner–Shirshov basis framework in various algebraic settings.

Keywords Divisible and annihilable ring · Semiring · Gröbner–Shirshov basis · Free resolution

2020 Mathematics Subject Classification 13P10, 13C10, 13P25

Introduction

In commutative algebra and algebraic geometry, the study of ideals and their properties is very important [2, 13]. Ideals are fundamental in understanding the geometric and algebraic structures associated with polynomial rings. One crucial aspect of studying ideals is their resolution, which provides valuable information about their structure and behavior. In particular, the minimal free resolution offers a powerful tool to investigate ideals' complex nature and associated modules. When investigating ideals within polynomial rings, one fundamental question arises: How can we describe an ideal in terms of a simpler, more transparent structure? The concept of resolution provides an answer to this question. A resolution of an ideal

---

G. M. Wamba (✉) · S. Diop · D. Sow
Département de Mathématiques et Informatique, Faculté des Sciences et Techniques, Université Cheikh Anta Diop, Dakar, Sénégal

C. T. Gueye et al. (eds.), *Mathematics of Computer Science, Cybersecurity and Artificial Intelligence*, Springer Proceedings in Mathematics & Statistics 468,
https://doi.org/10.1007/978-3-031-66222-5_5

allows us to break it down into simpler components or modules, providing a detailed description of the ideal's structure [3].

A minimal free resolution is a special type that captures an ideal's essential properties while minimizing the complexity of the modules involved. It aims to represent an ideal as a sequence of free modules, each connected by a homomorphism, which preserves certain algebraic properties [3, 14]. Importantly, minimal free resolutions possess several desirable properties that make them invaluable in various areas of mathematics, including algebraic geometry, commutative algebra, and algebraic topology. It is for example used to compute the cohomology group $\mathrm{Ext}_s^i(A, B)$ for further modules A, B [6, 8].

In the context of minimal free resolution, one powerful tool for constructing such resolutions and studying ideals is the theory of Gröbner bases. Gröbner bases systematically generate canonical representatives for the ideals, allowing us to perform computations and simplify the resolution process. The algorithmic nature of Gröbner bases facilitates the practical implementation and computational exploration of minimal free resolutions. It has been widely studied for submodules with coefficients over fields (see [1, 5, 9, 10]).

More recently, in 2019, S.Diop et al. [12] have been computing a minimal free resolution of $\mathbb{Z}$-ideal $\mathbb{Z}(I)_R$ of R as R module from ideal I of R where $R = \mathbb{Z}[x_1, \ldots, x_n]$.

In 2020, Ihsen Yengui et al. [7] computed a free resolution for an ideal of the Bézout rings.

In this chapter, we are interested in studying this problem in a polynomial semiring over the D-A rings. To do this, we revisit the method introduced in [11] for computing Gröbner–Shirshov bases for an ideal of semi-algebra in the commutative case. The main goal is to use this method to propose a technique for computing minimal free resolution for an ideal of semi-algebra. Indeed, we consider a semi-algebra $\mathrm{DRig}[X]$ where $\mathrm{Rig}[X] = (X^*, \cdot, 1, \theta)$, a commutative monomial semiring over $X = \{x_1, \ldots, x_n\}$ and D a D-A ring. This leads us to the Syzygy Theorem introduced in section "Properties of Gröbner–Shirshov Bases", and this theorem is very useful for this method and works as follows: Consider a representable order on D-A ring D. Using an admissible monomial order in the algebra $\mathrm{Rig}[X]$ and well-founded order in D-A ring D, we define a well-founded order on the semi-algebra $\mathrm{DRig}[X]$. Therefore, if $G = \{g_1, \ldots, g_k\}$ is a Gröbner–Shirshov basis for ideal of $\mathrm{DRig}[X]$ w.r.t. well-founded order, then Propositions 5.2 and 5.3 prove that the set of all syzygies obtained from $\mathcal{S}$-polynomials and $\mathcal{A}$-polynomials, $\mathrm{T} = \{T_1, \ldots, T_k\}$ with $k \geq 1$, is an $\mathcal{AS}$-reduced. These results are used to construct the Syzygy Theorem 5.2 (to obtain a Gröbner–Shirshov basis for syzygies submodule $\mathrm{syz}(g_1, \ldots, g_r)$ w.r.t. the order induced by well-founded order). Thus, it becomes easy to compute a free resolution for the ideal of semi-algebra $\mathrm{DRig}[X]$, and at each step, one uses some trick in [12] to obtain a minimal free resolution for an ideal of $\mathrm{DRig}[X]$.

In the last section of this chapter, we generalize the computation of free resolutions for a monomial semiring over D-A ring $\mathrm{DRig}[X]$. We end the paper with some examples, including an example where we compute a minimal free resolution on a semiring over a Gaussian integer ring modulo 12 (i.e., $\mathrm{D} = \mathbb{Z}_{12}[i]$, where $i = -1$).

Preliminaries and Notations

Let D be a commutative ring with the identity element 1_{D}. The ring D is called a divisible and annihilable ring, simply denoted by a $\mathrm{D} - \mathrm{A}$ ring, if:

(1) For each element in D, its representative form is computable.
(2) There exist a representable order < on D and a division algorithm, such that for any $a, b \in \mathrm{D} - \{0\}$, both $r = \mathrm{rem}(a, b) < b$ and $q = \mathrm{quot}(a, b)$ are computable such that $a = qb + r$.
(3) For any $c \in \mathrm{D} - \{0\}$, $\mathrm{ann}(c)$ is computable (note that $\mathrm{ann}(c)$ is the smallest generator of $\mathrm{ANN}(c)$, where $\mathrm{ANN}(c)$ is the set of the annihilators of c, and it is a principal ideal ring for a $\mathrm{D} - \mathrm{A}$ ring D).

Recall that a $\mathrm{D} - \mathrm{A}$ ring is a Noetherian ring. The following rings $\mathbb{Z}_n$ and $\mathbb{Z}_n[i]$ can be viewed as D-A rings. For the properties of $\mathrm{D} - \mathrm{A}$ rings, see [4, 11].

Let X be a finite and non-empty set, and denote X^* the set of words over X. We denote by $(X^*, \cdot, 1)$ free monoid generated by X, where the unit element (denoted "1") is the empty word. The semiring generated by X is $(X^*, \circ, \cdot, \theta, 1)$, where $(X^*, \circ, \theta)$ is a commutative monoid with neutral element θ and the law $\cdot$ is distributive with respect to $\circ$ from left and right. We denote the semiring by $\mathrm{Rig}[X] = (X^*, \cdot, \circ)$.

Total order $\prec$ on X^* is admissible if it satisfies the following condition: For any $u, v, w, t \in X^*, u \prec v \Rightarrow wut \prec wvt$.

An ordering $\prec$ on $\mathrm{Rig}\langle X\rangle$ is called a monomial ordering if:

(a) It is a total ordering, and it means u, v are always comparable under $\prec$ for any $u, v \in \mathrm{Rig}[X]$.
(b) It is a well ordering; there is no infinite decreasing sequence in $\mathrm{Rig}[X]$ with respect to $\prec$.
(c) It is compatible with the semiring structure: For any $u, v, w \in \mathrm{Rig}[X]$, $t', w' \in X^*$, we have

$$u \prec v \Rightarrow \begin{cases} u \circ w \prec v \circ w \\ t'uw' \prec t'vw' \end{cases}.$$

An admissible monomial order on $\mathrm{Rig}[X]$ and a representable order on D defines a well-founded order $<$ on polynomials in $\mathrm{DRig}[X]$ in a natural way.

For the properties and the notations on the semiring $\mathrm{Rig}[X]$, see [11].

In the following, we assume that D is a $\mathrm{D} - \mathrm{A}$ ring, and the groupoid algebra $\mathrm{DRig}[X]$ is the semiring algebra in the not necessarily commutative variables $X = \{x_1, \cdots, x_k\}$ with $k \geq 1$. Further, an admissible monomial order is defined such that we can define the leading monomial, the leading coefficient, and the leading term of a given polynomial $p \in \text{DRig}[X]$, denoted by $\text{Lm}(p)$, $\text{LC}(p)$, and $\text{LT}(p)$, respectively. Moreover, we denote $p - \text{LT}(p)$ by $\text{Rest}(p)$.

For convenience, if we say $\text{LT}(p) = cm$ is a term in $\text{DRig}[X]$, then usually it means that $c \in D - \{0\}$ is the leading coefficient and that m is the leading monomial in $\text{Rig}[X]$. Let m_1 and m_2 be monomials in $\text{Rig}[X]$, and we say $m_1|m_2$ if there exist $s, s' \in X^*$ and $u \in \text{Rig}[X]$ such that : $m_2 = sm_1s' \circ u$.

Moreover, let $\text{LT}(p_1) = c_1m_1$ and $\text{LT}(p_2) = c_2m_2$ be terms in $\text{DRig}\langle X\rangle$, and we say $\text{LT}(p_1)|\text{LT}(p_2)$ if $c_1|c_2$ in D and $m_1|m_2$ in $\text{Rig}[X]$.

Let $<$ be an admissible order on $\text{DRig}[X]$. For any two polynomials $f, g \in \text{DRig}[X]$, we say that $f < g$ if and only if:

(1) $\text{LM}(f) < \text{LM}(g)$ or
(2) $\text{LM}(f) = \text{LM}(g)$ and $\text{LC}(f) < \text{LC}(g)$ or
(3) $\text{LC}(f) = \text{LC}(g)$ and $\text{Rest}(f) < \text{Rest}(g)$

Let $f, g, p \in \text{DRig}[X]$ be three polynomials. Then, we say that:

(1) f is reduced to g modulo p, denoted $f \longrightarrow_p g$, if there exists a term am in f such that:

- $\text{Lm}(p)$ divides m, i.e., there exist $s, s \in X^*$, $u \in \text{Rig}[X]$ such that $m = s\text{LM}(p)s' \circ u$
- a is reducible modulo $\text{LC}(p)$, i.e., $a = q.\text{LC}(p) + r$ with $r = \text{rem}(a, \text{LC}(p)) < a$ and $q = \text{quot}(a, \text{LC}(p))$

and $g = f - qsps' \circ u$.

Put $f = am + f_1$, and then $g = rm + f_1 - qs\text{Rest}(p)s' \circ u$.

(2) f is reduced to g modulo G where G is a polynomial set, denoted $f \longrightarrow_G g$, if there exists p in G such that $f \longrightarrow_p g$.
(3) $g \in \text{DRig}\langle X\rangle$ is the form normal of f modulo G if $f \longrightarrow_G g$ and g not is reducible modulo G. Note that $\longrightarrow_G$ is Noetherian.
4) We denote by $\longrightarrow^*$ the reflexive and transitive closure of $\longrightarrow$, and $\longrightarrow^+$ is the transitive cloture of $\longrightarrow$. Moreover, $\longrightarrow^k$ means k consecutive steps of $\longrightarrow$ for a given integer k. We denote also by $\longleftrightarrow^*$ the smallest equivalence relation containing $\longleftrightarrow$.

Definition 5.1 (Gröbner–Shirshov Basis) Let G be a finite set of polynomials in $\text{DRig}[X]$. The set G is a Gröbner–Shirshov basis of $I = \langle G\rangle$ if every polynomial in $I = \langle G\rangle$ can reduce to 0 modulo G.

Given $c_1, c_2 \in D - \{0\}$, two comparable elements, the bigger element is reducible modulo the smaller. If $\text{rep}(c_1) \geq \text{rep}(c_2)$, then c_1 is reducible modulo c_2, i.e., $\text{rem}(c_1, c_2) < c_1$. We have the following definition:

Definition 5.2 ($\mathcal{S}$-polynomial) Let $p_i = c_im_i + \text{Rest}(p_i)$ with $\text{LM}(p_i) = m_i$ and $\text{LC}(p_i) = c_i$ for $i = 1, 2$. Let $\text{rep}(c_1) \geq \text{rep}(c_2)$, and if there exist $s_1, s_2 \in X^*$ and $u_1, u_2 \in \text{Rig}[X]$ such that

$$\omega = \mathrm{LCM}(s_1m_1, s_2m_2) = s_1m_1 \circ u_1 = s_2m_2 \circ u_2,$$

then we define the $\mathcal{S}$-polynomial of p_1 and p_2 with respect to ω as follows:

$$\mathcal{S}(p_1, p_2)_\omega = s_1p_1 \circ u_1 - qs_2p_2 \circ u_2.$$

ω is ambiguity of the $\mathcal{S}$-polynomial $\mathrm{S}(p_1, p_2)_\omega$, and $q = \mathrm{quot}(c_1, c_2)$.

Let $f = cm + \mathrm{Rest}(f) \in \mathrm{DRig}[X]$. The $\mathcal{A}$-polynomial of f is the polynomial

$$\mathrm{Apol}(f) = \mathrm{ann}(c).f = \mathrm{ann}(c).\mathrm{Rest}(f).$$

If the polynomial f in $\mathrm{DRig}[X]$ is nonzero and $\mathrm{ann}(\mathrm{LC}(f)) = 0$, then $\mathrm{Apol}(f) = 0$.

Let $g_1 = c_1m_1 + \mathrm{Rest}(g_1)$ be another polynomial in $\mathrm{DRig}[X]$ such that $g_1 = \mathrm{Apol}(f) = \mathrm{ann}(\mathrm{LC}(f)).(\mathrm{Rest}(f))$. If g_1 is nonzero in $\mathrm{DRig}[X]$, then we have $g_2 = \mathrm{Apol}(g_1) = \mathrm{Apol}^2(f)$ is the polynomial of g_1. If g_2 is nonzero, we can continue this process until $\mathrm{Apol}(g_k) = \mathrm{Apol}^{k+1}(f) = 0$ for some $k \in \mathbb{N}$. This process will terminate after at most n steps where n is the number of terms in f. We get a finite sequence of $\mathcal{A}$-polynomials, $g_1, \cdots, g_n$, where $g_k = \mathrm{Apol}^k(f) \neq 0$ for all $k = 1, \cdots, n$, and $\mathrm{Apol}(g_k) = 0$.

The notions of $\mathcal{A}$-polynomial are well known for Gröbner–Shirshov bases over the ring. They allow us to reduce a polynomial when its leading coefficient is a zero divisor. For the properties and the notations on $\mathcal{A}$-polynomials, see [4, 11].

Definition 5.3 Let G be a subset of semi-algebra $\mathrm{DRig}[X]$. Then G is:

(i) $\mathcal{A}$-reduced if every $\mathcal{A}$-polynomial in $\mathrm{SAP}(G)$ can reduce to 0 of the relation $\longrightarrow^*$.
(ii) $\mathcal{S}$-reduced if every $\mathcal{S}$-polynomial can reduce to 0 of the relation $\longrightarrow^*$, for every pair of polynomials of G.
(iii) $\mathcal{AS}$-reduced if it is $\mathcal{A}$-reduced and $\mathcal{S}$-reduced.

Properties of Gröbner–Shirshov Bases

In this section, we introduce key properties of Gröbner–Shirshov bases that will play a fundamental role in the following discussions. The results presented here are similar to those presented in a previous paper [11].

The results discussed in this section will form the basis of our further work. Rather than provide explicit proof in this section, we aim to succinctly articulate these properties and set the stage for their application and significance in the following parts of the paper.

Lemma 5.1 *Let G be a subset $\mathcal{S}$-reduced of* $\mathrm{DRig}[X]$ *and* $f \in \mathrm{DRig}[X]$ *such that* $f \longrightarrow^* 0$*, and there exists* $p \in G$ *such that* $\mathrm{LT}(p)$ *divides* $\mathrm{LT}(f)$.

Lemma 5.2 *Assume that G is $\mathcal{S}$-reduced. Let $p_i = c_i m_i + \text{Rest}(p_i)$, $i = 1, 2$, be the polynomials in G such that $s_i m_i \circ u_i = m$ with $m_i = \text{LM}(p_i)$, $s_i \in X^*$, and $u_i \in \text{Rig}[X]$. Then there exists $h \in G$ such that $\text{LM}(h)|m$ and $\text{LC}(h)|\text{rgcd}(\text{LC}(p_1), \text{LC}(p_2))$.*

Lemma 5.3 *Assume that G is $\mathcal{S}$-reduced. Let $p_i = c_i m_i + \text{Rest}(p_i)$, $i = 1, 2$, be the polynomials in G such that $s_i m_i \circ u_i = m$ with $m_i = \text{LM}(p_i)$, $s_i \in X^*$, and $u_i \in \text{Rig}[X]$. Then there exists $h \in G$ such that $\text{LM}(h)|m$ and $\text{LC}(h)|\text{rgcd}(\text{LC}(p_1), \text{LC}(p_2))$.*

Now we can extend the lemma 5.3 in the case of multiple polynomials in G. So the following lemma is its generalization.

Lemma 5.4 *Assume that G is $\mathcal{S}$-reduced. Let $p_i = c_i m_i + \text{Rest}(p_i)$, $1 \leq i \leq k$, be the polynomials in G such that $s_i m_i \circ u_i = m$ with $m_i = \text{LM}(p_i)$, $s_i \in X^*$ and $u_i \in \text{Rig}[X]$. Then there exists $h \in G$ such that $\text{LM}(h)|m$ and $\text{LC}(h)|\text{rgcd}(\text{LC}(p_1), \text{LC}(p_2), \cdots, \text{LC}(p_k))$.*

Now define the notion of the weak standard representation of a polynomial f. This is a rewriting of f with polynomials in G.

Definition 5.4 (Weak Standard Representation) Let $f \in \text{DRig}[X]$ and $p_i \in G$. The polynomial f has a weak standard representation w.r.t. G if there exist $s_{ij}, s'_{ij} \in X^*$ and $u_{ij} \in \text{Rig}[X]$ such that $f = \sum_{i=1}^{n} a_i s_i p_i \circ u_i$ with $a_i \in D$, which satisfy the following condition:

$$\text{LC}(a_i s_i p_i \circ u_i) \neq 0 \text{ and } \text{LM}(a_i s_i p_i \circ u_i) \leq \text{LM}(f)$$

Remark 5.1 for all $1 \leq i \leq n$. If $f = 0$, then we say $f = 0$ is the weak standard representation of f w.r.t. G.

The notion of weak standard representation is a powerful tool for understanding and characterizing elements within the semiring associated with a given Gröbner–Shirshov basis.

Lemma 5.5 *Let $f \in \text{DRig}[X]$ such that $f \longrightarrow^* 0$. Then f has a weak standard representation w.r.t. G.*

Lemma 5.6 *Let p be a polynomial in G and let $f = a.sp \circ u$ be a polynomial of $\text{DRig}[X]$, with $a \in D$, $s \in X^*$, $u \in \text{Rig}[X]$. If the subset G is $\mathcal{A}$-reduced, then f has a weak standard representation w.r.t. G.*

The following Theorem 5.1 provides a unified understanding of the conditions characterizing `Gröbner-Shirshov bases` in terms of $\mathcal{A}.\mathcal{S}$-reduction, weak standard representation, and leading term factorization.

Theorem 5.1 (Composition-Diamond Lemma for Commutative Semiring) *Let G be polynomials set in* DRig$[X]$ *and* $<$ *an well-founded order on* DRig$[X]$. *Then the following statements are equivalent:*

(1) *The set G is a Gröbner–Shirshov basis in* DRig$[X]$.
(2) *The set G is* $A.\mathcal{S}$*-reduced.*
(3) *For* $f \in I = \langle G \rangle$*, the polynomial* f *has a weak standard representation w.r.t.* G.
(4) *For* $f \in I = \langle G \rangle$*, there exists a polynomial h in G such that* $\mathrm{LT}(f) = s.\mathrm{LT}(h) \circ u$ *for some* $s, s' \in X^*$ *and* $u \in \mathrm{Rig}[X]$.

In Theorem 5.1, each of these conditions is equivalent to the others, offering different perspectives on the nature of Gröbner–Shirshov bases in commutative semirings.

Moreover, the study of the quotient semi-algebra DRig$[X]/I$ requires choosing a good representative. Hence, we can select $N(f|G)$ as the representative of $\bar{f}$, meaning $\bar{f} = N(f|G)$, because the normal form of f modulo I is unique. Thus, the following result.

Proposition 5.1 *Let* $I = \langle G \rangle$ *be an ideal of* DRig$[X]$ *and* $\bar{f}, \bar{g} \in$ DRig$[X]/I$. *Then, the following statements hold:*

(1) $\bar{f} = \bar{g}$ *if and only if* $N(f|G) = N(g|G)$.
(2) $\bar{f} = \overline{N(f|G)}$.

We present an algorithm designed to compute the Gröbner–Shirshov basis in the case of a monomial semiring DRig$[X]$ with coefficients in a D-A ring, using a well-founded order.

Algorithm 1 systematically computes the Gröbner–Shirshov basis for a given ideal in DRig$[X]$ using the given polynomial subset F. By iteratively applying $\mathcal{A}$-`polynomial` and $\mathcal{S}$-`polynomial` operations and performing `reduction` steps, Algorithm 1 `ComputeGröbnerShirshovBasis` transforms the set F into the `minimal Gröbner-Shirshov basis` G. The resulting basis G is representative of the reduced forms of polynomials in the given ideal, providing a valuable tool for algebraic computations in the given ring.

To illustrate the application of the proposed algorithm, we consider the following example in which we compute the `minimal Gröbner-Shirshov basis` for a given set of polynomials over a monomial semiring with coefficients in a D-A ring, with respect to well-founded order.

Example 5.1 Given $R = \mathbb{Z}_{12}[i]\mathrm{Rig}[X]$ where $i^2+1 = 0$ and $X = \{x, y\}$, compute a Gröbner–Shirshov basis of $F = \{f_1 = (5+3i)x^2y \circ x - y,\ f_2 = (3+2i)xy^2 \circ y - x\}$. We assume the length lexicographic order induced by $y > x$ and $<_{\mathbb{Z}_n[i]}$.

Since that $5 + 3i = 6 + 6i$, is a divisor of zero in $\mathbb{Z}_{12}[i]$, thus

$$\mathrm{Apol}(f_1) = \mathrm{ann}(\mathrm{LC}(f_1)) \cdot \mathrm{rest}(f_1) = (6 + 6i)y = f_3.$$

We have the new polynomial f_3 from f_1 and $F = \{f_1, f_2, f_3\}$.

Algorithm 1: COMPUTEGRÖBNERSHIRSHOVBASIS

Input: Ring $R = \text{DRig}[X]$, $X = \{x, y\}$;
polynomial set $F = \{f_i, f_j\}$ that generate ideal I;
A well-founded order $<$ on $\text{DRing}[X]$
Output: Gröbner-Shirshov basis G for the given ideal in $\text{DRig}[X]$

```
F_new ← F;
G ← {} ; // Initialize an empty set to store the Gröbner-Shirshov basis.
while F_new is not empty do
    for each polynomial f_i in F_new do
        if LC(f_i) has a non-trivial annihilator in D then
            Add Apol(f_i) to F_new ;                // Compute A-polynomials.
        end
    end
    for each pair (f_i, f_j) in F_new do
        Identify S−polynomial with respect to an ambiguity ω;
        Compute S(f_i, f_j)_ω;
        if S(f_i, f_j)_ω ≠ 0 then
            Add S(f_i, f_j)_ω to F_new ;            // Compute S−polynomials.
        end
    end
    for each polynomial f_i in F_new do
        Check if f_i is reducible modulo other polynomials in G;
        if reducible then
            Replace f_i with the reduced form ;     // Reduce polynomials.
        end
    end
    G ← F_new;
end
return G ;                                          // Final Gröbner-Shirshov basis.
```

$3 + 2i$ is not a zero divisor in $\mathbb{Z}_{12}[i]$, so no $\mathcal{A}$-polynomial is generated from f_2.

From polynomials f_1 and f_2, computation of $\mathcal{S}$-polynomial, starting firstly to look at the $\mathcal{S}$-polynomial with respect to ambiguity $\omega_0 = \text{LM}(f_1)y = x\text{LM}(f_2)$:

$$\mathcal{S}(f_1, f_2)_{\omega_0} = f_1 - qxf_2 = [(5+3i)x^2y \circ x - y]y - (-3-i)$$
$$x[(3+2i)xy^2 \circ y - x] = (-3-i)x^2 - y^2 = f_4$$

the new polynomial $f_4 = (-3-i)x^2 - y^2$ and the set F become $F = \{f_1, f_2, f_3, f_4\}$.

From the polynomials 1 and 3, computing of $\mathcal{S}$-polynomial, starting firstly to look the $\mathcal{S}$-polynomial with respect to ambiguity $\omega_1 = \text{LM}(f_1) = x^2\text{LM}(f_3) \circ x$:

$$\mathcal{S}(f_1, f_3)_{\omega_1} = qf_1 - x^2f_3 \circ x = 6[(5+3i)x^2y \circ x - y] - (6+6i)x^2y \circ x = 6y = f$$

where $q = \text{quot}(\text{LC}(f_3), \text{LC}(f_1)) = 6$ because $6 + 6i = 6(5 + 3i)$.

In fact, since that $f_3 = 6(5+3i)y = (5+3i)f$ is multiple of f, then f_3 is deleted and one fixed $f = f_{3'}$. Thus, we have the following polynomial $f_{3'} = 6y$ and the set F become $F = \{f_1, f_2, f_{3'}, f_4\}$.

Furthermore, as $\mathrm{LM}(f_{3'}) = y$ divide $\mathrm{LM}(f_1) = x^2y \circ x$, f_1 is reducible modulo $f_{3'}$, i.e., $f_1 \longrightarrow_{f_{3'}} f_{1'}$, which implies that

$$f_{1'} = s_1 f_1 - q f_{3'} s_2 \circ u = (5+3i)x^2y \circ x - y - 6x^2y \circ x = (-1+3i)x^2y \circ x - y,$$

where $5+3i = 6.(1)+(-1+3i)$ with $q = 1$ and $s_1 = x^2, s_2 = 1, u = x$. We delete the polynomial f_1, one obtains the following polynomial $f_{1'} = (-1+3i)x^2y\circ x - y$, and the set F becomes $F = \{f_{1'}, f_2, f_{3'}, f_4\}$.

From the polynomials $f_{1'}$ and $f_{3'}$, computing the $\mathcal{S}$-polynomial, starting by looking at the $\mathcal{S}$-polynomial with respect to ambiguity. $\omega_2 = \mathrm{LM}(f_{1'}) = x^2\mathrm{LM}(f_{3'}) \circ x$:

$$\begin{aligned}\mathcal{S}(f_{1'}, f_{3'})_{\omega_2} &= q f_{1'} - x^2 f_{3'} \circ x = (3-3i)[(-1+3i)x^2y \circ x - y] - 6x^2y \circ x \\ &= -(3-3i)y = f,\end{aligned}$$

where $6 = (-1+3i)(3-3i)$. Since that $\mathrm{LM}(f_{3'})$ divides, $\mathrm{LM}(f)$ then, f is reducible modulo $f_{3'}$, i.e., $f \longrightarrow_{f_{3'}} f'$, one has $f' = f - q s_1 f_{3'} s_2 \circ u = -(3-3i)y + 6y = (3+3i)y$, where $q = -1$ because $-3+3i = 6 \cdot (-1) + (-3-3i)$. In fact, since that polynomial $f_{3'} = 6y = i(-1+3i)(3+3i)y = i(-1+3i)f$ is multiple of f', one deletes $f_{3'}$. Whence the following polynomial $f' = f_{3'} = (3+3i)y$ and the set F become $F = \{f_{1'}, f_2, f_{3''}, f_4\}$.

From the polynomials $f_{1'}$ and $f_{3''}$, computing of $\mathcal{S}$-polynomial, starting firstly to look at the $\mathcal{S}$-polynomial with respect to ambiguity $\omega_3 = \mathrm{LM}(f_{1'}) = x^2\mathrm{LM}(f_{3''}) \circ x$:

$$\begin{aligned}\mathcal{S}(f_{1'}, f_{3''})_{\omega_3} &= q f_{1'} - x^2 f_{3''} \circ x = -3[(-1+3i)x^2y \circ x - y] + (3+3i)x^2y \circ x \\ &= 3y = f,\end{aligned}$$

where $3+3i = -3(-1+3i)$. Since that $f_{3''} = (3+3i)y = -3(-1+3i)y = (-1+3i)f$ is a multiple of f, one deletes the polynomial $f_{3''}$. Thus, we obtained the new polynomial $f_{3'''} = f = 3y$ and the set F become $F = \{f_{1'}, f_2, f_{3'''}, f_4\}$.

We have known also that $\mathrm{LM}(f_{3'''})$ divides $\mathrm{LM}(f_{1'})$ so $f_{1'} \longrightarrow_{f_{3'''}} f_{1''}$', i.e., $f_{1''} = f_{1'} - q s_1 f_{3'''} s_2 \circ u = -x^2y \circ x - y$, where $s_1 = x^2$, $s_2 = 1$, $u = x$, and $q = i$ because $-1+3i = 3.(i) - 1$. One deletes the rule $f_{1'}$, and one obtained the new following polynomial $f_{1''} = f = -x^2y \circ x - y$ and the set F become $F = \{f_{1''}, f_2, f_{3'''}, f_4\}$.

From the polynomials f_2 and $f_{3'''}$, computing of $\mathcal{S}$-polynomial, starting firstly to look at the $\mathcal{S}$-polynomial with respect to ambiguity $\omega_4 = \mathrm{LM}(f_2) = \mathrm{LM}(f_{3'''}) \circ xy^2$:

$$\mathcal{S}(f_2, f_{3'''})_{\omega_4} = f_2 - q f_{3''} \circ xy^2 = (3+2i)xy^2 \circ y - x - 3xy^2 \circ y = 2ixy^2 \circ y - x = f,$$

where $q = 1$ because $3 + 2i = -3.(1) + 2i$. Since that $\text{LM}(f_{3'''})$ divides, $\text{LM}(f)$ then, f is reducible modulo $f_{3'''}$, i.e., $f \longrightarrow_{f_{3'''}} f'$, one has $f' = f - q s_1 f_{3'''} s_2 \circ u = 2ixy^2 \circ y - 3ixy^2 \circ y = -ixy^2 \circ y - y$, where $q = 1$ because $2i = 3 \cdot (i) - i$, $s_1 = 1$, $s_2 = 1$, and $u = xy^2$. So one deletes the polynomial f. In fact, since that polynomial $f_2 = (3 + 2i)xy^2 - x = (2 - 3i)f'$ is a multiple of f', one deletes the polynomial f_2. Whence the new polynomial $f_{2'} = f_2 = -ixy^2 \circ y - x$ and F become $F = \{f_{1''}, f_{2'}, f_{3'''}, f_4\}$.

From the polynomials $f_{2'}$ and $f_{3'''}$, computing of $\mathcal{S}$-polynomial, starting firstly to look the $\mathcal{S}$-polynomial with respect to ambiguity $\omega_5 = \text{LM}(f_{2'}) = \text{LM}(f_{3'''}) \circ xy^2$:

$$\mathcal{S}(f_{2'}, f_{3'''})_{\omega_5} = qf_{2'} - f_{3''} \circ xy^2 = 3i[-ixy^2 \circ y - x] - 3xy^2 \circ y = -3ix = f,$$

where $q = 3i$ because $3 = -i.(3i)$. Whence the new polynomial $f_5 = f = -3ix$ and F become $F = \{f_{1''}, f_{2'}, f_{3'''}, f_4, f_5\}$.

From the polynomials $f_{2'}$ and f_5, computing of $\mathcal{S}$-polynomial, starting firstly to look at the $\mathcal{S}$-polynomial with respect to ambiguity $\omega_6 = \text{LM}(f_{2'}) = \text{LM}(f_5)y^2 \circ y$:

$$\mathcal{S}(f_{2'}, f_5)_{\omega_6} = qf_{2'} - f_5 y^2 \circ y = -3ixy^2 \circ y - 3x + 3ixy^2 \circ y = -3x = f,$$

where $q = 3$ because $-3i = -i.(3)$. Since that $f_5 = -3ix = if$ is a multiple of f, so we delete the f_5 and fixed the new polynomial $f_{5'} = f = -3x$ and F become $F = \{f_{1''}, f_{2'}, f_{3'''}, f_4, f_{5'}\}$. Furthermore, as $\text{LM}(f_{5'})$ divides $\text{LM}(f_4)$, f_4 is reducible modulo $f_{5'}$, i.e., $f_4 \longrightarrow_{f_{5'}} f_{4'}$, which implies that $f_{4'} = f_4 - q s_1 f_{5'} s_2 \circ u = (-3-i)x^2 - y^2 + 3x^2 = -ix^2 - y^2$, where $q = 1$ because $-3-i = -3.(1) - i$ and $s_1 = x$, $s_2 = 1$ $u = 0$. One deletes the polynomial f_4, and one obtains the new polynomial $f_{4'} = -ix^2 - y^2$ and the set F become $F = \{f_{1''}, f_{2'}, f_{3'''}, f_{4'}, f_{5'}\}$.

After the above steps, we get a `minimal Gröbner-Shirshov basis` of polynomials f_1 and f_2 over $(\mathbb{Z}_{12}[i])[x, y]$ with $i^2+1 = 0$, which consists of polynomials corresponding to $f_{1''}$, $f_{2'}$,$f_{3'''}$, $f_{4'}$, $f_{5'}$, i.e.,

$$F = \{f_{1''}, f_{2'}, f_{3'''}, f_{4'}, f_{5'}\}.$$

In Example 5.1, we demonstrated the computation of a `Gröbner-Shirshov basis` over the semi-algebra DRig$[X]$, where D is a ring of Gaussian integers modulo. Specifically, we considered the semi-algebra $R = \mathbb{Z}_{12}[i]\text{Rig}[X]$ with $i^2 + 1 = 0$ and the set of polynomials $F = \{f_1, f_2\}$, where $f_1 = (5 + 3i)x^2 y \circ x - y$ and $f_2 = (3 + 2i)xy^2 \circ y - x$. The computations were done with respect to a length lexicographic order induced by $y > x$ and $< \mathbb{Z}_n[i]$.

The `Gröbner-Shirshov basis` was obtained step by step, introducing new polynomials and eliminating redundant ones by computations of $\mathcal{S}$-polynomial and $\mathcal{A}$-polynomials. The final Gröbner–Shirshov basis for the ideal generated by F was determined to be $F = \{f_{1''}, f_{2'}, f_{3'''}, f_{4'}, f_{5'}\}$, where $f_{1''} = -x^2 y \circ x - y$, $f_{2'} = -ixy^2 \circ y - x$, $f_{3'''} = 3y$, $f_{4'} = -ix^2 - y^2$, and $f_{5'} = -3x$.

This example shows the effectiveness of the `Gröbner-Shirshov basis` calculation in handling polynomials over DRig[X] in the given algebraic setting. The resulting basis provides a concise representation of the ideal generated by the initial set of polynomials.

Syzygies Theorem

For $r > 1$, denote by $(e_1, \ldots, e_r)$ the standard basis for free R-module F $= R^r$.

Let $f_1, \ldots, f_r \in R$ and $<$ be an admissible order in Rig[X]. Our goal is to define Schreyer's ordering also called the ordering $>_1$ in F induced by $f_1, \ldots, f_r$ and $<$.

Definition 5.5 (Schreyer's Ordering) Given a monomial ordering $<$ on R and nonzero polynomials $f_1, \ldots, f_r \in R$, we define the Schreyer's ordering $>_1$ on F induced by $<$ and $f_1, \ldots, f_r$ as follows: Let $x^\alpha e_i$ and $x^\beta e_j$ be monomials in F. One says that

$$x^\alpha e_i >_1 x^\beta e_j \text{ if and only if:}$$

1. $x^\alpha \mathrm{LM}(f_i) > x^\beta \mathrm{LM}(f_j)$.
2. $x^\alpha \mathrm{LM}(f_i) = x^\beta \mathrm{LM}(f_j)$ and $i > j \ \forall\, \alpha, \beta \in \mathbb{N}^n$.

Example 5.2 Consider the polynomials $g_1 = 5xy^2 \circ y \circ x + xy \circ 1$, $g_2 = 2x^2y \circ y - 1 \circ xy$ in $R = \mathrm{DRig}[X]$ with $D = \mathbb{Z}_6$. Using length lexicographic ordering and the monomial ordering $>$, $\mathrm{LM}(g_1) = xy^2 \circ y \circ x$ and $\mathrm{LM}(g_2) = x^2y \circ y$. Let us compare the monomials xye_1 and $x^2y^2e_2$ with $>_1$ in R induced by g_1, g_2 and $>$:

1. $xy\mathrm{LM}(g_1) = x^2y^3 \circ xy^2xy^3 \circ x^2y$.
2. $x^2y^2\mathrm{LM}(g_2) = x^4y^3 \circ x^2y^3$.

Since $xy\mathrm{LM}(g_1) >_{\mathrm{Ln}} x^2y^2\mathrm{LM}(g_2)$, then $xy^2e_1 >_1 x^2y^2e_2$.

Remark 5.2 Consider a `Gröbner-Shirshov basis` $G = \{g_1, \ldots, g_r\}$ for the ideal I in the ring R. According to the Composition-Diamond Lemma for Gröbner–Shirshov structures (cf. Theorem 5.1), for any pair (i, j) with $1 \leq j < i \leq r$, the remainder obtained by the division of the $\mathcal{S}$-polynomial $\mathcal{S}(g_i, g_j)_\omega$ and the $\mathcal{A}$-polynomial Apol(p_i) by G is exactly zero. In such cases there exist elements $h_1, \ldots, h_s \in R$ which satisfy:

$$\mathcal{S}(g_i, g_j)_\omega = s_i g_i \circ u_i - q s_j g_j \circ u_j = \sum_{k=1}^{r} f_k^{ij} g_k \circ u_i + 0 \text{ and Apol}(g_i)$$

$$= \mathrm{ann}(\mathrm{LC}(g_i)).g_i = \sum_{k=1}^{r} h_k^i g_k + 0,$$

and this means that

$$\sum_{k=1}^{r} f_k^{ij} g_i \circ u_i - \frac{\mathrm{LC}(g_j)}{\mathrm{LC}(g_i)} s_i g_i \circ u_i + s_j g_j \circ u_j = 0 \text{ and } \sum_{k=1}^{r} h_k^i g_k - \mathrm{ann}(\mathrm{LC}(g_i)).g_i = 0,$$

where $\frac{\mathrm{LC}(g_i)}{\mathrm{LC}(g_j)} = \mathrm{quot}(\mathrm{LC}(g_i), \mathrm{LC}(g_j)) = q$. Thus, we have

$$f_1^{ij} g_1 \circ u_1 + \ldots + \ldots + g_j (f_j^{ij} + s_j) \circ u_j + g_i (f_i^{ij} - \frac{\mathrm{LC}(g_i)}{\mathrm{LC}(g_j)} s_j) \circ u_i + \ldots + f_r^{ij} g_r \circ u_r = 0$$

and

$$g_1 h_1 + \ldots + g_j h_j + \ldots + g_i (h_i - \mathrm{ann}(\mathrm{LC}(g_i))) + \ldots + g_r h_r = 0.$$

The following sets

$$G^{ij} = (f_1^{ij} \circ u_1, \ldots, (f_j^{ij} + s_j) \circ u_j, \ldots, (f_i^{ij} - \frac{\mathrm{LC}(g_j)}{\mathrm{LC}(g_i)} s_i) \circ u_i, \ldots, f_r^{ij} \circ u_r) \in R^r$$

and

$$K^i = (h_1, \ldots, (h_i - \mathrm{ann}(\mathrm{LC}(g_i))), \ldots, h_r) \in R^r$$

are syzygies of $g_1, \ldots, g_r$. Thus, the set of all syzygies obtained from our $\mathcal{S}$-polynomials and $\mathcal{A}$-polynomials is $T = \{G^{ij}, K^i \mid 1 \leq j < i \leq r\}$.

Therefore, by definition 5.2, $\mathcal{S}(g_i, g_j)_\omega = q s_i g_i \circ u_i - s_j g_j \circ u_j$ and $\mathrm{Apol}(g_i) = \mathrm{ann}(c).g_i$, respectively. This leads to $\frac{\mathrm{LC}(g_i)}{\mathrm{LC}(g_j)} s_i \mathrm{LT}(g_i) \circ u_i - s_j \mathrm{LT}(g_j) \circ u_j = 0$. That is,

$$s_i \mathrm{LM}(g_i) \circ u_i = s_j \mathrm{LM}(g_j) \circ u_j$$

observed that for all $i > j$, we have the following relation

$$\mathrm{LT}(f_k^{ij} g_k) \leq \mathrm{LT}(\mathcal{S}(g_i, g_j)) \leq s_j \mathrm{LT}(g_j) \circ u_j = \frac{\mathrm{LC}(g_i)}{\mathrm{LC}(g_j)} s_i \mathrm{LT}(g_i) \circ u_i.$$

Since that $s_i \mathrm{LM}(g_i) \circ u_i = s_j \mathrm{LM}(g_j) \circ u_j$, hence, for $i > j$, then by the Schreyer's ordering $<_1$, we have the leading term of G^{ij}

$$\mathrm{LT}(G^{ij}) = -\frac{\mathrm{LC}(g_j)}{\mathrm{LC}(g_i)} (s_i \circ u_i) e_i.$$

In the same way, for all i and each k, $\mathrm{LT}(h_k^i g_k) \leq \mathrm{LT}(\mathrm{Apol}(g_i)) \leq \mathrm{ann}(\mathrm{LC}(g_i))\mathrm{LT}(g_i)$. Thus, we have the leading term of K^i

$$\mathrm{LT}(K^i) = \mathrm{ann}(\mathrm{LC}(g_i))\mathrm{LT}(g_i)e_i.$$

The following Proposition 5.2 establishes a fundamental connection between Gröbner–Shirshov bases and the syzygy module of an ideal I in a ring R.

Proposition 5.2 (Division in Syzygy Module) *Let I be an ideal of R. Let $G = \{g_1, \ldots, g_r\}$ be a Gröbner–Shirshov basis for I with respect to the order $<$. Then the set $\mathrm{T} = \{T_1, \ldots, T_k\}$ is $\mathcal{S}$-reduced of G if for every vector polynomial $H \in \mathrm{syz}(G)$, then there exists a vector polynomial $Q \in T$ such that $\mathrm{LT}(Q)$ divides $\mathrm{LT}(H)$ w.r.t. Schreyer order $>_1$ induced by $<$ and $g_1, \ldots, g_r$.*

Proof Let $H = a.me+\mathrm{Rest}(H) \in G$ such that $H \longrightarrow^* 0$. We proceed by induction on H, by using Schreyer's order $>_1$.

Since $H \longrightarrow^* 0 \iff \exists k \geq 0$ such that $H \longrightarrow^k 0$:

(1) If $k = 0$, then $H = 0$. By definition 5.4, there exists $P \in \mathrm{T}$ such that $\mathrm{LT}(P)|\mathrm{LT}(H)$ w.r.t. Schreyer's ordering.
(2) Assumption of recurrence: Suppose that $Q \longrightarrow^* 0$ with $Q < H$:

- For $Q \longrightarrow^* 0$ with $Q < H$ according to Lemma 5.1, there exists $P \in T$ such that $\mathrm{LT}(P)|\mathrm{LT}(Q)$ w.r.t. Schreyer's ordering induced by $<$ and G.
- Further, since that relation $\longrightarrow^*$ is reflexive and transitive closure of relation $\longrightarrow$, we can consider $H \longrightarrow^+ 0$. By the rule satisfied by $\longrightarrow^+$, there exists $T_1 \in \mathrm{T}$ such that $\mathrm{LT}(T_1)|\mathrm{LT}(Q)$ w.r.t. the Schreyer's ordering, i.e., there exists $t = a.me$ in H such that $m = s_1.m_1e_1 \circ u_1$ where $m_1e_1 = \mathrm{LM}(T_1)$, $s_1 \in X^*$, $u_1 \in \mathrm{Rig}[X]$, and $a = q_1.\mathrm{LC}(T_1) + b_1$ with $b_1 = \mathrm{rem}(a, \mathrm{LC}(T_1)) < a = \mathrm{LC}(H)$ and $q_1 = \mathrm{quot}(a, \mathrm{LC}(T_1))$, set $\mathrm{LC}(T_1) = c_1$. By rule of the relation $\longrightarrow^*$, we can assume $H \longrightarrow^* H' \longrightarrow_{T_1} Q \longrightarrow^* 0$.

  If $b_1 \neq 0$, then $\mathrm{LT}(Q) = b_1.me$. Since $Q \longrightarrow^* 0$ with $Q < H$ by assumption of recurrence, there exists $T_2 \in T$ such that $\mathrm{LT}(T_2)|\mathrm{LT}(Q)$ w.r.t. Schreyer's ordering, i.e., $m = s_2.m_2e_2 \circ u_2$, where $m_2e_2 = \mathrm{LM}(T_2)$, $s_2 \in X^*$, $u_2 \in \mathrm{Rig}[X]$, and $a = q.\mathrm{LC}(T_2) + b_2$ with $b_2 = \mathrm{rem}(a, \mathrm{LC}(T_2)) < a = \mathrm{LC}(H)$ and $q = \mathrm{quot}(a, \mathrm{LC}(T_2))$, set $\mathrm{LC}(T_2) = c_2$. By the properties of ring D-A, we have $c_2 \leq b_1$. But since that, $m_1|m$ and $m_2|m$ so $s_1m_1e_1 \circ u_1 = me = s_2m_2e_2 \circ u_2$ (with $e = e_2$). We can so compute the $\mathcal{S}$-polynomial of T_1 and T_2 w.r.t. ω_1. Thus we have

$$\mathcal{S}(T_1, T_2)_{\omega_1} = s_1.T_1 \circ u_1 - q_2.s_2.T_2 \circ u_2,$$

  where $q_2 = \mathrm{quot}(c_1, c_2)$ and $c_1 = q_2.c_2 + b_2$ with $b_2 = \mathrm{rem}(c_1, c_2) < c_1$.

$$\mathcal{S}(T_1, T_2)_{\omega_1} = b_2.me + s_1.\mathrm{Rest}(T_1) \circ u - q_2.s_2.\mathrm{Rest}(T_2) \circ u_2,$$

  which leads to $\mathrm{LT}(\mathcal{S}(T_1, T_2)_{\omega}) = b_2.me = b_2.s_2m_2e_2 \circ u_2$, which means that

$$\begin{cases} \mathrm{LM}(\mathcal{S}(T_1, T_2)_{\omega_1}) = me = s_2 m_2 e_2 \circ u_2 \\ \mathrm{LC}(\mathcal{S}(T_1, T_2)_{\omega_1}) = b_2 < c_1 < a = \mathrm{LC}(H) \end{cases}.$$

If $b_2 \neq 0$, then $\mathcal{S}(T_1, T_2)_{\omega_1} < H$, and then by assumption of recurrence, there exists T_3 in T such that $\mathrm{LT}(T_3)$ divides $\mathrm{LT}(\mathcal{S}(T_1, T_2)_{\omega})$, i.e., $\mathrm{LM}(T_3) = m_3 e_3 | \mathrm{LM}(\mathcal{S}(T_1, T_2)_{\omega_1}) = me$ and $\mathrm{LC}(T_3) = c_3 | \mathrm{LC}(\mathcal{S}(T_1, T_2)_{\omega_1}) = b_2$. By the properties of the $\mathrm{D} - \mathrm{A}$ ring D, we have $c_3 < b_2$. We observe that $m_3 | m$ and $m_2 | m$ so $s_2.m_2 e_2 \circ u_2 = me = s_1 m_3 e_3 \circ u_3$ (here $e_3 = e = e_2$). We can compute the $\mathcal{S}$-polynomial of T_2 and T_3 w.r.t. ω_2. Thus, we have

$$\mathcal{S}(T_2, T_3)_{\omega_2} = s_2 T_2 \circ u_2 - q_3 s_3 T_3 \circ u_3,$$

where $q_3 = \mathrm{quot}(c_2, c_3)$ and $c_2 = q_3.c_3 + b_3$ with $b_3 = \mathrm{rem}(c_2, c_3) < c_2$

$$\mathcal{S}(T_2, T_3)_{\omega_2} = b_3.me + s_2 \mathrm{Rest}(T_2) \circ u_2 - q_3 s_3 \mathrm{Rest}(T_3) \circ u_3,$$

which leads to

$$\begin{cases} \mathrm{LM}(\mathcal{S}(T_2, T_3)_{\omega_2}) = me = s_1 m_3 e_3 \circ u_3 \\ \mathrm{LC}(\mathcal{S}(T_2, T_3)_{\omega_2}) = b_3 < c_2 < b_1 < a = \mathrm{LC}(H). \end{cases}$$

If $b_3 \neq 0$, then $\mathcal{S}(T_2, T_3)_{\omega_2} < H$. Since the ring D is Noetherian and that $\{b_i\}_{1 \leq i \leq N}$ is strictly decreasing, we can continue the above process until $b_{N+1} = 0$. Thus, we obtain a sequence of finite

$$a = b_0 > b_1 > b_2 > \cdots > b_N > b_{N+1} = 0$$

the corresponding polynomials in $T = \{T_1, T_2, \ldots, T_{N+1}\}$ with $T_i = c_i m_i e_i + \mathrm{Rest}(T_i)$, for all $1 \leq i \leq N+1$. According to Lemma 5.1 $c_{N+1} m_{k+1} e_{N+1} | a.me$, and for $1 \leq i \leq N+1$ that is to say $\mathrm{LT}(T_{N+1}) | \mathrm{LT}(H)$. It follows that T_{N+1} is the polynomial we were looking for in T.

In conclusion, Proposition 5.2 asserts that given a Gröbner–Shirshov basis G for I, a chosen set T is $\mathcal{S}$-reduced if it satisfies a crucial condition related to the leading terms of vector polynomials in the syzygy module. This condition ensures that, under Schreyer's order induced by the monomial order $<$ and the elements of the basis G, the chosen set T possesses a divisibility property crucial for studying the algebraic structure of the ideal. This proposition not only provides a theoretical framework for understanding the division properties within the syzygy module but also serves as a practical tool for computational algebraic tasks involving Gröbner–Shirshov bases and syzygies in ring theory.

The following algorithm shows that there exists a division for Syzygy module T in $\mathrm{DRig}[X]$.

Algorithm 2: DivisionSyzygyModule

Input: A Gröbner-Shirshov basis $G = \{g_1, \ldots, g_r\}$ for the ideal I in DRig$[X]$.
A set of vector polynomials $\mathrm{T} = \{T_1, \ldots, T_k\}$ in DRig$[X]$.
A well-founded order $<$ on DRig$[X]$ is used to compute the Gröbner-Shirshov basis.
Output: True if T is $\mathcal{S}$-reduced w.r.t. G. False otherwise.

for $k = 1, \ldots, s$ **do**
  Compute the $\mathcal{S}$ polynomials $\mathcal{S}(T)_\omega$ for all pairs of vector polynomials (T_i, T_j) in T with respect to ambiguity ω, using Schreyer's order $>_1$ and G as a basis:
  for *each pair* (T_i, T_j) *in* T **do**
    Compute $\mathcal{S}(T)_\omega$ using Schreyer's order $>_1$ and G.
  end
  Compute the normal form $N(H)$ for each vector polynomial H in syz(G), the syzygy module of G, using Schreyer's order $>_1$ and G as the basis:
  for *each vector polynomial H in syz(G)* **do**
    Compute $N(H)$ using Schreyer's order $>_1$ and G.
    Set found = False. **for** *each vector polynomial Q in* T **do**
      Compute LT(H) and LT(Q) using the Schreyer order $>_1$ induced by $<$ and G.
      if LT(Q) *divides* LT(H) *with respect to Schreyer's order* $>_1$ *and* G **then**
        Set found = True and break.
      end
    end
    if *found = False* **then**
      return *False* (T *is not $\mathcal{S}$-reduced).*
    end
  end
end
return *True* (T *is $\mathcal{S}$-reduced).*

Compute the $\mathcal{S}$-polynomials $\mathcal{S}(\mathrm{T})$ for all pairs of vector polynomials (Ti, Tj) in T, using Schreyer's order $>_1$ and G as a basis:

Now consider the vector polynomial H, written in the form $H = as.P \circ u$, where $P \in T$, $a \in$ D, $s \in X^*$, and $u \in$ Rig$[X]$. The following Proposition 5.3 establishes an important result concerning the weak representation of elements in the syzygy module syz(G) associated with a Gröbner–Shirshov basis G for an ideal I.

Proposition 5.3 (Weak Representation of Syzygy Module) *Let $G = \{g_1, \ldots, g_r\}$ be a Gröbner–Shirshov basis for I w.r.t. the ordering $<$. Assume that the set* $\mathrm{T} = \{T_1, \ldots, T_k\}$ *is $\mathcal{A}$-reduced of F. Then $H \in$* syz(G) $\subset F$ *has a weak standard representation w.r.t. the Schreyer's ordering $>_1$ induced by $<$ and $g_1, \ldots, g_r$.*

Proof Assume that set T is $\mathcal{A}$-reduced. If $H \in$ T, then Apol$(H) \longrightarrow^* 0 \iff \exists k \geq 0$ such that Apol$(H) \longrightarrow^k 0$. Reasoning by recurrence on f:

(1) If $H = 0$, it is obvious that by Remark 5.1 $H = 0$ is the weak standard representation of H w.r.t. T.
(2) If $H = a.sP \circ u$, let us show that H has a weak standard representation by recurrence by P using Schreyer's order $>_1$:

(a) Let P be nonzero minimal polynomial in T and assume that Apol(P) $\neq 0$. Since T is $\mathcal{A}$-reduced, then there exist some polynomial $Q \in$ T such that $\text{Apol}(P) = b_1 s_1 Q \circ u_1$ with $b_1 \in$ D, $s_1 \in X^*$, and $u_1 \in \text{Rig}[X]$. This means that $\text{LM}(Q) \leq \text{LM}(\text{Apol})(P) < \text{LM}(P)$, which leads to $\text{LM}(Q) < \text{LM}(P)$. Thus $Q < P$. This leads to a contradiction because Q is a polynomial in T. Hence, $\text{Apol}(P) = 0$:

i. If $\text{LC}(a.sP \circ u) = a.\text{LC}(P) \neq 0$, then H has a weak standard representation w.r.t. T.
ii. If $\text{LC}(H) = \text{LC}(a.sP \circ u) = a.\text{LC}(P) = 0$, then there exists $a_k \in D$ such that

$$H = a_k s \text{Apol}^k(P) \circ u \text{ and } a_k \text{LC}(\text{Apol}^k) \neq 0.$$

Since $\text{Apol}(p) = 0$, then $H = 0$. Hence, by Remark 5.1, H has a weak standard representation w.r.t. T.

(3) Assumption of recurrence: Assume that for all vector polynomials $L \in F$ such that vector polynomial $L = b.sQ \circ v$ where $Q \in F$ and $Q < H$, then L has a weak standard representation w.r.t. T.

1. Since the vector polynomial is equal, $H = a.sP \circ u$ with $P \in$ T. We discussed the following conditions:

(a) If $\text{LC}(H) = a\text{LC}(P) \neq 0$, then H has a weak standard representation w.r.t. T.
(b) If $\text{LC}(H) = a.\text{LC}(P) = 0$, then there exists $a_k \in$ D such that

$$H = a_k.s\text{Apol}^k(P) \circ u_1 \text{ and } a_k.\text{Apol}^k(P) \neq 0.$$

Since $P \in$ T so $\text{Apol}^k(P) \longrightarrow^* 0$. Thus, by Lemma 5.5, $\text{Apol}^k(P)$ has a weak standard representation w.r.t. T that is to say

$$\text{Apol}^k(P) = \sum_{i=1}^{n} b_i s_i P_i \circ u_i,$$

where $b_i \in$ D, $s_i \in X^*$, and $u_i \in \text{Rig}[X]$, $p_i \in \text{Rig}[X]$ and

$$\begin{cases} \text{LM}(b_i s_i P_i \circ u_i) \leq \text{LM}(\text{Apol}^k(P)) \\ \text{LC}(b_i s_i P_i \circ u_i) = b_i \text{LC}(\text{Apol}^k(P)) \neq 0. \end{cases}$$

Thus, we obtain

$$H = a_k.s(\sum_{i=1}^{n} b_i.s_i P_i \circ u_i) \circ u_1 = \sum_{i=1}^{n} a_k b_i s s_i P_i \circ s u_i \circ u,$$

and setting $v_i = su_i \circ u$, $z_i = s_1 s_i$, and $c_{ik} = a_k b_i$, we have

$$H = \sum_{i=1}^{n} c_{ik} z_i P_i \circ v_i.$$

Since $\mathrm{LM}(P_i) \leq \mathrm{LM}(c_{ijk} z_i P_i \circ v_i) \leq \mathrm{LM}(\mathrm{Apol}(P)) < \mathrm{LM}(P)$, so $\mathrm{LM}(P_i) < \mathrm{LM}(P)$, which implies $P_i < P$. Furthermore, by assumption of recurrence, each $c_{ik} z_i P_i \circ v_i$ has a standard representation w.r.t. T. While replacing each $c_{ik} z_i P_i \circ v_i$ by its standard representation, we obtain a weak standard representation of H.

Assuming that a specially chosen set T is $\mathcal{A}$-reduced with respect to the free module F, Proposition 5.3 asserts that every element H in syz(G) has a weak standard representation under Schreyer order $>_1$.

This result highlights the structured nature of syzygy modules and their connection to Gröbner–Shirshov bases, providing valuable insights into the algebraic properties of ideals in the context of computational algebra and ring theory. The notion of weak standard representation, as demonstrated in this chapter, serves as a powerful tool for understanding and characterizing elements within the syzygy module associated with a given Gröbner–Shirshov basis.

The following algorithm shows that there exists a weak standard representation for Syzygy module T in DRig$[X]$.

Using Propositions 5.2 and 5.3, we now give the following Theorem 5.2, which will be very useful for characterizing the existence of a Gröbner–Shirshov basis for Syz$(g_1, \ldots, g_r)$ with respect to Schreyer's ordering induced by the monomial order $<$ and G.

Theorem 5.2 establishes a fundamental equivalence between the properties of Gröbner–Shirshov bases and the syzygy module associated with an ideal I in a ring R.

Theorem 5.2 (Syzygy Theorem) *Let I be the ideal of R. Let $G = \{g_1, \ldots, g_r\}$ be a Gröbner–Shirshov basis for I w.r.t. the ordering $<$. Then the set* $\mathrm{T} = \{T_1, \ldots, T_k\}$ *forms a Gröbner–Shirshov basis for syz$(g_1, \ldots, g_r)$ w.r.t. the ordering $>_1$ induced by $<$ and $g_1, \ldots, g_r$ if and only if the subset G is $\mathcal{AS}$-reduced.*

This theorem 5.2 can be proved by using the same argument as in Propositions 5.2 and 5.3, respectively. We omit the details.

This theorem asserts that the set T, composed of carefully selected elements T_i, forms a Gröbner–Shirshov basis for the syzygy module syz$(g_1, \ldots, g_r)$ precisely when the Gröbner–Shirshov basis $G = \{g_1, \ldots, g_r\}$ is $\mathcal{AS}$-reduced. Theorem 5.2 highlights the intricate interplay between the algebraic structure of Gröbner–Shirshov bases and the syzygy module and provides a powerful criterion for determining when a chosen set becomes a Gröbner–Shirshov basis for the associated syzygy module. This theorem not only enhances our theoretical understanding of syzygies but also has practical implications for computational algebra and ring theory, making it a key result in the study of algebraic structures and ideals.

Algorithm 3: WeakRepresentationSyzygyModule

Input: A Gröbner-Shirshov basis $G = \{g_1, \dots, g_r\}$ for the ideal I in DRig[X].
A set of vector polynomials $\mathrm{T} = \{T_1, \dots, T_k\}$ that is $\mathcal{A}$-reduced with respect to G.
The vector polynomial $H = as.P \circ u$, where $P \in \mathrm{T}$, $a \in \mathrm{D}$, $s \in X^*$, and $u \in \mathrm{Rig}[X]$.
Output: True if H has a weak standard representation with respect to G. False otherwise.

for $k = 1, \dots, s$ **do**
  Compute the syzygy module syz(G) of the Gröbner-Shirshov basis G.
  Compute the $\mathcal{A}$-polynomial of all polynomials in T using the Schreyer's ordering $>_1$ and G:
  for *each nonzero vector polynomial* $H = a \cdot s \cdot P \circ u$ *in* T **do**
    For every nonzero minimal polynomial P in T, find a polynomial Q in T such that $\mathrm{Apol}(P) = b \cdot s \cdot Q \circ v$ with $Q \le P$.
    if *Q exists in* T **then**
      if $\mathrm{Apol}(P) = 0$ **then**
        return *True (H has a weak standard representation w.r.t* T*).*
      end
    end
    else if $\mathrm{LC}(H) = a \cdot \mathrm{LC}(P) \neq 0$ **then**
      return *True (H has a weak standard representation w.r.t* T*).*
    end
  end
end
if $\mathrm{LC}(H) = a \cdot \mathrm{LC}(P) = 0$ **then**
  Find $a_k \in \mathrm{D}$ such that $H = a_k \cdot s \cdot \mathrm{Apol}^k(P) \circ u$ and $a_k \cdot \mathrm{Apol}^k(P) \neq 0$.
  if $\mathrm{Apol}^k(P) \longrightarrow^* 0$ **then**
    if $\mathrm{Apol}(P) = 0$ **then**
      return *True ($H = 0$).*
    end
  end
end
return *True (H has a weak standard representation w.r.t* T*).*

The following algorithm computes a Gröbner–Shirshov basis for Syzygy module Syz(G) in DRig[X].

Example 5.3 Suppose we want to find a syzygy of the ring R for the ideal $I = \langle f_1 = (5+3i)x^2y \circ x - y, f_2 = (3+2i)xy^2 \circ y - x \rangle$ generated by f_1 and f_2 in R. We will fix the lexicographic ordering with $y > x$ as the well-founded order. In Example 5.1, it was demonstrated that the set G forms a `minimal Gröbner-Shirshov basis` for $I = \langle f_1, f_2 \rangle$ with respect to the well-founded order.

Let us arrange the polynomials in the set G with respect to the negative reverse lexicographic ordering d_s: $F = \{g_1 = f_{5'}, g_2 = f_{3''}, g_3 = f_{4'}, g_4 = f_{1''}, g_5 = f_{2'}\}$.

When computing the Gröbner–Shirshov basis, we observe that $\mathcal{S}(g_1, g_2)_{\omega_{12}} = \mathcal{S}(g_3, g_5)_{\omega_{35}} = 0$. Furthermore:

$$\mathcal{S}(g_1, g_3)_{\omega_{13}} = y^2 g_1 - 3x g_3 = -ix^2 g_1$$

Algorithm 4: ComputingGröbnerShirshovBasisSyzygy

Input: A set of vector polynomials $\mathrm{T} = \{T_1, \ldots, T_k\}$ in DRig$[X]$.
A `minimal Gröbner-Shirshov basis` $G = \{g_1, \ldots, g_r\}$ for the ideal I in DRig$[X]$.
A well-founded order $<$ on DRig$[X]$ is used to compute the Gröbner-Shirshov basis.
Output: True if T forms a Gröbner-Shirshov basis for syz$(g_1, ..., g_r)$. False otherwise.

```
for k = 1, ..., s do
    Compute the S-polynomials S(T) for all pairs of vector polynomials (T_i, T_j) in T with
    respect to a specific ambiguity ω, using Schreyer's ordering >_1 and G as the basis: for
    each pair (T_i, T_j) in T do
        Compute S(T_i, T_j)_ω using Schreyer's ordering >_1 and G.
    end
    Compute the normal forms N(S(T)) for each S-polynomial S(T) using Schreyer's
    ordering >_1 and G as the basis: for each S-polynomial S(T) do
        Compute N(S(T)) using Schreyer's ordering >_1 and G.
    end
    for each polynomial f in G do
        for each vector polynomial T in T do
            Compute the S-polynomial S(f, T)_ω using Schreyer's ordering >_1 and G as
            the basis:
            Compute N(S(f, T)_ω) using Schreyer's ordering >_1 and G.
            if any N(S(f, T)_ω) is nonzero then
                return False (T is not AS-reduced).
            end
        end
    end
end
return True (T forms a Gröbner-Shirshov basis for syz(g_1, ..., g_r)).
```

$$\mathcal{S}(g_2, g_3)_{\omega_{23}} = yg_2 + 3g_3 = ixg_1$$

$$\mathcal{S}(g_1, g_4)_{\omega_{14}} = (xy \circ 1)g_1 - 3g_4 = 3g_2$$

$$\mathcal{S}(g_2, g_4)_{\omega_{24}} = (x^2y \circ x)g_2 + 3yg_4 = -yg_2$$

$$\mathcal{S}(g_3, g_4)_{\omega_{34}} = (x^2y)g_3 - y^2g_4 = -ix^2(x^2y \circ x) + y^3 \longrightarrow_{g_4} -yg_3$$

$$\mathcal{S}(g_1, g_5)_{\omega_{15}} = (xy^2 \circ y)g_1 + 3ig_5 = ixg_1$$

$$\mathcal{S}(g_2, g_5)_{\omega_{35}} = (xy \circ 1)g_2 - 3ig_5 = -ig_1$$

$$\mathcal{S}(g_4, g_5)_{\omega_{45}} = yg_4 + ixg_5 = g_3.$$

Observe that the following relations: $G^{31} = (-(ix^2 + y^2), 0, 3x, 0, 0)$, $G^{32} = (ix, -y, -3, 0, 0)$, $G^{41} = (-(xy \circ 1), 1, 0, 3, 0)$, $G^{42} = (0, -(x^2y \circ x + y), 0, -3y, 0)$, $G^{43} = (0, 0, -(x^2y \circ x + y, y^2, 0)$, $G^{51} = (-(xy^2 \circ y - ix), 0, 0, 0, -3ix)$, $G^{52} = (-i, -(xy \circ 1), 0, 0, 3i)$, $G^{54} = (0, 0, 1,_{y}, -ix)$ are syzygies of g_1, g_2, g_3, g_4, g_5 w.r.t. Schreyer's order induced by well-founded order $<$. By the Syzygy theorem 5.2, the set

$$\mathrm{T} = \{G^{31}, G^{32}, G^{41}, G^{42}, G^{43}, G^{51}, G^{52}, G^{54}\}$$

forms a `minimal Gröbner-Shirshov basis` for syz$(g_1, g_2, g_3, g_4, g_5)$ w.r.t. Schreyer's order induced by $<$ and g_1, g_2, g_3, g_4, g_5.

Minimal Free Resolution in DRig*[X]*

Given an ideal $I = \langle G \rangle$ of $R = \mathbf{DRig}[X]$, this section presents a method for computing a minimal free resolution for I as an R module. The approach is to compute a minimum free resolution for I using syzygy theorems.

Let $I = \langle f_1, \ldots, f_r \rangle$ represent a two-sided ideal of R, and consider a minimal `Gröbner-Shirshov basis` $G = \{g_1, \ldots, g_r\}$ for I with respect to some monomial order $>$. We can then construct the following maps:

$$\mathrm{F} \xrightarrow{\mathrm{Im}(\phi_1)} I \longrightarrow 0$$

Since $\{g_1, \ldots, g_r\}$ form a minimal `Gröbner-Shirshov basis` for I, by the Syzygy Theorem 5.2), one can find a minimal Gröbner–Shirshov basis $\{T_1, \ldots, T_{r_1}\}$ for syz$(g_1, \ldots, g_r) \subset \mathrm{F}_1$.

Now consider the map $\alpha_1 : \mathrm{F}_1 \longrightarrow \ker(\phi_1) = \langle g_1, \ldots, g_r \rangle$. By the inclusion map i_1 we have the following commutative diagram:

$$\begin{array}{ccc} \ker(\phi_1) & \xrightarrow{i_1} & \mathrm{F} \\ \uparrow{\scriptstyle \alpha_1} & \nearrow{\scriptstyle \phi_2 = i_1 \circ \alpha_1} & \\ \mathrm{F}_1 & & \end{array}$$

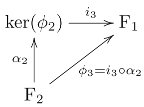
and adding these to the sequence, we obtain the following exact sequences:

$$\mathrm{F}_1 \xrightarrow{\phi_2} \mathrm{F} \xrightarrow{\phi_1} I \longrightarrow 0,$$

Now let us compute $\ker(\phi_2)$. Since $\{T_1, \ldots, T_{r_1}\}$ forms a `Gröbner-Shirshov basis` for syz$(g_1, \ldots, g_r)$ w.r.t. the Schreyer's ordering $<_1$ induced by $<$ and $\{g_1, \ldots, g_r\}$, then the remainder of the division's algorithm of $\mathcal{S}(T_i, T_j)_{1 \le j < i \le r_1}$ by $\{T_1, \ldots, T_{r_1}\}$ is zero. For each such $\mathcal{S}$-polynomials and $\mathcal{A}$-polynomials, one can collect the corresponding syzygy T_{ij} (playing the role of G^{ij} seen in Remark 5.2), and using Theorem 5.2, one sees that the set $\{T_{ij} \in \text{syz}(T_1, \ldots, T_{r_1}) / \ 1 \le j < i \le r_1\}$ forms a `Gröbner-Shirshov basis` for syz$(T_1, \ldots, T_{r_1}) =$ syz(syz$(g_1, \ldots, g_s)$) w.r.t. the Schreyer's ordering $<_2$ induced by $<_1$ and $\{T_1, \ldots, T_{r_1}\}$. Assume without loss of generalities that $\langle T_{ij} \in \text{syz}(T_1, \ldots, T_{r_1}) / \ i > j \rangle = \langle G_1, \ldots, G_{r_2} \rangle$ and the set $\{G_1, \ldots, G_{r_2}\}$ form a minimal `Gröbner-Shirshov basis` for syz(syz$(g_1, \ldots, g_s)$), and then:

$$\ker(\phi_2) = \mathrm{syz}(\mathrm{syz}(g_1, \dots, g_r)) = \langle G_1, \dots, G_{r_2} \rangle.$$

Consider the map $\alpha_2 : \mathrm{F}_{r_2} \longrightarrow \ker(\phi_2)$, and using the inclusion maps i_2 we get the following commutative diagrams:



Adding these to sequence precedent, we get the following exact sequences

$$\mathrm{F}_2 \xrightarrow{\mathrm{Im}(\phi_3)} \mathrm{F}_1 \xrightarrow{\mathrm{Im}(\phi_2)} \mathrm{F} \xrightarrow{\mathrm{Im}(\phi_1)} I \longrightarrow 0.$$

Continuing this way and so forth, this leads to a free resolution not necessarily finite. To obtain a finite and minimal free resolution, we use the technique used in [12] based on Hilbert's syzygies theorem. The degree reverses lexicographic ordering is the best choice for this method.

Theorem 5.3 *Let* $R = \mathrm{DRig}[X]$ *be a* $\mathrm{D} - \mathrm{A}$ *ring, and then any finitely generated* $R-$*module* F *has a minimal free resolution of length* Kdim R*,*

where Kdim R is the Krull dimension of the polynomial ring $R = \mathrm{DRig}[X]$.

If the ring D is a ring of integers modulo p,($\mathbb{Z}/p\mathbb{Z}$), the Krull dimension of D is 0. Therefore, the Krull dimension of R is equal to p. But, if the ring D is a Gauss integer ring modulo p ($\mathbb{Z}[i]/p\mathbb{Z} = \mathbb{Z}_p[i]$, where $i^2 + 1 = 0$), then we have two cases:

1. If the integer p is prime number, then the Krull dimension of D is 1.
2. If instead we take p to be a composite number, then Krull dimension of D will be either 1 or 2.

By applying the technique used in [12], we obtain the following short exact sequence:

$$0 \longrightarrow F_k \xrightarrow{\mathrm{Im}(\phi_{k+1})} F_{k-1} \xrightarrow{\mathrm{Im}(\phi_k)} \cdots F_2 \xrightarrow{\mathrm{Im}(\phi_3)} F_1 \xrightarrow{\mathrm{Im}(\phi_2)} F \xrightarrow{\mathrm{Im}(\phi_1)} I \longrightarrow 0$$

Theorem 5.4 (Existence of Minimal Free Resolution) *Let* $I = \langle f_1, \dots, f_r \rangle$ *be an ideal of* $R = \mathbf{DRig}[X]$*, and let* $G = \{g_1, \dots, g_r\}$ *be a minimal Gröbner–Shirshov basis for* I *with respect to a monomial order* $>$*. Suppose m and n are dimensioned such that* $m \leq n$ *and* I *have a minimal free resolution of length between m and n. Then, there exists a minimal free resolution for* I*:*

$$0 \to R^n \xrightarrow{\phi_n} \dots \xrightarrow{\phi_{k+1}} R^{n-k} \xrightarrow{\phi_k} \dots \xrightarrow{\phi_2} R^m \xrightarrow{\phi_1} I \to 0.$$

Algorithm 5: ComputeMinimalFreeResolution

Input: Ideal $I = \langle f_1, \ldots, f_r \rangle$ of $R = \mathbf{DRig}[X]$,
Monomial order $>$ on R,
A polynomial set $G = \{f_1, \ldots, f_r\}$ that generate ideal I;
Dimensions m and n such that $m \leq n$.
Output: Minimal free resolution for I

1 **Procedure:**

1. Initialize resolution: $F \leftarrow \emptyset$.
2. Set $k = 1$.
3. **while** $k \leq n$:
   (a) `ComputeGröbnerShirshovBasis(G)  // Compute Gröbner-Shirshov basis for I.`
   (b) `ComputingGröbnerShirshovBasisSyzygy(G) // Compute Gröbner-Shirshov basis for syzygy` T_k.
   (c) Set $\mathrm{T}_k = \{H_1, \ldots, H_s\}$, where each H_i is a syzygy polynomial.
   (d) Update resolution: $F \leftarrow F \cup \mathrm{T}_k$.
   (e) Increment k.
   (f) Arrange T_k `// Arranged according to degree reverses lexicographic ordering` d_s.
4. **Return** the sequence $0 \to R^n \xrightarrow{\phi_n} \ldots \xrightarrow{\phi_2} R^m \xrightarrow{\phi_1} I \to 0$, where each ϕ_i corresponds to the matrices formed by the polynomials in F.

Theorem 5.4 establishes the existence of a minimal free resolution and provides a constructive algorithm for its computation, combining the theoretical and computational aspects of minimal free resolution in the context of Gröbner–Shirshov bases.

Moreover, the resolution can be constructed algorithmically using Algorithm 5, which computes both Gröbner–Shirshov basis for ideal I and Gröbner–Shirshov basis for syzygies and forms a minimal free resolution.

To illustrate the application of the proposed algorithm `ComputeMinimalFree Resolution` 5, we consider the following example in which we compute the `minimal free resolution` for a given set of polynomials over a monomial semiring with coefficients in a D-A ring, with respect to well-founded order.

Example 5.4 Let us compute a `minimal free resolution` for I in $R = \mathrm{DRig}[x, y]$ where the ideal $I = \langle f_1 = (5+3i)x^2y \circ x - y,\ f_2 = (3+2i)xy^2 \circ y - x \rangle$ and $D = \mathbb{Z}_{12}[i]$ with $i^2 + 1 = 0$. Since $\dim(R)$ is between 3 and 4, we know in advance that I has a minimal free resolution of length between 3 and 4. In Example 5.1, the set $G = \{f_{1''}, f_{2'}, f_{3'''}, f_{4'}, f_{5'}\}$ is a minimal Gröbner–Shirshov basis for I in R. This leads to the exact following sequence:

$$R^5 \longrightarrow I \longrightarrow 0$$

and note that: $\ker(\phi_1) = \mathrm{syz}(g_1, \ldots, g_r)\ \ (*)$. This sequence can be represented as follows:

$$R^5 \xrightarrow{Im(\phi_1)} I \longrightarrow 0$$

In Example 5.3 we have seen that T $= \{G^{32}, G^{41}, G^{42}, G^{43}, G^{51}, G^{52}, G^{54}\}$ is the `minimal Gröbner-Shirshov basis` for $\mathrm{syz}(g_1, g_2, g_3, g_4, g_5)$ w.r.t. the Schreyer's ordering induced by $<$ and g_1, g_2, g_3, g_4, g_5. These lead to the exact sequences

$$R^8 \longrightarrow R^5 \xrightarrow{\begin{pmatrix} G^{31} \\ G^{32} \\ G^{41} \\ G^{42} \\ G^{43} \\ G^{51} \\ G^{52} \\ G^{54} \end{pmatrix}} I \longrightarrow 0$$

Observe the following leading monomials of the set T, $\mathrm{LM}(G^{31}) = y^2e_1$, $\mathrm{LM}(G^{32}) = ye_2$, $\mathrm{LM}(G^{41}) = (xy \circ 1)e_1$, $\mathrm{LM}(G^{42}) = (x^2y \circ x)e_2$, $\mathrm{LM}(G^{43}) = (x^2y \circ x)e_3$, $\mathrm{LM}(G^{51}) = (xy^2 \circ y)e_1$, $\mathrm{LM}(G^{52}) = (xy \circ 1)e_2$, $\mathrm{LM}(G^{54}) = ye_4$. Let us arrange these polynomials as follows: $H_1 = G^{32}, H_2 = G^{54}, H_3 = G^{31}, H_4 = G^{41}, H_5 = G^{52}, H_6 = G^{51}, H_7 = G^{42}, H_8 = G^{43}$. Since the set $T_1 = \{H_1, H_2, H_3, H_4, H_5, H_6, H_7, H_8\}$ is a `minimal Gröbner-Shirshov basis` for $\mathrm{syz}(g_1, g_2, g_3, g_4, g_5)$ w.r.t. $<_1$ induced by $>$ and g_1, g_2, g_3, g_4, g_5, then all the $\mathcal{S}$-vector polynomials and $\mathcal{A}$-vector polynomials reduce to 0 modulo T. Observe that:

$$\mathcal{S}(H_2, H_1)_{\omega_{21}} = \mathcal{S}(H_5, H_1)_{\omega_{51}} = \ldots = \mathcal{S}(H_5, H_4)_{\omega_{54}} = \mathcal{S}(H_6, H_2)_{\omega_{62}}$$
$$= \mathcal{S}(H_6, H_3)_{\omega_{63}} = S(H_6, H_5)_{\omega_{65}} = 0.$$

Furthermore, since $\mathcal{S}(H_8, H_1) = \ldots = \mathcal{S}(H_8, H_7) = 0$, hence,

$$\begin{aligned}
\mathcal{S}(H_4, H_3)_{\omega_{43}} &= y^2H - (xy \circ 1)H_3 = y^2(e_2 + 3e_4) - (x^2y \circ x) \\
&\quad (-ixe_1 + 3e_3) \longrightarrow_{H_1} yH_7; \\
\mathcal{S}(H_7, H_5)_{\omega_{75}} &= H_7 - xH_5 = 3(-ye_4 - ixe_5) + (-ye_2 + ixe_1) \longrightarrow_{H_2} H_1; \\
\mathcal{S}(H_7, H_1)_{\omega_{71}} &= yH_7 - (x^2y \circ x)H_1 = -y^2(ye_2 + 3ye_4) \\
&\quad -(x^2y \circ x)(ixe_1 - 3e_3) \longrightarrow_{H_7} (xy \circ 1)H_3, \\
\mathcal{S}(H_6, H_4)_{\omega_{64}} &= H_6 - yH_4 = (ixe_1 - ye_2) - 3(ixe_5 + ye_4) \longrightarrow_{H_1} 3H_2; \\
\mathcal{S}(H_6, H_3)_{\omega_{63}} &= yH_6 - (xy \circ 1)H_3 = xy(ie_1 - 3ie_5) - (xy \circ 1) \\
&\quad (-ix^2e_1 + 3xe_3) \longrightarrow_{H_5} (x^2y \circ x)H_1 \\
\mathcal{S}(H_5, H_1)_{\omega_{51}} &= yH_5 - (xy \circ 1)H_1 = y(-ie_1 + 3ie_5) - (xy \circ 1) \\
&\quad (ixe_1 - 3e_3) \longrightarrow_{H_1} yH_5
\end{aligned}$$

and thus, the following relations: $G'^{43} = (0, 0, -(xy \circ 1), y^2, 0, 0, -y, 0)$, $G'^{51} = (0, 0, 0, 0, y, 0, 0, 0)$, $G'^{75} = (1, 0, 0, 0, x, 0, 1, 0)$, $G'^{71} = (-(x^2y \circ x), 0, -(xy \circ 1), 0, x, 0, y, 0)$, $G'^{63} = (-(x^2y \circ x), 0, -(xy \circ 1), 0, 0, y, 0, 0)$, and $G'^{64} = (0, -3, 0, -y, 0, 1, 0, 0)$ are syzygies for $H_1, H_2, H_3, H_4, H_5, H_, H_8$, and by the Syzygy Theorem 5.2, the set from a minimal Gröbner–Shirshov basis for $\mathrm{syz}(H_1, H_2, H_3, H_4, H_5, H_6, H_7, H_8)$ w.r.t. the Schreyer's ordering $<_2$ induced by $<_1$ and $H_1, \ldots, H_8$. These lead to the exact sequences:

$$R^5 \longrightarrow R^8 \xrightarrow{\begin{pmatrix} G'^{43} \\ G'^{51} \\ G'^{63} \\ G'^{64} \\ G'^{71} \\ G'^{75} \end{pmatrix}} R^5 \xrightarrow{\begin{pmatrix} G^{31} \\ G^{32} \\ G^{41} \\ G^{42} \\ G^{43} \\ G^{51} \\ G^{52} \\ G^{54} \end{pmatrix}} I \longrightarrow 0$$

and set $T_1 = \{G'^{43}, G'^{51}, G'^{63}, G'^{64}, G'^{71}, G'^{75}\}$, where $\mathrm{LM}(G'^{43}) = (xy \circ 1)e_3$, $\mathrm{LM}(G'^{51}) = ye_5$, $\mathrm{LM}(G'^{63}) = (x^2y \circ x)e_1$, $\mathrm{LM}(G'^{64}) = ye_4$, $\mathrm{LM}(G'^{71}) = (x^2y \circ x)e_1$, $\mathrm{LM}(G'^{75}) = xe_5$. Let us arrange these polynomials as follows: $P_1 = G'^{75}, P_2 = G'^{64}, P_3 = G'^{51}, P_4 = G'^{43}$, $P_5 = G'^{71}$, $P_6 = G'^{61}$. Since the set $T_1 = \{P_1, P_2, P_3, P_4, P_5, P_6\}$ is a `minimal Gröbner-Shirshov basis` for $\mathrm{syz}(H_1, H_2, H_3, H_4, H_5, H_6, H_7, H_8)$ w.r.t. $<_1$ induced by $>$ and $H_1, \ldots, H_8$, these lead to

$$\mathcal{S}(P_4, P_1) = \ldots = \mathcal{S}(P_4, P_3) = \mathcal{S}(P_6, P_1) = \ldots = \mathcal{S}(P_6, P_4)$$

$$= \mathcal{S}(P_5, P_1) = \ldots = \mathcal{S}(P_5, P_4) = 0.$$

Further, we have the following $\mathcal{S}$-polynomials, $\mathcal{S}(P_3, P_1)_{\omega_{31}} = xP_3 + yP_1 = y(e_7 - e_6) \longrightarrow_{P_1} yP_3$ and $\mathcal{S}(P_6, P_5)_{\omega_{65}} = P_6 - yP_5 = y^2(e_6 - e_7) \longrightarrow_{P_1} -xP_3$. Thus, these relations $G''^{31} = (y, 0, 0, 0, 0, 0)$ and $G''^{65} = (0, 0, x, 0, x, 0, -y, 1)$ are also the syzygies for $P_1, \ldots, P_6$. Thus, the set $T_2 = \{G''^{31}, G''^{65}\}$ is a `minimal Gröbner-Shirshov basis` for $\mathrm{syz}(P_1, P_2, P_3, P_4, P_5, P_6)$ w.r.t. $<_2$ induced by $<_1$ and $P_1, \ldots, P_6$. These lead to the exact sequences:

$$R^2 \longrightarrow R^6 \xrightarrow{\begin{pmatrix} G''^{31} \\ G''^{65} \end{pmatrix}} R^8 \xrightarrow{\begin{pmatrix} G'^{43} \\ G'^{51} \\ G'^{63} \\ G'^{64} \\ G'^{71} \\ G'^{75} \end{pmatrix}} R^5 \xrightarrow{\begin{pmatrix} G^{31} \\ G^{32} \\ G^{41} \\ G^{42} \\ G^{43} \\ G^{51} \\ G^{52} \\ G^{54} \end{pmatrix}} I \longrightarrow 0$$

Let us arrange these polynomials as follows: $Q_1 = G'''^{65}$, $Q_2 = G'''^{64}$. Since $\mathcal{S}(Q_2, Q_1)_{\omega_{21}} = 0$, $\mathrm{syz}(Q_2, Q_1) = 0$. Thus, we have the following minimal free resolution

$$0 \longrightarrow R^2 \xrightarrow{(0)} R^6 \xrightarrow{\begin{pmatrix} G''^{31} \\ G''^{65} \end{pmatrix}} R^8 \xrightarrow{\begin{pmatrix} G'^{31} \\ G'^{32} \\ G'^{41} \\ G'^{42} \\ G'^{43} \\ G'^{51} \end{pmatrix}} R^5 \xrightarrow{\begin{pmatrix} G^{31} \\ G^{32} \\ G^{41} \\ G^{42} \\ G^{43} \\ G^{51} \\ G^{52} \\ G^{54} \end{pmatrix}} I \longrightarrow 0$$

In conclusion, we have successfully computed a `minimal free resolution` for the given ideal I in the ring $R = \mathrm{DRig}[x, y]$, where the ideal $I = \langle f_1 = (5 + 3i)x^2y \circ x - y, f_2 = (3 + 2i)xy^2 \circ y - x\rangle$.

Using a `minimal Gröbner-Shirshov basis`, we constructed the exact sequences representing the minimal free resolution. This `minimal free resolution` has been organized into a sequence of free modules and homomorphisms:

$$0 \longrightarrow R^2 \xrightarrow{\phi_4} R^6 \xrightarrow{\phi_3} R^8 \xrightarrow{\phi_2} R^5 \xrightarrow{\phi_1} I \longrightarrow 0$$

This example shows that the ideal I in the given semiring with coefficients in the Gaussian integer ring modulo 12 ($R = \mathrm{DRig}[x, y]$ with $\mathrm{D} = \mathbb{Z}_{12}[i]$ where $i^2 + 1 = 0$) has a minimal free resolution of length 4. The calculation included both `Gröbner-Shirshov basis` for ideal $I = \langle G\rangle$ and `Gröbner-Shirshov basis` for Syzygy, demonstrating their effectiveness in resolving ideals in this algebraic setting.

References

1. Berkesch C; Schreyer F,-O. *Syzygies, finite length modules, and random curves*. arXiv. 1403.0581 (2014).
2. Cox. D, J. O'Sheave. *Ideals, Varieties and Algorithms: An introduction to Computational Algebraic Geometry and Commutative*. Undergraduate Texts in Mathematics, Second edition, Springer, 1997;
3. David Eisenbud. *Geometry of Syzygies*. Second Course in Commutative Algebra and Algebraic geometry, Springer Verlag New York, 2005;
4. Deepak Kapur and Yongyang Cai: *An Algorithm for Computing a Gröbner Basis of a Polynomial Ideal over a Ring avec Zero Divisors*. Mathematics in Computer Science. pages 601–634. December 2009;
5. Erocal B., Motsak O., Schreyer F-O., and Steenpass A.: *Refined Algorithms to Compute Syzygies*. arxiv:1502.01654v2[math.AC] (2016).
6. E.Graham.E, Phillip G. *Syzygies*. London Mathematical Society, Springer Verlag New York, 1985;

7. Maroua.G, Henri.L, Sefan.N, I.Yengui *The syzygy theorem for Bézout rings.* J.Algebra. 301(2),447–458(2021);
8. J.P. Serre, *Local Algebra*, J.Springer, 1964;
9. Schreyer F-O. *A standard basis approach to syzygies of canonical curves*.J. Reine Angew. Math. 421 (1991), 83–123.
10. Schreyer F-O. *Die Berechnung von Syzygien mit dem verallgemeinerten weierstrasschen divisionssatz*. Diplomarbeit, Hamburg (1980).
11. S.Diop, M.wamba, D.Sow. *Gröbner-Shirshov basis monomials semiring over D-A rings* JP Journal of Algebra, Number Theory and Applications, 2023;
12. S.Diop.M.wamba.M.Bouesso, D. Sow, *On the computation of minimal free resolutions with integer coefficients* springer in Algebra, Codes and Cryptology.(2019);
13. Ralf.Fröberg, *An introduction to Gröbner bases*, University of Innsbruck, 1997;
14. Roger Wiegand, *What is ... a Syzygies* American Mathematical Society (2006). 456–457;

Chapter 6
Schur Complement and Inequalities of Eigenvalues on Block Hadamard Product

Abdoulaye Mbaye and Etienne Mbaye

Abstract The Schur complement theory is very important in many areas such as statistics, matrix analysis, numerical analysis, and control theory. It is a powerful tool to discuss many significant results. In this chapter, we establish two inequalities on the eigenvalues of Schur complement of the block Hadamard product, deduce one important corollary, and illustrate them in numerical examples.

Keywords Block matrices · Division algebra · Block matrices · Block Hadamard product · Eigenvalue

Introduction

In paper [2] the authors generalize Kronecker product for block matrices, mention some of its properties, and apply it to the study of a block Hadamard product of positive semidefinite matrices which was defined in [3]. Also generalizations of Schur's theorem were obtained in [4].

By using the definition and the properties of block Hadamard product, in [1] useful inequalities were obtained, and some numerical examples which confirm the theoretical analysis were given.

Let $\mathbb{N}$ be the set of positive integers. For $n,\ p,\ q \in \mathbb{N}$ let M_n be the linear space of $n \times n$ matrices with complex entries and let $\mathbf{M}_{p,q}(M_n)$ be the space of $p \times q$ block matrices $\mathbf{A} = (A_{\alpha\beta})_{\alpha=1,\ldots,p}^{\beta=1,\ldots,q}$, whose $\alpha,\ \beta$ entry belongs to M_n.

Let $\mathbf{A} = (A_{\alpha\beta})$, $\mathbf{B} = (B_{\alpha\beta}) \in \mathbf{M}_{pq}(M_n)$, where each block is an $n \times n$ matrix with complex entries. The block Hadamard product of $\mathbf{A}$ and $\mathbf{B}$ is $\mathbf{A}\square\mathbf{B} = (A_{\alpha\beta}B_{\alpha\beta})$, where $A_{\alpha\beta}B_{\alpha\beta}$ denotes the usual matrix product.

A. Mbaye (✉) · E. Mbaye
Université Cheikh Anta Diop, Dakar, Sénégal
e-mail: abdoulaye.mbaye@esp.sn

C. T. Gueye et al. (eds.), *Mathematics of Computer Science, Cybersecurity and Artificial Intelligence*, Springer Proceedings in Mathematics & Statistics 468,
https://doi.org/10.1007/978-3-031-66222-5_6

Let $\alpha \subset \{1, 2, \cdots, p\}$, $\beta \subset \{1, 2, \cdots, q\}$ be the index sets and $\alpha^c = \{1, 2, \cdots, p\} \setminus \alpha$, $\beta^c = \{1, 2, \cdots, q\} \setminus \beta$ be the complements of α and β, and their cardinalities are $|\alpha|$ and $|\beta|$.

Let $A(\alpha, \beta)$ denote the submatrix of A with the block rows indexed by α and block columns indexed by β. We will write $A(\alpha)$ for $A(\alpha, \alpha)$. If $|\alpha| = |\beta|$ and $A(\alpha, \beta)$ is nonsingular, the block of the Schur complement of $A(\alpha, \beta)$ is

$$A/A(\alpha, \beta) = A\left(\alpha^c, \beta^c\right) - A\left(\alpha^c, \beta\right)(A(\alpha, \beta))^{-1} A\left(\alpha, \beta^c\right).$$

$A/A(\alpha)$ is noted A/α.

If every $n \times n$ block of A commutes with every $n \times n$ block of B, we call these matrices block commuting, and we denote this by $A_{bc}B$.

Let A be a Hermitian matrix in M_n. We denote by $\lambda_{min}(A)$ and $\lambda_{max}(A)$ the smallest and largest eigenvalues, respectively, of the matrix A.

Theorem 6.1 *See Theorem 1 in [1].*
Let $\mathbf{A}$ *and* $\mathbf{B} \in \mathbf{M}_p$ *be positive semidefinite and* $\mathbf{A}_{bc}\mathbf{B}$*, and then*

$$(\mathbf{A}\Box\mathbf{B})/\alpha \geq \mathbf{A}/\alpha\Box\mathbf{B}/\alpha$$

.

Proposition 6.1 *See Proposition* 3.2 *in [2].*
If $\mathbf{A}$ *and* $\mathbf{B} \in \mathbf{M}_p$ *are positive semidefinite and* $\mathbf{A}_{bc}\mathbf{B}$*, then*

$$\lambda_{min}(\mathbf{A})\lambda_{min}(\mathbf{B}) \leq \lambda_{min}(\mathbf{A}\Box\mathbf{B})$$

$$\lambda_{max}(\mathbf{A}\Box\mathbf{B}) \leq \lambda_{max}(\mathbf{A})\lambda_{max}(\mathbf{B}).$$

Corollary 6.1 *Let* $\mathbf{A}$, $\mathbf{B} \in \mathbb{M}_p(M_n)$ *and* $\mathbf{A}_{bc}\mathbf{B}$.
If $\mathbf{A}$ *and* $\mathbf{B}$ *are positive semidefinite, then* $\mathbf{A}\Box\mathbf{B}$ *is positive semidefinite.*
If $\mathbf{A}$ *and* $\mathbf{B}$ *are positive definite, then* $\mathbf{A}\Box\mathbf{B}$ *is positive definite.*

Theory and Main Results

Theorem 6.2 *Let* $\mathbf{A}$ *and* $\mathbf{B} \in \mathbb{M}_p(M_n)$ *be positive definite matrices such that* $\mathbf{A}_{bc}\mathbf{B}$. *Then*

$$\lambda_{min}\left[(\mathbf{A}\Box\mathbf{B})/\alpha - \mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right] \leq \lambda_{min}\left[(\mathbf{A}\Box\mathbf{B})/\alpha\right] - \lambda_{min}\left(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right)$$

$$\lambda_{max}\left[(\mathbf{A}\Box\mathbf{B})/\alpha - \mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right] \geq \lambda_{max}\left[(\mathbf{A}\Box\mathbf{B})/\alpha\right] - \lambda_{max}\left(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right).$$

Proof Let $\mathbf{A}$ and $\mathbf{B} \in \mathbf{M}_p(M_n)$ be positive definite matrices such that $\mathbf{A}_{bc}\mathbf{B}$:

1. $\mathbf{A}/\alpha\Box\mathbf{B}/\alpha \geq \left[\lambda_{min}(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha)\right] I_n - \left[\lambda_{min}(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha)\right] I_n \geq -\mathbf{A}/\alpha\Box\mathbf{B}/\alpha$
   $(\mathbf{A}\Box\mathbf{B})/\alpha - \left[\lambda_{min}(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha)\right] I_n \geq (\mathbf{A}\Box\mathbf{B})/\alpha - \mathbf{A}/\alpha\Box\mathbf{B}/\alpha$
   $\lambda_{min}\left[(\mathbf{A}\Box\mathbf{B})/\alpha - \left[\lambda_{min}(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha)\right] I_n\right] \geq \lambda_{min}\left[(\mathbf{A}\Box\mathbf{B})/\alpha - \mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right]$
   $\lambda_{min}\left[(\mathbf{A}\Box\mathbf{B})/\alpha\right] - \lambda_{min}\left[(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha)\right] \geq \lambda_{min}\left[(\mathbf{A}\Box\mathbf{B})/\alpha - \mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right].$
2. $[\lambda_{max}(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha)] I_n \geq \mathbf{A}/\alpha\Box\mathbf{B}/\alpha$
   $-[\lambda_{max}(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha)] I_n \leq -\mathbf{A}/\alpha\Box\mathbf{B}/\alpha$
   $(\mathbf{A}\Box\mathbf{B})/\alpha - [\lambda_{max}(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha)] I_n \leq (\mathbf{A}\Box\mathbf{B})/\alpha - \mathbf{A}/\alpha\Box\mathbf{B}/\alpha$
   $\lambda_{max}\left[(\mathbf{A}\Box\mathbf{B})/\alpha - [\lambda_{max}(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha)] I_n\right] \leq \lambda_{max}\left[(\mathbf{A}\Box\mathbf{B})/\alpha - \mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right]$
   $\lambda_{max}\left[(\mathbf{A}\Box\mathbf{B})/\alpha\right] - \lambda_{max}\left[(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha)\right] \leq \lambda_{max}\left[(\mathbf{A}\Box\mathbf{B})/\alpha - \mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right].$

□

Corollary 6.2 *Let* $\mathbf{A}$ *and* $\mathbf{B} \in \mathbf{M}_p(M_n)$ *be positive definite matrices such that* $\mathbf{A}_{bc}\mathbf{B}$. *Then:*

1. $\lambda_{min}\left[(\mathbf{A}\Box\mathbf{B})/\alpha\right] \geq \lambda_{min}\left(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right).$
2. $\lambda_{max}\left[(\mathbf{A}\Box\mathbf{B})/\alpha\right] \geq \lambda_{max}\left(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right).$

Proof

1. $(\mathbf{A}\Box\mathbf{B})/\alpha \geq \mathbf{A}/\alpha\Box\mathbf{B}/\alpha$
   $(\mathbf{A}\Box\mathbf{B})/\alpha \geq \mathbf{A}/\alpha\Box\mathbf{B}/\alpha \geq \left[\lambda_{min}(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha)\right] I_n$
   $(\mathbf{A}\Box\mathbf{B})/\alpha \geq \left[\lambda_{min}(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha)\right] I_n$
   $\lambda_{min}\left[(\mathbf{A}\Box\mathbf{B})/\alpha\right] \geq \lambda_{min}\left[\left[\lambda_{min}(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha)\right] I_n\right] = \lambda_{min}(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha)$
   $\lambda_{min}\left[(\mathbf{A}\Box\mathbf{B})/\alpha\right] \geq \lambda_{min}(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha).$
2. $(\mathbf{A}\Box\mathbf{B})/\alpha \geq \mathbf{A}/\alpha\Box\mathbf{B}/\alpha$
   $\left[\lambda_{max}(\mathbf{A}\Box\mathbf{B})/\alpha\right] I_n \geq (\mathbf{A}\Box\mathbf{B})/\alpha \geq \mathbf{A}/\alpha\Box\mathbf{B}/\alpha$
   $\left[\lambda_{max}(\mathbf{A}\Box\mathbf{B})/\alpha\right] I_n \geq \mathbf{A}/\alpha\Box\mathbf{B}/\alpha$
   $\lambda_{max}[(\mathbf{A}\Box\mathbf{B})/\alpha] = \lambda_{max}\left[\left[\lambda_{max}(\mathbf{A}\Box\mathbf{B})/\alpha\right] I_n\right] \geq \lambda_{max}[\mathbf{A}/\alpha\Box\mathbf{B}/\alpha]$
   $\lambda_{max}[(\mathbf{A}\Box\mathbf{B})/\alpha] \geq \lambda_{max}[\mathbf{A}/\alpha\Box\mathbf{B}/\alpha].$

Theorem 6.3 *Let* $\mathbf{A}$ *and* $\mathbf{B} \in M_p(M_n)$ *be positive definite matrices such that* $\mathbf{A}_{bc}\mathbf{B}$. *Then*

$$\lambda_{min}\left[\left((\mathbf{A}\Box\mathbf{B})/\alpha\right)^{-1}\right] \leq \lambda_{min}\left[\left(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right)^{-1}\right] \leq \lambda_{min}\left[\left(\mathbf{A}/\alpha\right)^{-1}\right]\lambda_{min}\left[\left(\mathbf{B}/\alpha\right)^{-1}\right].$$

Proof Let $\mathbf{A}$ and $\mathbf{B} \in \mathrm{M}_p(M_n)$ be positive definite matrices such that $\mathbf{A}_{bc}\mathbf{B}$.

$(\mathbf{A}\Box\mathbf{B})/\alpha$ and $\mathbf{A}/\alpha\Box\mathbf{B}/\alpha$ are positive definite.

This means that $\lambda_{min}\left[(\mathbf{A}\Box\mathbf{B})/\alpha\right] > 0$ et $\lambda_{min}\left[\mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right] > 0$.

By using Corollary 6.1, we have $\lambda_{max}\left[(\mathbf{A}\Box\mathbf{B})/\alpha\right] \geq \lambda_{max}\left[\mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right]$.

Then $\left[\lambda_{max}\left[(\mathbf{A}\Box\mathbf{B})/\alpha\right]\right]^{-1} \leq \left[\lambda_{max}\left[\mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right]\right]^{-1}$.

$\Leftrightarrow \quad \lambda_{max}^{-1}\left[(\mathbf{A}\Box\mathbf{B})/\alpha\right] \leq \lambda_{max}^{-1}\left[\mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right]$

$\Leftrightarrow \quad \lambda_{min}\left[\left((\mathbf{A}\Box\mathbf{B})/\alpha\right)^{-1}\right] \leq \lambda_{min}\left[\left(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right)^{-1}\right] \quad (*)$

$\mathbf{A}/\alpha$, $\mathbf{B}/\alpha$ and $\mathbf{A}/\alpha\Box\mathbf{B}/\alpha$ are positive definite, then $\lambda_{min}(\mathbf{A}/\alpha) > 0$, $\lambda_{min}(\mathbf{B}/\alpha) > 0$ and $\lambda_{min}\left(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right) > 0$.

By using Proposition 6.1 $\lambda_{min}(\mathbf{A}/\alpha)\lambda_{min}(\mathbf{B}/\alpha) \leq \lambda_{min}\left(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right)$

$0 < \lambda_{min}(\mathbf{A}/\alpha)\lambda_{min}(\mathbf{B}/\alpha) \leq \lambda_{min}\left(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right)$

$\left[\lambda_{min}(\mathbf{A}/\alpha)\lambda_{min}(\mathbf{B}/\alpha)\right]^{-1} \geq \left[\lambda_{min}\left(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right)\right]^{-1}$

$\lambda_{min}^{-1}\left(\mathbf{A}/\alpha\right)\lambda_{min}^{-1}\left(\mathbf{B}/\alpha\right) \geq \lambda_{min}^{-1}\left(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right)$

$\lambda_{max}\left[\left(\mathbf{A}/\alpha\right)^{-1}\right]\lambda_{max}\left[\left(\mathbf{B}/\alpha\right)^{-1}\right] \geq \lambda_{max}\left[\left(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right)^{-1}\right] \quad (**)$.

After $(*)$ and $(**)$, we have: $\lambda_{min}\left[\left((\mathbf{A}\Box\mathbf{B})/\alpha\right)^{-1}\right] \leq \lambda_{min}\left[\left(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha\right)^{-1}\right] \leq \lambda_{min}\left[\left(\mathbf{A}/\alpha\right)^{-1}\right]\lambda_{min}\left[\left(\mathbf{B}/\alpha\right)^{-1}\right]$.

Example 6.1 See Example 1 dans [1].

Let A et B be $\mathbf{M}_2(\mathbb{M}_2)$ positive definite block matrices.

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} 5 & 2 & . & 1 & 0 \\ 2 & 3 & . & 0 & 1 \\ . & . & . & . & . \\ 1 & 0 & . & 5 & 2 \\ 0 & 1 & . & 2 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} 3 & 1 & . & 1 & 0 \\ 1 & 2 & . & 0 & 1 \\ . & . & . & . & . \\ 1 & 0 & . & 3 & 1 \\ 0 & 1 & . & 1 & 2 \end{bmatrix}$$

$$A\Box B = \begin{bmatrix} A_{11}B_{11} & A_{12}B_{12} \\ A_{21}B_{21} & A_{22}B_{22} \end{bmatrix} = \begin{bmatrix} 17 & 9 & . & 1 & 0 \\ 9 & 8 & . & 0 & 1 \\ . & . & . & . & . \\ 1 & 0 & . & 17 & 9 \\ 0 & 1 & . & 9 & 8 \end{bmatrix}$$

$A/\alpha = A/A_{11} = A_{22} - A_{21}A_{11}^{-1}A_{12} = \begin{bmatrix} 5 & 1 \\ 2 & 3 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 5 & 1 \\ 2 & 3 \end{bmatrix}^{-1}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 52/11 & 24/11 \\ 24/11 & 28/11 \end{bmatrix}$

$B/\alpha = B/B_{11} = B_{22} - B_{21}B_{11}^{-1}B_{12} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}^{-1}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 13/5 & 6/5 \\ 6/5 & 7/5 \end{bmatrix}$

$(A\Box B)/\alpha = (A\Box B)/A_{11}\Box B_{11} = A_{22}\Box B_{22} - A_{21}\Box B_{21}(A_{11}\Box B_{11})^{-1}A_{12}\Box B_{12}$

$(A\square B)/\alpha = \begin{bmatrix} 17 & 9 \\ 9 & 8 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 17 & 9 \\ 9 & 8 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 927/55 & 504/55 \\ 504/55 & 423/55 \end{bmatrix}.$

The characteristic polynomial of $(A\square B)/\alpha$: $p_{[(A\square B)/\alpha]}(\lambda) = 55\lambda^2 - 1350\lambda + 2511$.

The eigenvalues of $(A\square B)/\alpha$ are: $\lambda_1 = \frac{135}{11} - \frac{252\sqrt{5}}{55} = 2,0274703576$ and $\lambda_2 = \frac{135}{11} + \frac{252\sqrt{5}}{55} = 22,517984188$.

$A/\alpha\square B/\alpha = \begin{bmatrix} 52/11 & 24/11 \\ 24/11 & 28/11 \end{bmatrix} \square \begin{bmatrix} 13/5 & 6/5 \\ 6/5 & 7/5 \end{bmatrix} = \begin{bmatrix} 52/11 & 24/11 \\ 24/11 & 28/11 \end{bmatrix} \begin{bmatrix} 13/5 & 6/5 \\ 6/5 & 7/5 \end{bmatrix} = \begin{bmatrix} 164/11 & 96/11 \\ 96/11 & 68/11 \end{bmatrix}.$

The characteristic polynomial of $A/\alpha\square B/\alpha$: $p_{[A/\alpha\square B/\alpha]}(\lambda) = 11\lambda^2 - 232\lambda + 176$.

The eigenvalues of $A/\alpha\square B/\alpha$ are $\lambda_1 = \frac{116}{11} - \frac{48\sqrt{5}}{11} = 0,78806700727$ and $\lambda_2 = \frac{116}{11} + \frac{48\sqrt{5}}{11} = 20,302842084$.

$(A\square B)/\alpha - A/\alpha\square B/\alpha = \begin{bmatrix} 927/55 & 504/55 \\ 504/55 & 423/55 \end{bmatrix} - \begin{bmatrix} 164/11 & 96/11 \\ 96/11 & 68/11 \end{bmatrix} = \begin{bmatrix} 107/55 & 24/55 \\ 24/55 & 83/55 \end{bmatrix}.$

The characteristic polynomial of $(A\square B)/\alpha - A/\alpha\square B/\alpha$: $p_{[(A\square B)/\alpha - A/\alpha\square B/\alpha]}(\lambda) = 55\lambda^2 - 190\lambda + 151$.

The eigenvalues of $(A\square B)/\alpha - A/\alpha\square B/\alpha$ are $\lambda_1 = \frac{19}{11} - \frac{12\sqrt{5}}{55} = 1,2394033504$ et $\lambda_2 = \frac{19}{11} + \frac{12\sqrt{5}}{55} = 2,2151421042$.

The matrix $(A\square B)/\alpha - A/\alpha\square B/\alpha$ is Hermitian, and its eigenvalues are positive. Then $(A\square B)/\alpha - A/\alpha\square B/\alpha$ is positive definite.

$\lambda_{min}\left[(A\square B)/\alpha\right] - \lambda_{min}\left(A/\alpha\square B/\alpha\right) = \frac{135}{11} - \frac{252\sqrt{5}}{55} - \left(\frac{116}{11} - \frac{48\sqrt{5}}{11}\right) = \frac{19}{11} - \frac{12\sqrt{5}}{55}.$

Then we have $\lambda_{min}\left[(A\square B)/\alpha - A/\alpha\square B/\alpha\right] = \lambda_{min}\left[(A\square B)/\alpha\right] - \lambda_{min}\left(A/\alpha\square B/\alpha\right)$

$\lambda_{max}\left[(A\square B)/\alpha\right] - \lambda_{max}\left(A/\alpha\square B/\alpha\right) = \frac{135}{11} + \frac{252\sqrt{5}}{55} - \left(\frac{116}{11} + \frac{48\sqrt{5}}{11}\right) = \frac{19}{11} + \frac{12\sqrt{5}}{55}.$

Then we have $\lambda_{max}\left[(A\square B)/\alpha - A/\alpha\square B/\alpha\right] = \lambda_{max}\left[(A\square B)/\alpha\right] - \lambda_{max}\left(A/\alpha\square B/\alpha\right)$.

Conjecture

Let $\mathbf{A}$ and $\mathbf{B} \in \mathbb{M}_p(M_n)$ be positive definite matrices such that $\mathbf{A}_{bc}\mathbf{B}$. Then we conjecture that

$$\lambda_{min}\Big[(\mathbf{A}\Box\mathbf{B})/\alpha - \mathbf{A}/\alpha\Box\mathbf{B}/\alpha\Big] = \lambda_{min}\Big[(\mathbf{A}\Box\mathbf{B})/\alpha\Big] - \lambda_{min}\Big(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha\Big)$$

$$\lambda_{max}\Big[(\mathbf{A}\Box\mathbf{B})/\alpha - \mathbf{A}/\alpha\Box\mathbf{B}/\alpha\Big] = \lambda_{max}\Big[(\mathbf{A}\Box\mathbf{B})/\alpha\Big] - \lambda_{max}\Big(\mathbf{A}/\alpha\Box\mathbf{B}/\alpha\Big).$$

References

1. Mustafa Ozel and Ayca Ileri, Several Schur Complement inequalities on block Hadamard product, New Trends in Mathematical Sciences, 242–247(2017).
2. M. Gunter, L, Klotz, Schur's theorem for a block Hadamard product, Linear Algebra and its Application, 437(2012)948–956 .
3. R.A. Horn, R. Mathias, Y. Nakamura, Inequalities for unitarily invariant norms and bilinear matrix products, Linear and Multilinear Algebra 30 (1991) 303–314.
4. F. Zhang (editor). The Schur Complement and its Applications. Numerical Methods and Algorithms, Springer-Verlag, New York, 2005.

Chapter 7
A Perturbed Mann-Type Algorithm for Zeros of Maximal Monotone Mappings

Oumar Abdel Kader Aghrabatt, Aminata D. Diene, and Ngalla Djitte

Abstract Let E be a uniformly convex and uniformly smooth real Banach space and E^* its dual. Let $A : E \to E^*$ be a bounded maximal monotone mapping such that $A^{-1}(0) \neq \emptyset$. We first introduce the algorithm: For given $x_1 \in E$, let $\{x_n\}$ be generated by the formula: $x_{n+1} = x_n - \lambda_n J^{-1} A x_n - \lambda_n \theta_n (x_n - x_1)$, $n \geq 1$, where J is the normalized duality mapping from E into E^* and λ_n and θ_n are positive real numbers in (0, 1) satisfying suitable conditions. Next, we obtain the strong convergence of the sequence $\{x_n\}$ to the solution of the equation $Au = 0$ closest to the initial point x_1. Using this result, we deal with the convex minimization problem. Our results improve and unify most of the ones that have been proved in this direction for this important class of nonlinear mappings. Furthermore, our new technique of proof is of independent interest.

Keywords Maximal monotone mapping · Zeros · Convex minimization problem

Backround

Let H be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle_H$ and norm $\| \cdot \|_H$. An operator $A : H \to H$ with domain $D(A)$ is called *monotone* if for every $x, y \in D(A)$, the following inequality holds:

$$\langle x - y, Ax - Ay \rangle_H \geq 0, \tag{7.1}$$

O. A. K. Aghrabatt · N. Djitte (✉)
Gaston Berger University, Saint Louis, Sénégal
e-mail: abdkader@aims-senegal.org; ngalla.djitte@ugb.edu.sn

A. D. Diene
Université Cheikh Anta Diop, Ecole Polytechnique Supérieur, Dakar, Sénégal
e-mail: aminatadiop.diene@esp.sn

C. T. Gueye et al. (eds.), *Mathematics of Computer Science, Cybersecurity and Artificial Intelligence*, Springer Proceedings in Mathematics & Statistics 468,
https://doi.org/10.1007/978-3-031-66222-5_7

and it is called *strongly monotone* if there exists $k \in (0, 1)$ such that every $x, y \in D(A)$ satisfies

$$\langle x - y, Ax - Ay \rangle_H \geq k\|x - y\|_H^2. \tag{7.2}$$

Such operators have been studied extensively (see, e.g., Bruck Jr [5], Chidume [9], Martinet [31], Reich [39], Rockafellar [51]) because of their role in convex analysis, in certain partial differential equations, in nonlinear analysis and optimization theory.

The extension of the *monotonicity* definition to operators defined from a Banach space has been the starting point for the development of nonlinear functional analysis. The monotone maps constitute the most manageable class because of the very simple structure of the monotonicity condition. The monotone mappings appear in a rather wide variety of contexts since they can be found in many functional equations. Many of them also appear in calculus of variations as subdifferential of convex functions (see, e.g., Pascali and Sburian [37], p. 101, Rockafellar [51]).

The *first* extension involves mappings A from E to E^*. Here and in the sequel, $\langle \cdot, \cdot \rangle$ stands for the duality pairing between (a possible normed linear space) E and its dual E^*. Let E be a real normed space. A mapping $A : E \to E^*$ with domain $D(A)$ is called *monotone* if for each $x, y \in D(A)$, the following inequality holds:

$$\left\langle x - y, Ax - Ay \right\rangle \geq 0, \tag{7.3}$$

and it is called *strongly monotone* if there exists $k \in (0, 1)$ such that for each $x, y \in D(A)$, the following inequality holds:

$$\langle x - y, Ax - Ay \rangle \geq k\|x - y\|^2. \tag{7.4}$$

The *second* extension of the notion of monotonicity to real normed spaces involves mappings A from E *into itself*. Let E be a real normed space. For $q > 1$, define the multivalued map $J_q : E \to 2^{E^*}$ by

$$J_q(x) := \left\{u^* \in E^* : \langle x, u^* \rangle = \|x\|.\|u^*\|,\ \|u^*\| = \|x\|^{q-1}\right\}$$

The map J_q is called the generalized duality map on E. If $q = 2$, J_2 is called *normalized duality map* and is denoted by J. In a real Hilbert space H, J is the identity map on H. It is easy to see from the definition that

$$J_q(x) = \|x\|^{q-2}J(x), \quad and \quad \langle x, j_q(x) \rangle = \|x\|^q,\ \forall\, j_q(x) \in J_q(x).$$

A mapping $A : E \to E$ with domain $D(A)$ is called *accretive* if for all $x, y \in D(A)$, the following inequality is satisfied:

$$\|x - y\| \leq \|x - y + s(Ax - Ay)\| \ \ \forall s > 0. \tag{7.5}$$

It is called *m-accretive* if, in addition, the graph of A is not properly contained in the graph of any other accretive operator $A' : E \to E$. It is well known that A is m-accretive if and only if A is accretive and $R(I + tA) = E$ for all $t > 0$. If E is a real Hilbert space, accretive mappings are called *monotone* mappings and m-accretive mappings are called *maximal monotone mappings*.

Such operators have been studied extensively (see, e.g., [20–22, 40, 51]) because of their role in convex analysis, in certain partial differential equations, in nonlinear analysis and optimization theory.

As a consequence of a result of Kato [28], it follows that A is accretive if and only if for each $x, y \in D(A)$, there exists $j(x - y) \in J(x - y)$ such that

$$\langle Ax - Ay, j(x - y)\rangle \geq 0. \tag{7.6}$$

Finally, A is called *strongly accretive* if there exists $k \in (0, 1)$ such that for each $x, y \in D(A)$, there exists $j(x - y) \in J(x - y)$ such that

$$\langle Ax - Ay, j(x - y)\rangle \geq k\|x - y\|^2. \tag{7.7}$$

In a Hilbert space, the normalized duality map is the identity map. Hence, in Hilbert spaces, *monotonicity* and *accretivity* coincide.

For accretive-type operator A, solutions of the equation $Au = 0$, in many cases, represent *equilibrium state* of some dynamical system (see, e.g., [9], p.116).

For approximating a solution of $Au = 0$ (assuming existence), where $A : E \to E$ is of accretive type, Browder [4] defined an operator $T : E \to E$ by $T := I - A$, where I is the identity map on E. He called such an operator *pseudo-contractive*. One can observe that zeros of A correspond to fixed points of T. For Lipschitz strongly pseudo-contractive maps, Chidume [13] proved the following theorem.

Theorem C1 (Chidume, [13]) *Let $E = L_p$, $2 \leq p < \infty$, and $K \subset E$ be nonempty closed convex and bounded. Let $T : K \to K$ be a strongly pseudo-contractive and Lipschitz map. For arbitrary $x_0 \in K$, let a sequence $\{x_n\}$ be defined iteratively by $x_{n+1} = (1 - \lambda_n)x_n + \lambda_n T x_n$, $n \geq 0$, where $\{\lambda_n\} \subset (0, 1)$ satisfies the following conditions: (i) $\sum_{n=1}^{\infty} \lambda_n = \infty$, (ii) $\sum_{n=1}^{\infty} \lambda_n^2 < \infty$. Then, $\{x_n\}$ converges strongly to the unique fixed point of T.*

By setting $T := I - A$ in Theorem C1, the following theorem for approximating a solution of $Au = 0$ where A is a strongly accretive and bounded operator can be proved.

Theorem C2 *Let $E = L_p$, $2 \leq p < \infty$. Let $A : E \to E$ be a strongly accretive and bounded map. Assume $A^{-1}(0) \neq \emptyset$. For arbitrary $x_0 \in K$, let a sequence $\{x_n\}$ be defined iteratively by $x_{n+1} = x_n - \lambda_n A x_n$, $n \geq 0$, where $\{\lambda_n\} \subset (0, 1)$ satisfies the following conditions: (i) $\sum_{n=1}^{\infty} \lambda_n = \infty$, (ii) $\sum_{n=1}^{\infty} \lambda_n^2 < \infty$. Then, $\{x_n\}$ converges strongly to the unique solution of $Au = 0$.*

The main tool used in the proof of Theorem C1 is an inequality of Bynum [6]. This theorem signaled the return to extensive research efforts on inequalities in

Banach spaces and their applications to iterative methods for solutions of nonlinear equations. Consequently, Theorem C1 has been generalized and extended in various directions, leading to flourishing areas of research, for the past 30 years or so, for numerous authors (see, e.g., Censor and Reich [8], Chidume [13], Chidume [11, 12], Chidume and Ali [10], Chidume et al. [14], Chidume and Chidume [15, 16], Chidume and Osilike [19], Chidume and Djitte [17, 18], Deng [23], Zhou and Jia [60], Liu [30], Qihou [38], Reich [41–43], Reich and Sabach [44, 45], Weng [47], Xiao [50], Xu [52, 56, 58], Berinde et al. [3], Moudafi [33–35], Moudafi and Thera [36], Xu and Roach [54], Xu et al. [55], Zhu [61], and a host of other authors). Recent monographs emanating from these researches include those by Berinde [2], Chidume [9], Goebel and Reich [24], and William and Shahzad [49].

Unfortunately, the success achieved in using geometric properties developed from the mid-1980s to early 1990s in approximating zeros of *accretive-type mappings* has not carried over to approximating zeros of *monotone-type operators* in general Banach spaces.

Attempts have been made to overcome this difficulty by introducing the inverse of the normalized duality mapping in the recursion formulas for approximating zeros of monotone-type mappings.

Let E be a normed linear space. A monotone mapping $A : E \to 2^{E^*}$ is said to be *maximal* if its graph $G(A) = \{(x, y) : y \in Ax\}$ is not properly contained in the graph of any other monotone mapping. We know that if A is maximal monotone, then the zero set of A, $A^{-1}(0) := \{x \in E : 0 \in Ax\}$, is closed and convex. It is also known (see, e.g., Kohshada and Takahashi [29] for more details) that if E is reflexive, strictly convex, and smooth, then a monotone mapping A from E into E^* is maximal if and only if $R(J + \lambda A) = E^*$ for each $\lambda > 0$.

A function $F : E \to (-\infty, +\infty]$ is said to be *proper* if the set $\{x \in \mathbb{R} : F(x) \in \mathbb{R}\}$ is nonempty. A proper function $F : E \to (-\infty, +\infty]$ is said to be convex if for all $x, y \in E$ and $\alpha \in [0, 1]$ the following holds

$$F\Big(\alpha x + (1 - \alpha)y\Big) \leq \alpha F(x) + (1 - \alpha)F(y).$$

Also F is said to be *lower semicontinuous* if the set $\{x \in \mathbb{R} : F(x) \leq r\}$ is closed in E for all $r \in \mathbb{R}$. For the proper lower semicontinuous function $F : E \to (-\infty, +\infty]$, Rockafellar [51] proved that the *subdifferential* mapping, $\partial f : E \to 2^{E^*}$ of f defined by

$$\partial f(x) = \Big\{x^* \in E^* : f(y) - f(x) \geq \langle y - x, x^*\rangle \ \ \forall\, y \in E\Big\},$$

is maximal monotone.

Let E be a Banach space and $A : E \to 2^{E*}$ be a maximal monotone mapping. Then we consider the problem of finding a point $u \in E$ such that $0 \in Au$. Such a problem is connected with the *convex minimization problem*. In fact, if $F : E \to (-\infty, +\infty]$ is a proper lower semicontinuous convex function, then we have that $0 \in Au$ if and only if $F(u) = \min\limits_{x \in E} F(x)$.

A well-known method for solving the inclusion $0 \in Au$ in Hilbert space H is the *Proximal Point Algorithm*:

$$x_1 \in H, \quad x_{n+1} = J_{r_n} x_n, \; n \geq 1, \tag{7.8}$$

where $\{r_n\} \in (0, \infty)$ and $J_r = (I + rA)^{-1}$, $r > 0$. This algorithm was first introduced by Martinet [31]. In 1976, Rockafellar [51] proved that if $\liminf r_n > 0$ and $A^{-1}(0) \neq \emptyset$, then the sequence $\{x_n\}$ defined by (7.8) converges weakly to an element of $A^{-1}(0)$. Later, many researchers have studied the convergence of the sequence defined by (7.8) in a Hilbert space (see, for instance, Güler [25], Solodov and Svaiter [46], Kamimura and Takahashi [26], Lehdili and Moudafi [46]). In particular, Kamimura and Takahashi [48] obtained the following strong convergence:

Theorem KT *Let H be a real Hilbert space and $A : H \to 2^H$ be a maximal monotone mapping. For $u \in H$, let $\{x_n\}$ be a sequence defined by*

$$x_{n+1} = \alpha_n u + (1 - \alpha) J_{r_n} x_n, \; n \geq 1,$$

where $\{\alpha_n\} \in (0, 1)$ and $\{r_n\} \in (0, \infty)$ satisfy: $\lim_{n\to\infty} \alpha_n = 0$, $\sum \alpha_n = \infty$, *and* $\lim_{n\to\infty} r_n = \infty$. *If $A^{-1}(0) \neq \emptyset$, then the sequence $\{x_n\}$ converges strongly to Pu, where P is the metric projection of H onto $A^{-1}(0)$.*

In the case of Banach spaces, for finding zeros point of a maximal monotone mappings by using the Proximal Point Algorithm, Kohshada and Takahashi [29] introduced the following iterative sequence for a monotone mapping $A : E \to 2^{E^*}$:

$$x_1 = u \in E, \; x_{n+1} = J^{-1}\Big(\alpha_n Ju + (1 - \alpha) J J_{r_n} x_n\Big), \; n \geq 1, \tag{7.9}$$

where $J_{r_n} := (J + r_n A)^{-1}$, and J the duality mapping from E into E^*, $\{\alpha_n\} \in (0, 1)$ and $\{r_n\} \in (0, \infty)$ satisfy $\lim_{n\to\infty} \alpha_n = 0$, $\sum \alpha_n = \infty$ and $\lim_{n\to\infty} r_n = \infty$. They proved that if E is smooth and uniformly convex and A maximal monotone with $A^{-1}(0) \neq \emptyset$, then the sequence $\{x_n\}$ converges strongly to an element of $A^{-1}(0)$. This result extends Theorem KT to Banach spaces. However, the algorithm requires the computation of $(J + r_n A)^{-1} x_n$ at each step of the process, which makes its implementation difficult for applications.

Following the work of Kohshada and Takahashi [29], in [59], Zegeye introduced an iterative scheme for approximating zeros of maximal monotone mappings defined in uniformly smooth and 2-uniformly convex real Banach spaces. In fact, he proved the following theorem.

Theorem Z (Zegeye [59]) *Let E be a uniformly smooth and 2-uniformly convex real Banach space with dual E^*. Let $A : E \to E^*$ be a Lipschitz continuous monotone mapping with constant $L > 0$ and $A^{-1}(0) \neq \emptyset$. For given $u, x_1 \in E$, let $\{x_n\}$ be generated by the algorithm*

$$x_{n+1} = J^{-1}\Big(\beta_n Ju + (1-\beta_n)(Jx_n - \alpha_n Ax_n)\Big),\ \ n \geq \ 1,$$

where J is the normalized duality mapping from E into E^* and $\{\alpha_n\}$ and $\{\beta_n\}$ are real sequences in $(0, 1)$ satisfying (i) $\lim_{n\to\infty} \beta_n = 0$, (ii) $\sum \beta_n = \infty$, and (iii) $\alpha_n = o(\beta_n)$. Suppose that $B_{min} \cap \left(AJ^{-1}\right)^{-1}(0) \neq \emptyset$. Then $\{x_n\}$ converges strongly to $x^* \in A^{-1}(0)$ and that $R(Ju) = Jx^* \in \left(AJ^{-1}\right)^{-1}(0)$, where R is a sunny generalized nonexpansive retraction of E^* onto $\left(AJ^{-1}\right)^{-1}(0)$.

Remark 1 In Theorem Z, the author imposed the condition $B_{\min}\cap\left(AJ^{-1}\right)^{-1}(0) \neq \emptyset$. This condition is not easy to check because the set $B_{\min}$ and $\left(AJ^{-1}\right)^{-1}(0)$ are not known precisely.

Remark 2 It is well known that if E is a reflexive real Banach space and $A : E \to E^*$ is monotone and continuous such that $D(A) = E$, then A is maximal monotone. Therefore, the class of *bounded maximal monotone* mappings constitutes a superclass of that of *Lipschitz monotone* mappings, used in Theorem Z.

In a recent work, Mendy et al. introduced and studied a new iterative scheme for approximating zeros of bounded maximal monotone mappings. In fact, they consider the following iterative algorithm defined as follows: For given $x_1 \in E$, let $\{x_n\}$ be generated by the formula:

$$x_{n+1} = J^{-1}[Jx_n - \lambda_n Ax_n - \lambda_n\theta_n(Jx_n - Jx_1)],\ n \geq 1, \tag{7.10}$$

where J is the normalized duality mapping from E into E^* and λ_n and θ_n are positive real numbers in $(0, 1)$ satisfying suitable conditions. Then they obtained the following strong convergence result.

Theorem MA (Mendy et al. [32]) *For $q > 1$, let E be a 2-uniformly convex and q-uniformly smooth real Banach space and $A : E \to E^*$ be a bounded maximal monotone mapping such that $A^{-1}(0) \neq \emptyset$. Then, there exists $\gamma_0 > 0$ such that if $\lambda_n < \gamma_0\theta_n$ for all $n \geq 1$, the sequence $\{x_n\}$ given by (7.10) converges strongly to some $x^* \in A^{-1}(0)$.*

Recently, Sene et al. in [53] introduced a new Krasnoselskii-type algorithm for approximating zeros of Lipschitz and strongly mappings in some Banach spaces. They proved the following results.

Theorem SA1 (Sene et. al. [53]) *Let E be a 2-uniformly smooth Banach space. Let $A : E \to E^*$ be a L-Lipschitz and k-strongly monotone mapping with $A^{-1}(0) \neq \emptyset$ and such that*

$$L^2(d_2 - 1) < k^2.$$

For given $x_1 \in E$, let $\{x_n\}$ be a sequence defined as follows:

$$x_{n+1} = x_n - \lambda J^{-1} A x_n,\ n \geq 1, \tag{7.11}$$

where $\lambda \in (\alpha_1, \alpha_2)$ *with* $\alpha_1 = \frac{k}{L^2}$ *and* $\alpha_2 = \frac{k+\sqrt{k^2 - L^2(d_2-1)}}{L^2}$. *Then, the sequence* $\{x_n\}$ *converges strongly to* x^**, the unique solution of* $Au = 0$.

Theorem SA2 (Sene et al. [53]) *Let* $E = L_p$, $1 < p \leq 2$ *and let* $A : E \to E^*$ *be an* L*-Lipschitz and* k*-strongly monotone mapping with* $A^{-1}(0) \neq \emptyset$ *and such that*

$$L^2(d_2 - 1) < k^2.$$

Assume that $2 - \frac{k^2}{L^2} < p \leq 2$. *For arbitrary* $x_1 \in E$, *let* $\{x_n\}$ *be a sequence defined as follows:*

$$x_{n+1} = x_n - \lambda J^{-1}(A x_n),\ n \geq 1, \tag{7.12}$$

where $\lambda \in (\beta_1, \beta_2)$ *with* $\beta_1 = \frac{k}{L^2}$ *and* $\beta_2 = \frac{k+\sqrt{k^2 - L^2(2-p)}}{L^2}$.

Then, the sequence $\{x_n\}$ *converges strongly to* x^**, the unique solution of* $Au = 0$.

Following the work of Sene et al. in [53], Adoum et al. in [1] proposed a Mann-type algorithm for approximating zeros of bounded strongly monotone mappings in certain Banach spaces. They obtained the following result.

Theorem AA (Adoum et al. [1]) *For* $1 < p \leq 2$, *let* E *be a uniformly smooth and* p*-uniformly convex real Banach space with dual space* E^*. *Let* $A : E \to E^*$ *be a bounded and strongly monotone map. Assume that* $J^{-1}A$ *is strongly accretive. For arbitrary* $x_1 \in E$, *let* $\{x_n\}$ *be a sequence defined iteratively as follows:*

$$x_{n+1} = x_n - \lambda_n J^{-1} A x_n,\ n \geq 1, \tag{7.13}$$

where $\lambda_n \in (0, 1)$ *is a decreasing real sequence satisfying* (i) $\lim \lambda_n = 0$ *and* $(ii) \sum \lambda_n = \infty$ *and* $\sum \lambda_n^2 < \infty$. *Then, there exists* $\delta > 0$ *such that, if* $\lambda_n \leq \delta$, *the sequence* $\{x_n\}$ *converges strongly to* $x^* \in E$, *the unique solution of* $Au = 0$.

In this chapter, we introduce a *new iterative algorithm* to approximate zeros of *bounded maximal monotone mappings* defined in some real Banach spaces. The algorithm proposed here is simpler than the one introduced by Mendy et al. in [32] and does not require the computation of resolvent in the process and no continuity assumption is made. The results obtained in this work extend and unify those obtained by Kohshada and Takahashi [29], Zegeye [59], Sene et al. [53], Adoum et. al. [1], and most of the results that have been proved in this direction for this important class of nonlinear mappings. Then, we apply our results to the convex minimization problem. Finally, our method of proof is of independent interest.

Preliminaries

Let E be a normed linear space. E is said to be smooth if

$$\lim_{t \to 0} \frac{\|x + ty\| - \|x\|}{t} \tag{7.14}$$

exists for each $x, y \in S_E$. (Here $S_E := \{x \in E : ||x|| = 1\}$ is the unit sphere of E.) E is said to be uniformly smooth if it is smooth and the limit is attained uniformly for each $x, y \in S_E$, and E is Fréchet differentiable if it is smooth and the limit is attained uniformly for $y \in S_E$.

Let E be a real normed linear space of dimension ≥ 2. The *modulus of smoothness* of E , ρ_E, is defined by

$$\rho_E(\tau) := \sup \left\{ \frac{\|x + y\| + \|x - y\|}{2} - 1 : \|x\| = 1, \|y\| = \tau \right\}; \quad \tau > 0.$$

A normed linear space E is called *uniformly smooth* if

$$\lim_{\tau \to 0} \frac{\rho_E(\tau)}{\tau} = 0.$$

If there exist a constant $c > 0$ and a real number $q > 1$ such that $\rho_E(\tau) \leq c\tau^q$, then E is said to be *q-uniformly smooth*.

A normed linear space E is said to be strictly convex if:

$$\|x\| = \|y\| = 1, \ x \neq y \ \Rightarrow \ \left\| \frac{x + y}{2} \right\| < 1.$$

The modulus of convexity of E is the function $\delta_E : (0, 2] \to [0, 1]$ defined by

$$\delta_E(\epsilon) := \inf \left\{ 1 - \frac{1}{2}\|x + y\| \ : \|x\| = \|y\| = 1, \ \|x - y\| \geq \epsilon \right\}.$$

E is uniformly convex if and only if $\delta_E(\epsilon) > 0$ for every $\epsilon \in (0, 2]$. For $p > 1$, E is said to be p-uniformly convex if there exists a constant $c > 0$ such that $\delta_E(\epsilon) \geq c\epsilon^p$ for all $\epsilon \in (0, 2]$. Observe that every p-uniformly convex space is uniformly convex.

Typical examples of such spaces are the L_p, ℓ_p, and W_p^m spaces for $1 < p < \infty$, where

$$L_p \ (or \ l_p) \ or \ W_p^m \text{ is } \begin{cases} 2-\text{uniformly smooth and } p-\text{uniformly convex if } 2 \leq p < \infty; \\ 2-\text{uniformly convex and } p-\text{uniformly smooth if } 1 < p < 2. \end{cases}$$

It is well known that E is smooth if and only if J is single valued. Moreover, if E is a reflexive smooth and strictly convex Banach space, then J^{-1} is single valued, one-to-one, surjective, and it is the duality mapping from E^* into E.

Remark 3 Note also that a duality mapping exists in each Banach space. From [27], this mapping is known precisely in l_p, L_P, $W^{m,p}$ spaces, $1 < p < \infty$, and is given in such spaces by:

(i) $l_p \ : \ Jx = \|x\|_{l_p}^{2-p} y \in l_q, \ x = (x_1, x_2, \cdots, x_n, \cdots), \ y = (x_1|x_1|^{p-2}, x_2|x_2|^{p-2}, \cdots, x_n|x_n|^{p-2}, \cdots)$

(ii) $L_p : \ Ju = \|u\|_{L_p}^{2-p}|u|^{p-2}u \in L_q$

(iii) $W^{m,p} : \ Ju = \|u\|_{W^{m,p}}^{2-p} \sum_{|\alpha \leq m|}(-1)^{|\alpha|} D^{\alpha}\left(|D^{\alpha}u|^{p-2}D^{\alpha}u\right) \in W^{-m,q}$

where $1 < q < \infty$ is such that $1/p + 1/q = 1$.

In the sequel, we shall need the following results.

Theorem 1 (H. K. Xu [56])
Let $q > 1$ and E be a real Banach space. Then the following are equivalent:

(i) *E is q-uniformly smooth .*
(ii) *There exists a constant $d_q > 0$ such that for all $x, y \in E$,*

$$||x + y||^q \leq ||x||^q + q\langle y, j_q(x)\rangle + d_q||y||^q. \tag{7.15}$$

Lemma 1 (Chidume et. al [14]) *Let E be a real normed linear space and $q > 1$. Then, the following inequality holds:*

$$||x+y||^q \leq ||x||^q + q\langle y, j_q(x+y)\rangle \ \ \forall\, j_q(x+y) \in J_q(x+y), \ \ \forall\, x, y \in E. \tag{7.16}$$

Lemma 2 (Xu [57]) *Let $\{\rho_n\}$ be a sequence of nonnegative real numbers satisfying the following inequality:*

$$\rho_{n+1} \leq (1 - \alpha_n)\rho_n + \alpha_n\sigma_n + \gamma_n, \tag{7.17}$$

where $\{\alpha_n\}$, $\{\sigma_n\}$ and $\{\gamma_n\}$ are real sequences satisfying: (i) $\{\alpha_n\} \subset (0, 1)$, $\sum \alpha_n = \infty$.
(ii) $\limsup_{n\to\infty} \sigma_n \leq 0$ and (iii) $\gamma_n \geq 0$, $\sum \gamma_n < \infty$. Then, the sequence (ρ_n) converges to zero as $n \to \infty$.

Main Results

For $q > 1$, let E be a q-uniformly smooth and strictly convex real Banach space with norm $\|\cdot\|$ and dual space E^*. For $A : E \to E^*$ a mapping, let the sequence $\{x_n\}$ be generated iteratively from $x_1 \in E$ by

$$x_{n+1} = x_n - \lambda_n J^{-1} A x_n - \lambda_n \theta_n (x_n - x_1),\ n \geq 1, \tag{7.18}$$

where J is the normalized duality mapping from E into E^* and $\{\lambda_n\}$, $\{\theta_n\}$ are real sequences in (0, 1) satisfying, here and elsewhere, the following conditions:

$$(i)\ \lim_{n\to\infty} \theta_n = 0;\ (ii)\ \sum_{n=1}^{\infty} \lambda_n \theta_n = \infty, \quad \lambda_n^{q-1} = o(\theta_n);\ (iii)$$

$$\limsup_{n\to\infty} \frac{\left(\frac{\theta_{n-1}}{\theta_n} - 1\right)}{\lambda_n \theta_n} \leq 0,\ \sum_{n=1}^{\infty} \lambda_n^q < \infty.$$

Remark 4 Real sequences that satisfy conditions *(i)–(iii)* are $\lambda_n = (n+1)^{-a}$ and $\theta_n = (n+1)^{-b}$, $n \geq 1$ with $0 < b < (q-1)a$, $\frac{1}{q} < a < 1$ and $a + b < 1$.

In fact, (i), (ii), and the second part of (iii) are easy to check. For the first part of condition (iii), using the fact that $(1+x)^s \leq 1 + sx$, for $x > -1$ and $0 < s < 1$, we have

$$0 \leq \frac{\left(\frac{\theta_{n-1}}{\theta_n} - 1\right)}{\lambda_n \theta_n} = \left[\left(1 + \frac{1}{n}\right)^b - 1\right] \cdot (n+1)^{a+b}$$

$$\leq b \cdot \frac{(n+1)^{a+b}}{n} = b \cdot \frac{n+1}{n} \cdot \frac{1}{(n+1)^{1-(a+b)}} \to 0 \text{ as } n \to \infty.$$

The following result will be useful.

Theorem 2 (Reich, [7]) *Let E be a uniformly smooth real Banach space, and let $A : E \to 2^E$ be m-accretive with $D(A) = E$. Let $J_t x := (I + tA)^{-1} x$, $t > 0$ be the resolvent of A, and assume that $A^{-1}(0)$ is nonempty. Then for each $x \in E$, $\lim_{t\to\infty} J_t x$ exists and belongs to $A^{-1}(0)$.*

The following is a consequence of Theorem 2.

Lemma 3 *Let E be a uniformly convex and uniformly smooth real Banach space, $x_1 \in E$, and let $A : E \to E^*$ be a maximal monotone mapping such that $A^{-1}(0) \neq \emptyset$. Assume that $J^{-1}A : E \to E$ is accretive. Then, there exists a sequence $\{y_n\}$ in E such that:*

$$\theta_n (y_n - x_1) + J^{-1} A y_n = 0,\ \forall n \geq 1; \tag{7.19}$$

$$y_n \to y^* \textit{ with } y^* \in A^{-1}(0), \tag{7.20}$$

where J is the normalized duality mapping from E into E^*.

Proof Since E uniformly convex and uniformly smooth, then the duality mapping J from E into E^* is single valued, onto, and one-to-one, and its inverse J^{-1} exists and is the duality mapping of E^*. Let $x_1 \in E$ and set $y_n := (I + t_n J^{-1} A)^{-1} x_1$, where $t_n = (\theta_n)^{-1}\ \ \forall\, n \geq 1$. Then we have:

$$\theta_n(y_n - x_1) + J^{-1}Ay_n = 0,\ \ n \geq 1, \tag{7.21}$$

and from Theorem 2, it follows that

$$y_n \to y^* \ \text{ with } \ y^* \in (J^{-1}A)^{-1}(0), \tag{7.22}$$

which implies that $Ay^* = 0$.

We now prove the following theorem.

Theorem 3 *For $q > 1$, let E be a q-uniformly smooth and strictly convex real Banach space with norm $\|\cdot\|$ and dual space E^*. Let $A : E \to E^*$ be a bounded mapping such that $A^{-1}(0) \neq \emptyset$. Assume that $J^{-1}A : E \to E$ is m-accretive. Then, there exists $\gamma_0 > 0$ such that if $\lambda_n < \gamma_0\theta_n$ for all $n \geq 1$, the sequence $\{x_n\}$ given by (7.18) converges strongly to some $x^* \in A^{-1}(0)$.*

Proof Let $x^* \in E$ be a solution of the equation $Ax = 0$. There exists $r > 0$ sufficiently large such that $x_1 \in B(x^*, \frac{r}{2})$. Define $B = \bar{B}(x^*, r)$. Since A is bounded, it follows that $J^{-1}A(B)$ is bounded. Consequently,

$$M_0 := \sup\Big\{\|J^{-1}Ax + \theta(x - x_1)\|^q : \ x \in B\,, 0 < \theta \leq 1\Big\} + 1 < \infty.$$

Set

$$M_q := d_q M_0;\ \ \gamma := \Big(\frac{2^q - 1}{2^q M_q}\Big)r^q,$$

where d_q denotes the constant appearing in Theorem 1.

Now, assume that $\lambda_n^{q-1} \leq \gamma\theta_n$ for all $n \geq 1$.

Step 1. *We prove that the sequence $\{x_n\}$ is bounded*. In fact, we prove that $x_n \in B$ for all $n \geq 1$. The proof is by induction. By construction, $x_1 \in B$. Suppose that $x_n \in B$ for some $n \geq 1$. We prove that $x_{n+1} \in B$.

Using the recursion formula (7.18) and (ii) of Theorem 1, we have

$$\begin{aligned}\|x_{n+1} - x^*\|^q &= \|x_n - x^* - \lambda_n(J^{-1}Ax_n + \theta_n(x_n - x_1))\|^q \\ &\leq \|x_n - x^*\|^q - q\lambda_n\langle J^{-1}Ax_n + \theta_n(x_n - x_1), j_q(x_n - x^*)\rangle \\ &\quad + d_q\lambda_n^q\|J^{-1}Ax_n + \theta_n(x_n - x_1)\|^q \\ &\leq \|x_n - x^*\|^q - q\lambda_n\langle J^{-1}Ax_n + \theta_n(x_n - x_1), j_q(x_n - x^*)\rangle + \lambda_n^q M_q.\end{aligned}$$

Using the fact that $J^{-1}A$ is accretive, we obtain

$$\langle J^{-1}Ax_n + \theta_n(x_n - x_1), j_q(x_n - x^*)\rangle \geq \theta_n\|x_n - x^*\|^q + \theta_n\langle x^* - x_1, j_q(x_n - x^*)\rangle.$$

Therefore, we have the following estimates:

$$\begin{aligned}\|x_{n+1}-x^*\|^q &\le (1-q\lambda_n\theta_n)\|x_n-x^*\|^q - q\lambda_n\theta_n\langle x^*-x_1, j_q(x_n-x^*)\rangle + \lambda_n^q M_q\\ &\le (1-q\lambda_n\theta_n)\|x_n-x^*\|^q + q\lambda_n\theta_n\|x^*-x_1\|\|x_n-x^*\|^{q-1} + \lambda_n^q M_q\\ &\le (1-q\lambda_n\theta_n)\|x_n-x^*\|^q + q\lambda_n\theta_n\Big(\frac{1}{q}\|x^*-x_1\|^q + \frac{1}{q'}\|x_n-x^*\|^q\Big)\\ &\quad +\lambda_n^q M_q,\end{aligned}$$

with $1/q + 1/q' = 1$. Thus,

$$\|x_{n+1}-x^*\|^q \le (1-\lambda_n\theta_n)\|x_n-x^*\|^q + \lambda_n\theta_n\|x^*-x_1\|^q + \lambda_n^q M_q.$$

So, using the induction assumption, the fact that $x_1 \in B(x^*, r/2)$ and the condition $\lambda_n^{q-1} \le \gamma\theta_n$, we obtain

$$\|x_{n+1}-x^*\|^q \le r^q.$$

Therefore $x_{n+1} \in B$. Thus by induction, $x_n \in B$ for all $n \ge 1$.

Step 2. *We prove that* $\|x_{n+1}-y_n\| \to 0$ as $n \to 0$. From step 1, we have $\{x_n\} \subset B$. Since $\{y_n\}$ is bounded (being a convergent sequence) and A is bounded on B, there exists some positive constant M such that:

$$\begin{aligned}\|x_{n+1}-y_n\|^q &= \|x_n-y_n-\lambda_n(J^{-1}Ax_n+\theta_n(x_n-x_1))\|^q\\ &\le \|x_n-y_n\|^q - q\lambda_n\langle J^{-1}Ax_n+\theta_n(x_n-x_1), j_q(x_n-y_n)\rangle\\ &\quad + d_q\lambda_n^q\|J^{-1}Ax_n+\theta_n(x_n-x_1)\|^q\\ &\le \|x_n-y_n\|^q - q\lambda_n\langle J^{-1}Ax_n+\theta_n(x_n-x_1), j_q(x_n-y_n)\rangle + M_q\lambda_n^q.\end{aligned}$$

Using (7.21) and the fact that A is accretive, we have

$$\begin{aligned}\langle J^{-1}Ax_n+\theta_n(x_n-x_1), j_q(x_n-y_n)\rangle &= \langle J^{-1}Ax_n - J^{-1}Ay_n, j_q(x_n-y_n)\rangle\\ &\quad +\theta_n\|x_n-y_n\|^q\\ &\quad +\langle J^{-1}Ay_n+\theta_n(y_n-x_1), j_q(x_n-y_n)\rangle\\ &\ge \frac{\theta_n}{q}\Big\|x_n-y_n\Big\|^q.\end{aligned}$$

Therefore,

$$\|x_{n+1}-y_n\|^q \le (1-\lambda_n\theta_n)\|x_n-y_n\|^q + M_q\lambda_n^q. \tag{7.23}$$

Using again the fact that A is accretive, we have

$$\|y_{n-1} - y_n\| \leq \left\| y_{n-1} - y_n + \frac{1}{\theta_n}\Big(J^{-1}Ay_{n-1} - J^{-1}Ay_n\Big)\right\|.$$

Observing from (7.21) that

$$y_{n-1} - y_n + \frac{1}{\theta_n}\Big(J^{-1}Ay_{n-1} - J^{-1}Ay_n\Big) = \frac{\theta_n - \theta_{n-1}}{\theta_n}(y_{n-1} - x_1),$$

it follows that

$$\|y_{n-1} - y_n\| \leq \frac{\theta_{n-1} - \theta_n}{\theta_n}\|y_{n-1} - x_1\|. \tag{7.24}$$

By Lemma 1, we have

$$\begin{aligned}\|x_n - y_n\|^q &= \|(x_n - y_{n-1}) + (y_{n-1} - y_n)\|^q \\ &\leq \|x_n - y_{n-1}\|^q + q\langle y_{n-1} - y_n, j_q(x_n - y_n)\rangle.\end{aligned}$$

Using Schwartz's inequality, we obtain

$$\|x_n - y_n\|^q \leq \|x_n - y_{n-1}\|^q + q\|y_{n-1} - y_n\|\|x_n - y_n\|^{q-1}. \tag{7.25}$$

Using (7.23), (7.24), (7.25), and the fact that $\{x_n\}$ and $\{y_n\}$ are bounded, we have

$$\begin{aligned}\|x_{n+1} - y_n\|^q &\leq (1 - \lambda_n\theta_n)\|x_n - y_{n-1}\|^q + C\Big(\frac{\theta_{n-1} - \theta_n}{\theta_n}\Big) + M_q\lambda_n^q \\ &= (1 - \lambda_n\theta_n)\|x_n - y_{n-1}\|^q + (\lambda_n\theta_n)\sigma_n + \gamma_n\end{aligned}$$

for some positive constant C where

$$\sigma_n := \frac{C\Big(\frac{\theta_{n-1} - \theta_n}{\theta_n}\Big)}{\lambda_n\theta_n} = C\Big(\frac{\frac{\theta_{n-1}}{\theta_n} - 1}{\lambda_n\theta_n}\Big), \quad \gamma_n := M_q\lambda_n^q.$$

Thus, by Lemma 2, $x_{n+1} - y_n \to 0$. Using (7.22), it follows that $x_n \to y^*$ and $0 \in Ay^*$. This completes the proof.

Remark 5 For $E = L_p$ or $W^{m,p}$, $1 < p < \infty$, if $A : E \to E^*$ is maximal monotone, then $J^{-1}A : E \to E$ is m-accretive, where J is the duality mapping of E.

From Remark 5 and the fact that L_p and $W^{m,p}$ spaces are 2-uniformly smooth for $2 \leq p < \infty$ and p-uniformly smooth for $1 < p < 2$ and using Remark 3, we have the following corollaries.

Corollary 1 *Let $E = W^{m,p}$, $1 < p < \infty$, and let $A : E \to E^*$ be a bounded and maximal monotone mapping such that $A^{-1}(0) \neq \emptyset$. Then, there exists $\gamma_0 > 0$ such* *that if* $\lambda_n < \gamma_0\theta_n$ *for all* $n \geq 1$*, the sequence* $\{x_n\}$ *given by* (7.18) *converges strongly to some* $x^* \in A^{-1}(0)$*.*

Corollary 2 *Let* $E = L_p$*,* $1 < p < \infty$*, and* L^q *its dual space. Let* $A : L_p \to L_q$ *be a bounded and maximal monotone mapping such that* $A^{-1}(0) \neq \emptyset$*. For arbitrary* $x_1 \in E$*, let* $\{x_n\}$ *be a sequence defined iteratively as follows:*

$$x_{n+1} = x_n - \lambda_n \|Ax_n\|_{L_q}^{2-q} |Ax_n|^{q-2} Ax_n - \lambda_n\theta_n(x_n - x_1),\ n \geq 1, \tag{7.26}$$

where J *is the normalized duality mapping from* E *into* E^* *and* $\{\lambda_n\}$*,* $\{\theta_n\}$ *are real sequences in* $(0, 1)$ *satisfying, here and elsewhere, the following conditions:*

$$(i)\ \lim_{n\to\infty}\theta_n = 0;\ (ii)\ \sum_{n=1}^{\infty}\lambda_n\theta_n = \infty, \quad \lambda_n^{q-1} = o(\theta_n);\ (iii)$$

$$\limsup_{n\to\infty}\frac{\left(\frac{\theta_{n-1}}{\theta_n} - 1\right)}{\lambda_n\theta_n} \leq 0,\ \sum_{n=1}^{\infty}\lambda_n^q < \infty.$$

Then the sequence $\{x_n\}$ *defined by* (7.26) *converges strongly to* $x^* \in E$*, the unique solution of* $Au = 0$*.*

Corollary 3 *Let* H *be a real Hilbert space and let* $A : H \to H$ *be a bounded and maximal monotone mapping such that* $A^{-1}(0) \neq \emptyset$*. For arbitrary* $x_1 \in H$*, let* $\{x_n\}$ *be a sequence defined iteratively as follows:*

$$x_{n+1} = x_n - \lambda_n Ax_n - \lambda_n\theta_n(x_n - x_1),\ n \geq 1, \tag{7.27}$$

where J *is the normalized duality mapping from* E *into* E^* *and* $\{\lambda_n\}$*,* $\{\theta_n\}$ *are real sequences in* $(0, 1)$ *satisfying, here and elsewhere, the following conditions:*

$$(i)\ \lim_{n\to\infty}\theta_n = 0;\ (ii)\ \sum_{n=1}^{\infty}\lambda_n\theta_n = \infty, \quad \lambda_n^{q-1} = o(\theta_n);\ (iii)$$

$$\limsup_{n\to\infty}\frac{\left(\frac{\theta_{n-1}}{\theta_n} - 1\right)}{\lambda_n\theta_n} \leq 0,\ \sum_{n=1}^{\infty}\lambda_n^q < \infty.$$

Then the sequence $\{x_n\}$ *defined by* (7.27) *converges strongly to* $x^* \in E$*, the unique solution of* $Au = 0$*.*

Proof The proof follows from the fact that Hilbert spaces are 2-uniformly smooth and in addition the duality map J is identity map.

Application to Differentiable Convex Minimization Problems

In this section, we study the problem of finding a minimizer of a differentiable convex function f defined from a real Banach space E to $\mathbb{R}$.

In the sequel we will need the following result.

Lemma 4 (Chudime et al. [14]) *Let E be normed linear space and $f : E \to \mathbb{R}$ a real-valued convex function. Assume that f is bounded. Then the subdifferential map $\partial f : E \to 2^{E^*}$ is bounded on bounded subsets of E.*

Lemma 5 (Rockafellar [51]) *Let E be normed linear space and $f : E \to \mathbb{R}$ a proper, lower semicontinuous convex function. Then the subdifferential map, $\partial f : E \to 2^{E^*}$, is maximal monotone.*

We now prove the following theorem.

Theorem 4 *Let E be a q-uniformly smooth and strictly convex real Banach. Let $f : E \to \mathbb{R}$ be a differentiable, bounded convex real-valued function which satisfies the growth condition: $f(x) \to +\infty$ as $\|x\| \to +\infty$. For arbitrary $x_1 \in E$, let $\{x_n\}$ be the sequence defined iteratively by*

$$x_{n+1} = x_n - \lambda_n J^{-1}(df(x_n)) - \lambda_n \theta_n (x_n - x_1), \quad n \geq 1, \tag{7.28}$$

where J is the normalized duality mapping from E into E^* and $\{\lambda_n\}$, $\{\theta_n\}$ are real sequences in $(0, 1)$ satisfying, here and elsewhere, the following conditions:

$$(i) \lim_{n\to\infty} \theta_n = 0; (ii) \sum_{n=1}^{\infty} \lambda_n \theta_n = \infty, \quad \lambda_n^{q-1} = o(\theta_n); (iii)$$

$$\limsup_{n\to\infty} \frac{\left(\frac{\theta_{n-1}}{\theta_n} - 1\right)}{\lambda_n \theta_n} \leq 0, \ \sum_{n=1}^{\infty} \lambda_n^q < \infty.$$

Then, f has a unique minimizer $x^* \in E$, and the sequence $\{x_n\}$ defined by (7.28) converges strongly to a^*.

Proof Since E is q-uniformly smooth, then it is reflexive. Therefore, from the growth condition, the continuity, and the strict convexity of f, f has a unique minimizer x^* characterized by $df(x^*) = 0$. Finally, from Lemma 4 and Lemma 5, the differential map $df : E \to E^*$ is bounded and maximal monotone. Therefore, the proof follows from Theorem 3.

Acknowledgments The authors thank the referees and the editors of the Journal for their work and their valuable suggestions that helped to improve the presentation of this paper.

References

1. Adoum, A. A., Sene, M., Ndiaye, M., Djitte, N., *A Mann Type iterative algorithm for bounded and strongly monotone mappings in certain Banach spaces*, In: Seck, D., Kangni, K., Nang, P., Salomon Sambou, M. (eds) Nonlinear Analysis, Geometry and ApplicationsTrends in Mathematics. Birkhäuser, Cham, 2023
2. Berinde, V., *Iterative Approximation of Fixed Points*, Lecture Notes in Mathematics, Springer, London, UK, 2007.
3. Berinde, V., St. Maruster, I.A. Rus, *An abstract point of view on iterative approximation of fixed points of nonself operators,* J. Nonlinear Convex Anal. 15 (2014), no. 5, 851–865.
4. Browder, F.E., *Nonlinear mappings of nonexpansive and accretive type in Banach spaces*, Bull. Amer. Math. Soc. vol. 73 (1967), 875–882.
5. Bruck Jr, R. E., *A strongly convergent iterative solution of* $0 \in U(x)$ *for a maximal monotone operator* U *in Hilbert spaces*, J. Math. Anal. Appl., vol. 48 (1974), 114–126.
6. Bynum, W. L., *Weak parallelogram laws for Banach spaces*, Canad. Math. Bull. 19(3) (1976), 269–275.
7. Reich S.; *Strong Convergence theorems for resolvents of accretive operators in Banach spaces*, in *J. Math. Anal. Appl.* **183** (1994), 118–120.
8. Censor Y., Reich, S., *Iterations of paracontractions and firmly nonexpansive operators with applications to feasibility and optimization*, Optimization, vol. 37, no. 4 (1996), pp. 323–339.
9. Chidume, C. E., *Geometric Properties of Banach Spaces and Nonlinear iterations*, vol. 1965 of Lectures Notes in Mathematics, Springer, London, UK, 2009.
10. Chidume, C. E., Ali Bashir, *Approximation of common fixed points for finite families of nonself asymptotically nonexpansive mappings in Banach spaces*, J. Math. Anal. Appl., vol. 326 (2007), pp. 960–973.
11. Chidume, C. E., *An approximation method for monotone Lipschitzian operators in Hilbert-spaces*, Journal of the Australian Mathematical Society series A-pure mathematics and statistics, vol. 41 (1986), pp. 59–63.
12. Chidume, C. E., *Convergence theorems for asymptotically pseudo-contractive mappings*, Nonlinear Analysis-Theory Methods & Applications, vol. 49 (2002), pp. 1–11.
13. Chidume, C. E., *Iterative approximation of fixed points of Lipschitzian strictly pseudo-contractive mappings*, Proc. Amer. Math. Soc., vol. 99 (1987), no. 2,pp. 283–288.
14. Chidume, C. E., Djitte, N., Sene, M., *I*terative algorithm for zeros of multivalued accretive operators in certain Banach spaces, Afr. Mat., 26(2015), no. 3–4, 357–368.
15. Chidume, C. E., Chidume, C. O., *Convergence theorems for fixed points of uniformly continuous generalized Phi-hemi-contractive mappings*, J. Math. Anal. Appl., vol. 303 (2005), pp. 545–554.
16. Chidume, C. E., Chidume, C. O., *Convergence theorem for zeros of generalized Phi-quasi-accretive operators*, Proc. Amer. Math. Soc., vol. 134 (2006), pp. 243–251.
17. Chidume, C. E., Djitte, N., *Iterative algorithm for zeros of bounded* m*-Accretive nonlinear operators*, *to appear*, J. Nonlinear and convex analysis.
18. Chidume, C. E., Djitte, N., *Strong convergence theorems for zeros of bounded maximal monotone nonlinear operators*, J. Abstract and Applied Analysis, **Volume 2012**, Article ID 681348, 19 pages, doi:10.1155/2012/681348.
19. Chidume, C. E., Osilike, M. O., *Iterative solutions of nonlinear accretive operator equations in arbitrary Banach spaces*, Nonlinear Analysis-Theory Methods & Applications, vol. 36 (1999), pp. 863–872.
20. Chidume C. E. and Zegeye H.; *Approximate fixed point sequence and convergence theorems for Lipschitz pseudo-contractive maps*, Proc. Amer. Math. Soc. **132** (2004), no. 831–840.
21. Chidume C. E., Mutangadura S. A., *An example on the Mann itération method for Lipschitz pseudocontrations,* Proc. Amer. Math. Soc. **129** (2001), no. 8, 2359–2363.
22. Chidume C. E., *Geometric Properties of Banach spaces and Nonlinear Iterations*, Springer Verlag Series: Lecture Notes in Mathematics, Vol.1965 (2009), ISBN 978-1-84882-189-7.

23. L. Deng, *On Chidume's open question*, J. Math. Appl. vol. 174 (1993), no. 2, 441–449.
24. Goebel, K., Reich, S., *Uniform convexity, hyperbolic geometry, and nonexpansive mappings.* Monographs and Textbooks in Pure and Applied Mathematics, vol. 83, Marcel Dekker, inc., New York, 1984.
25. G uler. O., *On the convergence of the proximal point algorithm for convex minimization*, SIAM Journal on Control and Optimization, (1991) vol. 29, no. 2, pp. 403–419.
26. Kamimura, S., Takahashi, W., *Approximating solutions of maximal monotone operators in Hilbert spaces*, Journal of Approximation Theory, vol. 106 (2000), no. 2, pp. 226–240.
27. Kamimura, S., Takahashi, W., *Strong convergence of a proximal-type algorithm in a Banach space*, SIAMJ. Optim., vol. 13 (2002), no. 3, pp. 938–945.
28. Kato, T., *Nonlinear semi groups and evolution equations*, *J. Math. Soc. Japan* **19** (1967), 508–520.
29. Kohsaka, F., Takahashi, W., *Strong convergence of an iterative sequence for maximal monotone operator in Banach space*,Abstr. Appl. Annal. 3 (2004) 239–349.
30. Liu, L., *Ishikawa and Mann iterative process with errors for nonlinear strongly accretive mappings in Banach spaces*, J. Math. Anal. Appl., vol. 194 (1995), no. 1, pp. 114–125.
31. Martinet, B., *regularization d' inéquations variationelles par approximations successives*, Revue Francaise d'informatique et de Recherche operationelle,**4**(1970), 154–159.
32. Mendy J.T., Sene M., Djitte N., *Algorithm for zeros of maximal monotone mappings in classical Banach spaces*, Int. J. Math. Anal., Vol. 11, 2017, no. 12, 551–570.
33. Moudafi, A., *Alternating CQ-algorithm for convex feasibility and split fixed-point problems,* J. Nonlinear Convex Anal. 15 (2004), no. 4, 809–818.
34. Moudafi, A., *A relaxed alternating CQ-algorithm for convex feasibility problems,* Nonlinear Anal. 79 (2003), 117–121.
35. Moudafi, A., *Proximal methods for a class of bilevel monotone equilibrium problems,* J. Global Optim. 47 (2010), no. 2, 45–52.
36. Moudafi, A., Thera, M., *Finding a zero of the sum of two maximal monotone operators,* J. Optim. Theory Appl. 94 (1997), no. 2, 425–448.
37. Pascali, D., Sburian, S., *Nonlinear mappings of monotone type*, Editura Academia Bucuresti, Romania, 1978.
38. Qihou, L., *The convergence theorems of the sequence of Ishikawa iterates for hemi-contractive mapping*, J. Math. Anal. Appl., vol. 148 (1990), pp. 55–62.
39. Reich, S., *A weak convergence theorem for the alternating methods with Bergman distance*, in: A. G. Kartsatos (Ed.), Theory and Applications of Nonlinear Operators of Accretive and Monotone Type, in Lecture notes in pure and Appl. Math., vol. 178 (1996), Dekker, New York. pp. 313–318.
40. Reich, S., *Iterative methods for accretive sets,* Nonlinear Equations in Abstract Spaces, pp. 317–326, Academic Press, New York, 1978.
41. Reich, S., *Extension problems for accretive sets in Banach spaces*, J. Functional Anal. 26 (1977), 378–395.
42. Reich, S., *Iterative methods for accretive sets in Banach Spaces,* Academic Press, New York, 1978, 317–326.
43. Reich, S., *Constructive techniques for accretive and monotone operators*, Applied non-linear analysis, Academic Press, New York (1979), pp. 335–345.
44. Reich, S., and Sabach, S., *A strong convergence theorem for a proximal-type algorithm in reflexive Banach spaces*, Journal of Nonlinear and Convex Analysis, vol. 10 (2009), no. 3, pp. 471–485.
45. Reich, S., and Sabach, S., *Two strong convergence theorems for a proximal method in reflexive Banach spaces*, Numerical Functional Analysis and Optimization, vol. 31 (2010), no. 1–3, pp. 22–44.
46. Solodov, M, V., and Svaiter, B. F., *Forcing strong convergence of proximal point iterations in a Hilbert space*, Mathematical Programming, vol. 87 (2000), pp. 189–202.
47. Weng, X. L., *Fixed point iteration for local striclty pseudo-contractive mappings*, Proc. Amer. Math. Soc., vol. 113 (1991), no. 3, 727–731.

48. Kamimura, S., Takahashi, W., *Approximating solutions of maximal monotone operators in Hilbert spaces*, J. Approx. Theory 106 (2000) 226–240.
49. William, K., Shahzad, N., *Fixed point theory in distance spaces*, Springer Verlag, 2014.
50. Xiao, R., *Chidume's open problems and fixed point theorems*, Xichuan Daxue Xuebao, vol. 35 (1998), no. 4, pp. 505–508.
51. Rockafellar, R. T., *Monotone operator and the proximal point algorithm*, SIAM J. Control Optim. **14** (1976) 877–898.
52. Xu, Z. B., *A note on the Ishikawa iteration schemes*, J. Math. Anal. Appl., vol. 167 (1992), pp. 582–587.
53. M. Sene, M. Ndiaye and N. Djitte, *A new Krasnoselskii's type algorithm for zeros of strongly monotone and Lipschitz mappings*, CREAT. MATH. INFORM. Volume 31 (2022), No. 1, Pages 109–120.
54. Xu, Z. B., Roach, G. F., *Characteristic inequalities for uniformly convex and uniformly smooth Banach space*, J. Math. Anal. Appl., vol 157 (1991), pp. 189–210.
55. Xu, Z. B., Jiang, Y. L., Roach, G. F., *A further necessary and sufficient condition for strong convergence of nonlinear contraction semigroups and of iteration methods for accretive operators in Banach spaces*, Proc. Edinburgh Math. Soc., vol. 38 (1995), no. 2, pp. 1–12.
56. Xu, H. K., *Inequalities in Banach spaces with applications*, Nonlinear Anal., Vol. 16 (1991), no. 12, pp. 1127–1138.
57. H. K., *Iterative algorithms for nonlinear operators,* J. London Math. Soc. **66**(2002), no2, 240–256.
58. Xu, Y., *Existence and convergence for fixed points of mappings of the asymptotically nonexpansive type*, Nonlinear Anal., vol. 16 (1991), pp. 1139–1146.
59. Zegeye, H. *Strong convergence theorems for maximal monotone mappings in Banach spaces*, J. Math. Anal. Appl.,343 (2008) 663–671.
60. Zhou, H., Jia, Y., *Approximating the zeros of accretive operators by the Ishikawa iteration process*, Abstr. Appl. Anal. vol. 1 (1996), no. 2, 153–167.
61. Zhu, L., *Iteration solution of nonlinear equations involving m-accretive operators in Banach spaces*, J. Math. Anal. Appl., vol. 188 (1994), 410–415.

Chapter 8
On Rickart and Baer Semimodules

Mamadou Lamine Diallo, Jean Raoult Tsiba, and Djiby Sow

Abstract This chapter generalizes the Rickart (resp., Baer) property on semirings and semimodules. We introduce weak Rickart (resp., Baer) semimodules and then identify i-Rickart semimodules as a specific subclass of the former. Basic links between the different Baer and Rickart semimodules are discussed. A characterization of the Rickart semimodules by their endomorphism semiring is provided.

Keywords Weak Rickart semirings · Weak Rickart semimodules · Weak Baer semirings · Weak Baer semimodules · i-Rickart semirings · i-Rickart semimodules · i-Baer semirings · i-Baer semimodules

Introduction

An idempotent e of a ring R is such that $e^2 = e$. Then clearly the corresponding decomposition $R = eR \oplus (1 - e)R$ is useful to determining the structure of R. The Rickart and Baer properties are based on connections of idempotents to annihilators of a ring. It is useful in solving a linear equation in one unknown $ax = b$, where $a \neq 0$. The general class of rings in which the last equation is solvable includes that of Rickart and Baer. Again a good setting to solve completely a finite system of linear equations is R such that the direct sum of copies of R inherits the Rickart and Baer property.

In 1946, C. E. Rickart studied $C^\star$-algebras, Banach algebras with involution * such that $||xx^\star|| = ||x||^2$, which satisfy that the right annihilator of every single element is generated by a projection ($e^2 = e$ and $e^\star = e$).

In 1960, S. Maeda [17] defined Rickart rings in an arbitrary setting after Kaplansky's work on Baer semirings in 1955 [5]. A ring is called right Rickart if an idempotent generates the right annihilator of every single element. Hattori

M. L. Diallo (✉) · J. R. Tsiba · D. Sow
Department of Mathematics and Computer Sciences, Cheikh Anta Diop University of Dakar, Dakar, Sénégal

C. T. Gueye et al. (eds.), *Mathematics of Computer Science, Cybersecurity and Artificial Intelligence*, Springer Proceedings in Mathematics & Statistics 468,
https://doi.org/10.1007/978-3-031-66222-5_8

[10] introduced in 1960 right p.p. rings, rings with the property that every principal right ideal of R is projective; it was later shown that right p.p. rings are precisely right Rickart rings. Much work was developed by authors such as Berberian, E.P. Armendariz [2], S.K. Berberian, G.M. Bergman [3], S. Endo [6], M.W. Evans [7], I. Kaplansky [12], and W.K. L.Small [19]).

In 2004, Baer modules were introduced, for the first time, by Rizvi and Roman [18]. M is called right Baer semimodule if the right annihilator in M of every subset of the endomorphism semiring S of M is generated by an idempotent of S. Replacing every subset by every element of S, Rizvi and Roman defined again Rickart modules in 2010 [15]. Although many works generalize the Rickart and Baer properties in module theory, as far as we know, Gupta et al. first introduce Rickart semirings theory [8]. Semirings generalize rings and distributive bounded lattices not necessarily with subtraction. Semirings and their semimodules differ from rings and modules, respectively, in that they do not necessarily admit subtraction. They have important uses in many areas of Computer Science and Mathematics, such as Automata Theory, Tropical Geometry, and Idempotent Analysis. Our reference is Golan's book [9] for semirings and their semimodules, Wisbauer's book [20] for modules, and Birkenmeier–Park–Rizvi's book [4] for the Rickart property.

In our work we introduce the Rickart (resp., Baer) property on semirings and semimodules. We study weak Rickart (resp., weak Baer) semimodules and i-Rickart semimodules. Many properties on Baer and Rickart semirings and semimodules over semirings are discussed.

This chapter is divided into three sections as it follows.

In section "Introduction", we present some preliminaries.

In section "Preliminaries", we characterize some Rickart (Baer) semimodules by their endomorphisms.

In section "Characterizations of i-Rickart Semimodules", we exhibit links between the properties and generalize a well-known result of L. Small.

Preliminaries

Basic Notions

Throughout this chapter, R is a semiring with unit and M is a unital right R-semimodule. For a right R-semimodule M, $S = End_R(M)$ will denote the endomorphism semiring of M. Then M can be viewed as a left S-right R-bisemimodule. For $\varphi \in S$, $\ker(\varphi)$ and $im\varphi$ (or φM) stand for the kernel and the image of φ (or proper image), respectively. The notations $N \subseteq M$, $N \leqslant M$, $N \leqslant^{esse} M$, $N \leqslant^{\oplus} M$, $N \leqslant^{\bar{\oplus}} M$, and $N \trianglelefteq M$ mean that N is a subset, a subsemimodule, an essential subsemimodule, a direct summand, a weak summand, and a fully invariant subsemimodule of M, respectively. By $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{Z}$, and $\mathbb{N}$, we denote the ring of reals, rationals, integers, and natural numbers, respectively. $\mathbb{N}_n$

will denote $\mathbb{N}/\mathbb{N}_n$ and $Mat_{n\times n}(R)$ denotes an $n \times n$ matrix semiring over R. For M and M' two R-semimodules, $\hom_R(M, M')$ and $\hom_R(M, M) = End_R(M)$ denote the semiring of homomorphism from M to M' and the endomorphism semiring of M, respectively. For $g \in \hom_R(M, M')$ and $f \in \hom_R(M', M'')$ such that $g(M) \subseteq M'$, the composition of maps fog is denoted fg, and then $fog(m) = fg(m) = f(g(m)), \forall m \in M$. We also denote $r_M(I) = \{m \in M|Im = 0\}$, $r_S(I) = \{\varphi \in S|I\varphi = 0\}$ for $\emptyset \neq I \subseteq S$; $r_R(N) = \{r \in R|Nr = 0\}$, $l_S(N) = \{\varphi \in S|\varphi N = 0\}$ for $N \leqslant M$.

Recall the following definitions and properties:

(a) Let M, N be R-semimodules and $f \in \hom_R(M, N)$:

1. $f(M) = imf = \{f(m) : m \in M\}$ is said to be the proper image of f.
2. $Imf = \{b \in M : b + f(m) = f(m^{'}),$ for some $m, m^{'} \in M\}$ the image or extended image of f. $Imf = \{x \in M : x =_{imf} 0_M\}$. $imf \subseteq Imf$. $\ker(f) = \{m \in M : f(m) = 0_N\}$ is the kernel of f.
3. f is said to be k-regular, if $f(x_1) = f(x_2)$ implies $x_1 + k_1 = x_2 + k_2$, for some $k_1, k_2 \in \ker(f)$. f is said to be i-regular if $f(M) = Im(f)$. f is said to be regular if f is i-regular and k-regular.
4. $f(M) \subsetneq Imf$. And $f(M) = Imf$ iff f is i-regular.

(b) Let R be a semiring, $a \in R$, and $S \subseteq R$ a subset:

1. The right annihilator of $a \in R$ is the ideal $r_R(a) = \{r \in R : ar = 0\}$. Similarly, $l_R(a)$ means the left annihilator of a.
2. R is called right Baer semiring if for every nonempty subset S of R, there exists an idempotent $e \in R$ such that $r_R(S) = eR$ [8].
3. R is called right Rickart semiring (or p.p. semiring) if for every $a \in R$, there exists an idempotent $e \in R$ such that $r_R(a) = eR$ [8].

(c) Let N_1 and N_2 be subsemimodules of M:

1. If $M = N_1 + N_2$ and $N_1 \cap N_2 = \{0\}$, then M is a weak sum of N_1 and N_2, and we denote $M = N_1 \bar{\oplus} N_2$. The decomposition of $m \in M$ into $x_i \in N_i$ is not necessarily unique.
2. Let $M = N_1 + N_2$ and $N_1 \cap N_2 = \{0\}$ such that every element m of M decomposes uniquely $m = n_1 + n_2$, where $n_i \in N_i$. Then M is called direct sum and N_i a direct summand of M. This is denoted by $M = N_1 \oplus N_2$. By Remark 2.2. in [1], a direct summand of M is subtractive.
3. For every subsemimodule N of M, the Bourne congruence relation on M is defined by: $m_1, m_2 \in M, m_1 \equiv_N m_2$ if $m_1 + n_1 = m_2 + n_2$, with $n_1, n_2 \in N$. The restriction of the Bourne equivalence relation $\equiv_{N_1}$ to N_2 and $\equiv_{N_2}$ to N_1 is trivial, meaning that the projection p_1 of M on N_1 along N_2 and that of M on N_2 along N_1 , p_2, is well defined on M.

(d) Let $I \leq R$ be an ideal and $N \leq M$ a subsemimodule:

1. I is subtractive, if $a + b \in I$ and $a \in I$ imply that $b \in I$ for all $a, b \in R$. Similarly, N is subtractive, if $n + n' \in N$ and $n \in N$ imply that $n' \in N$ for all $n, n' \in M$.
2. Direct summands of M in [1], $\ker(f)$, $Im(f)$, and annihilators are subtractive.
3. If $M = M_1 \oplus M_2 = M$ and $N \leq M$ is subtractive, then $N = N_1 \oplus N_2$, where $N_i = N_i \cap M_i$ from Lemma 2.3. of [1] (semi-modularity law).

(e) For the definition of a Baer module and a Rickart module, refer to [16] and [18].

Introduction to Rickart Semirings and Semimodules

We introduce weak Rickart semirings and weak Rickart semimodules.

Weak Rickart Semirings

M and N are R-semimodules, I an ideal of a semiring S, and $e \in \hom_R(M, N)$. Recall that:

1. $\overline{e(M)} = \{n \in N / m + e(m_1) = e(m_2),\ \ m_1, m_2 \in M\} = \{n \in N / n \equiv_{eM} 0\}$.
2. $\overline{I} = \{s \in S / s + i_1 = i_2,$ where $i_1, i_2 \in I\} = \{s \in N / s \equiv_I 0\}$.

$\overline{e(M)}$, denoted $Im(e)$ or $Im_N(e)$, and I are the subtractive closure of eM and I, respectively. For $a \in S$, let $I = aS$, and then $\overline{aS}$ allows defining the notion of i-regular element of S.

Definition 8.1 Let a be an element of a semiring S:

i a is said to be left (resp., right) i-regular element, if $aS = \overline{aS}$ (resp., $Sa = \overline{Sa}$).
ii a is i-regular, if a is left and right regular, and in others words $\overline{aS} = aS = Sa = \overline{Sa}$.

Note that if S is commutative, then a left i-regular is right i-regular, and thus it is i-regular.

Recall that in [17] a ring R is right Rickart, if for every $a \in R$, $r_R(a) = eR$, with $e^2 = e$ in R. Then $r_R(a) = eR$ is a direct summand of R, which is not the case in coming generalizations.

Gupta et al. in [8] have proposed a definition of Rickart semiring that extends the property of Rickart rings in semirings theory. In their sense a semiring S is said to be a Rickart semiring if, for every $r \in S$, there exists an idempotent e of S such that $r_S(r) = eS$. The following new definition generalizes Rickart semiring in semiring theory. In fact the generating idempotent e is not necessarily i-regular.

Definition 8.2 A commutative semiring S is said to be a weak-Rickart (w-Rickart briefly) semiring if, for every $r \in S$, there exists an idempotent e of S such that $r_S(r) = Im_S(e)$.

Definition 8.3 Let M be a R-semimodule and $\varphi \in End_R(M)$. M is said to be a weak-Rickart, w-Rickart briefly, semimodule if, for every $\varphi \in End_R(M)$, there exists an idempotent e of $End_R(M)$ such that $\ker(\varphi) = Im(e)$.

Example 8.1 Let $M = (\{0, 1, 2, 3\}, \gcd, 0)$ and $\mathbb{B} = (\{0, 1\}, +, \times, 0, 1)$, the boolean semiring. Then M is a $\mathbb{B}$-semimodule, with scalar multiplication defined, for every $(m, r) \in M \times \mathbb{B}$, by $m.r = 0$, if $r = 0$, else $m.r = m$. $\{0\}$, $\{0, 2\}$, $\{0, 3\}$ and M stand for kernel of some endomorphism of M and are generated by non-i-regular idempotent. One can verify that φ_2, such that $\varphi_2(1) = \varphi_2(3) = 1$ and $\varphi_2(2) = 2$, is an idempotent of $End(M)$. Then $\ker(0_S) = \overline{\varphi_2(M)} = \overline{\{0, 1, 2\}} = Im_M(\varphi_2) = \{0, 1, 2, 3\}$. It is true that $\gcd(3, \varphi_2(1)) = \varphi_2(1) \Leftrightarrow \gcd(3, 1) = 1$, and then $3 \in Im_M(\varphi_2)$, while $3 \notin \varphi_2(M)$.

We propose a semiring which is not Rickart.

Example 8.2 The semiring $(R = \begin{pmatrix} \mathbb{N} & \mathbb{N} \\ 0 & \mathbb{N} \end{pmatrix}, +, \times, 0_R, 1_R)$ has $End_R(R) \cong R$. First, the set of idempotents of R is $Id(R) = \{0_R, e_{10}, e_{01}, e_{11}, 1_R\} \cup \{e_p;\, p \in \mathbb{N}\}$, where : $0_R = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, $e_{10} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $e_{01} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, $e_{11} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, $e_p = \begin{pmatrix} 0 & p \\ 0 & 1 \end{pmatrix}$, and $1_R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Let $r_0 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, and then $r_R(x_0) = \begin{pmatrix} 0 & \mathbb{N} \\ 0 & 0 \end{pmatrix}$. Since $\forall e \in Id(R)$, $r_R(x_0) \neq Im_R(e)$, R is not Rickart.

Example 8.13 proves that the class of i-Rickart semimodules is strictly contained in that of w-Rickart semimodules.

Example 8.3 Let R be the semiring as in Example 8.13, and it is remarkable that R is a w-Rickart R-semimodule without being an i-Rickart semimodule.

Characterizations of i-Rickart Semimodules

Endomorphism Semiring of i-Rickart Semimodules

If M is a w-Rickart R-semimodule, then the Rickart property on M may go through $S = End_R(M)$. This happens, if for every $\varphi \in S$, $\ker(\varphi) = Im_M(e)$ with e an i-regular idempotent of S. We identify a subclass of w-Rickart semirings and w-Rickart semimodules too, using the newly introduced notion of i-regular idempotent. Example 8.13 (resp.,8.1) proves that the class of i-Rickart semirings (resp., semimodules) is strictly contained in that of w-Rickart semirings (resp., semimodules).

Definition 8.4 A semiring S is said to be an i-Rickart semiring if, for every $r \in S$, there exists an i-regular idempotent e of S such that $r_S(r) = eS$.

The next two examples are i-Rickart semirings as integral semidomains. Let icm(x, y) be the smallest common multiple of the numbers x and y.

Example 8.4 *a.* The semiring $(R = \{0, 1, 2\}, \max, \min, 0, 2)$
b. The semiring $(R = \{0, 1, 2\}, \gcd, \mathrm{lcm}, 0, 1)$, with $\gcd(0, 0) = 0$

In the following, we specify i-Rickart semimodules as a subclass of w-Rickart semimodules.

Definition 8.5 A R-semimodule M is said to be an i-Rickart semimodule if, for every $\varphi \in End_R(M)$, there exists an i-regular idempotent e of $End_R(M)$ such that $\ker(\varphi) = Im_M(e) = e(M)$.

Example 8.5 Let $\mathbf{B} = \{0, 1\}$ be the boolean semiring and $A = \prod_{i=1}^{\infty} \mathbf{B}$ the usual semiring product. It is a commutative and idempotent semiring.

Consider the next infinite sequence of 0 and 1:

- $T_0 = \{(a_n)_{n=1}^{\infty} \in A$ such that $\exists\, k \in \mathbf{N}, \forall n \geqslant k, a_n = 0\}$.
- $T_1 = \{(a_n)_{n=1}^{\infty} \in A$ such that $\exists\, k \in \mathbf{N}, \forall n \geqslant k, a_n = 1\}$.
- $\mathbf{B}^{\infty} = T_0 \cup T_1$.

As $\mathbf{B}^{\infty}$ is closed under laws on A and contains the zero sequence and the unit one, it is a subsemiring of $\prod_{i=1}^{\infty} \mathbf{B}$, and then it is a semiring. Let us prove that S is i-Rickart semiring using the definition.

Note that every sequence $a = (a_n)_{n=1}^{\infty}$ has a unique complement $a^{\perp} = (a_n^{\perp})_{n=1}^{\infty}$ sequence, where $aa^{\perp} = a^{\perp}a = 0$ and $a^{\perp} + a = 1_A$. Thus for every a, $r_{\mathbf{B}^{\infty}}(a) = a^{\perp}(\mathbf{B}^{\infty})$, so that $\mathbf{B}^{\infty}$ is i-Rickart.

Lemma 8.1 *Let M and N be R-semimodules and $f \in \hom_R(M, N)$ an R-homomorphism. Assume that $\ker(f) = eM$ where e is an idempotent of $End_R(M)$. Then $fe = 0$, and for each $h \in End_R(M)$, we have: $fh = 0$ implies that $h = eh$.*

Proof $\ker(f) = eM$ implies that $fe(M) = f(eM) = f(\ker(f)) = 0 \Leftrightarrow fe = 0$.

For the second part, we have $fh = 0$ implies that $h(M) \subseteq \ker(f) = eM$. Let $m \in M$, then $h(m) \in h(M) \subseteq eM$, and thus there exists $m' \in M$ such that $h(m) = e(m') \Rightarrow eh(m) = ee(m') = e(m') = h(m)$. We conclude that $h = eh$.

The recall of in part **a** obviously leads to the following useful remark.

Remark 8.1 Let $f, g \in End_R(M)$, where M is an R-semimodule. If $\ker(f) = Im_M(g)$ and g is i-regular, then $\ker(f) = Im_M(g) = gM$.

Proposition 8.1 *Let M be an i-Rickart R-semimodule. Then $End_R(M)$ is an i-Rickart semiring.*

Proof To prove that S is an i-Rickart semiring, let us justify that every nonzero $\varphi \in S$ $r_S(\varphi)$ is generated by an i-regular idempotent $e \in S$, clearly $r_S(\varphi) = Im_S(e)$.

M is an i-Rickart semimodule and then $r_M(\varphi) = Im_M(e)$, for some i-regular idempotent $e \in S$. Hence $\varphi(Im_M(e)) = 0$. Since $eM \subseteq Im_M(e)$, then $\varphi(eM) = 0 = \varphi e(M)$ so that $\varphi e = 0$ or $e \in r_S(\varphi)$. We wish to show that $r_S(\varphi) = Im_S(e)$:

- Let $z \in Im_S(e)$, and then $z + e(s_1) = e(s_2)$ so that $\varphi(z) + \varphi e(s_1) = \varphi e(s_2) \Leftrightarrow \varphi(z) + 0 = 0 \Leftrightarrow z \in r_S(\varphi)$. Then we do have: $Im_S(e) \subseteq r_S(\varphi)$.
- Conversely, let $\psi \in r_S(\varphi)$, and then $\varphi\psi = 0 \Rightarrow \varphi(\psi(M)) = 0$, which leads to $\psi(M) \subseteq r_M(\varphi)$ and $\psi(m) \in r_M(\varphi)$, for all $m \in M$. Since $r_M(\varphi) = Im_M(e) = eM$, as e is i-regular, then $\psi(M) \subseteq eM$. Due to Lemma 8.1 and Remark 8.1, we obtain that $\psi = e\psi \in eS \subseteq Im_S(e)$, proving $r_S(\varphi) = Im_S(e) = eS$, which justifies $End_R(M)$ is an i-Rickart semiring.

Example 8.6 Let M be the R-semimodule in Example 8.1. $End_R(M)$ can be summarized as follows:
$\varphi_4^2 = \varphi_0$ and $\forall i \neq 4$, $\varphi_i^2 = \varphi_i$; $\varphi_2\varphi_3 = \varphi_4\varphi_3 = \varphi_4\varphi_5 = \varphi_2$; $\varphi_5\varphi_3 = \varphi_3\varphi_2 = \varphi_3$; $\varphi_2\varphi_4 = \varphi_2\varphi_5 = \varphi_4$ and $\varphi_3\varphi_4 = \varphi_3\varphi_5 = \varphi_3$. Note that φ_3 and φ_5 are not i-regular.

Let us introduce a generalized version of the notion of k-local retractability for semimodules and prove that it is the necessary condition to ensure that M and $End_R(M)$ are simultaneously w (resp., i)-Rickart. Zelmanowitz regular modules are proved to be k-local-retractable [21].

Definition 8.6 A R-semimodule M is called w-k-local-retractable (resp., i-k-local-retractable) if for every $\varphi \in End_R(M)$ and every nonzero element $m \in r_M(\varphi) = \ker(\varphi)$, there exists a homomorphism $\psi_m : M \to r_M(\varphi)$ such that $m \in Im_M(\psi_m) \subseteq r_M(\varphi)$ (resp., $m \in \psi_m(M) \subseteq r_M(\varphi)$).

Proof

(a) Every w-Rickart R-semimodule M is w-k-local-retractable.
(b) Every i-Rickart R-semimodule M is i-k-local-retractable.

Proof Let M be a w-Rickart semimodule, $\varphi \in End_R(M)$, and $0 \neq m \in r_M(\varphi)$. We have $r_M(\varphi) = Im_M(e)$, as M is a w-Rickart semimodule, and thus $m \in Im_M(e) \subseteq r_M(\varphi)$. Hence M is w-k-local-retractable.
Similarly, we prove (b).

The following lemma is proved similarly as Lemma 8.1.

Lemma 8.2 *Let R be a semiring and $r \in R$. Assume that $r_R(r) = eS$ with e an idempotent of S. Then $re = 0$, and for $r' \in S$, we have : $rr' = 0$ implies that $r' = er'$.*

Proposition 8.2 *For a R-semimodule M, the following assertions are equivalent:*

(a) M is an i-Rickart R-semimodule.
(b) $End_R(M)$ is an i-Rickart semiring and M is i-k-local-retractable.

Proof $a) \Rightarrow b)$ comes out of Proposition 8.1 and Proposition .

$b) \Rightarrow a)$ Let $\varphi \in End_R(M) = S$. Let us prove : $r_M(\varphi) = \ker(\varphi) = Im_M(e)$, where $e^2 = e$ is i-regular.

Since S is i-Rickart, $r_S(\varphi) = Im_S(e) = eS$, for some i-regular $e^2 = e$. Then $\varphi e = 0$, which implies that $Im_M(e) \subseteq ker\varphi$. Conversely, let $m \in r_M(\varphi)$ be nonzero. Since M is i-k-local-retractable, there exists ψ_m such that $m \in Im_M(\psi_m) \subseteq \ker(\varphi)$, and thus we deduce $\varphi\psi_m = 0$, so $\psi_m \in r_S(\varphi)$. Since S is i-Rickart, then $r_S(\varphi) = Im_S(e) = eS$ implies that $\psi_m \in eS$, so that by Lemma 8.1 $\psi_m = e\psi_m$. Moreover $m \in Im_M(\psi_m) \Leftrightarrow m \in Im_M(e\psi_m) \Leftrightarrow \exists x_1, x_2 : m + e\psi_m(x_1) = e\psi_m(x_2) \Leftrightarrow \exists x_1, x_2 : m + e(\psi_m(x_1)) = e(\psi_m(x_2))$. Hence $m \in Im_M(e)$ and $r_M(\varphi) \subseteq Im_M(e)$.

Endomorphism Semiring of i-Baer Semimodules

We extend the Baer property to semirings by identifying w-Baer semirings and i-regular Baer semirings, a subclass of the former. Here we require the annihilator of any subset to have an idempotent generator.

Definition 8.7 Let S be a semiring:

i. S is a weak-Baer (w-Baer, briefly) semiring if, for every subset $S' \in S$, there exists an idempotent e of S such that $r_S(S') = Im_S(e)$.
ii. S is an i-Baer semiring if, for every $S' \in S$, there exists an i-regular idempotent e of S such that $r_S(S') = eS$.

Similarly, we define weak-Baer semimodules and i-regular Baer semimodules.

Definition 8.8 Let M be a R-semimodule, and $S' \subseteq S = End_R(M)$:

i. M is a w-Baer semimodule if $r_M(S') = Im(e)$ for some idempotent $e \in S$.
ii. M is an i-Baer if $r_M(S') = Im(e)$ for some i-regular idempotent e of $End(M)$.

Example 8.7

- $\mathbb{N}$ and $\mathbb{Q}^+$ are integral semidomains. Thus $\mathbb{N}$ and $\mathbb{Q}^+$ are Baer (Rickart) semirings. Hence $\mathbb{N}_\mathbb{N}$ and $\mathbb{Q}^+_{\mathbb{Q}^+}$ are Baer (Rickart) semimodules.
- $\mathbb{B}_2$ the boolean semiring is a Baer $\mathbb{N}$-semimodule as $S = End_\mathbb{B}(\mathbb{Z}_2) = \{0_S, 1_S\}$, so that $\ker(\varphi)$ is trivial for every $\varphi \in S$.
- To show that $\mathbb{Q}^+$ is a Baer $\mathbb{N}$-semimodule, it suffices to prove that $\ker(\varphi) = 0$ for every nonzero $\varphi \in End_\mathbb{N}(\mathbb{Q}^+)$.

Theorem 8.1 *Let M be an i-Baer semimodule. Then $End_R(M)$ is an i-Baer semiring.*

Proof Let $I \subseteq End_R(M)$ be a subset. We wish to prove that $r_S(I)$ is generated by an i-idempotent of $End_R(M)$ to satisfy the definition of an i-Baer semiring. We already know that $r_M(I) = Im_M(e)$ with e an i-regular idempotent of $End_R(M)$, as M is an i-Baer semimodule. Let us prove that $r_S(I) = Im_S(e) = eS$.

$eS \subseteq r_S(I)$ comes out of what follows : $r_M(I) = Im_M(e) = eM$. Since $eM \subseteq Im_M(e)$, then $I(eM) \subseteq I(Im_M(e)) = 0 \Rightarrow Ie(M) = I(eM) = 0$. Thus $Ie =$

$0 \Leftrightarrow e \in r_S(I) \Leftrightarrow eS \subseteq r_S(I)$. It remains to prove the reverse inclusion, $r_S(I) \subseteq eS$. Take $\varphi \in r_S(I)$ and $m \in M$, and then $I\varphi = 0 \Leftrightarrow I\varphi(M) = I(\varphi M) = 0 \Leftrightarrow \varphi(M) \subseteq r_M(I)$, since $r_M(I) = eM$, then $\varphi(M) \subseteq eM$ and by lemma 8.1 $e\varphi = \varphi \Leftrightarrow \varphi \in eS$. Therefore $r_S(I) = eS$, as desired.

We generalize the notions of retractability and quasi-retractability. Examples of retractable or quasi-retractable modules are in [22]. Baer semimodules are quasi-retractable.

Definition 8.9 Let M be an R-semimodule:

i M is said to be retractable of type 1 if for every nonzero subsemimodule N of M, there exists a nonzero endomorphism φ such that $\varphi(M) \subseteq N$.
ii M is said to be retractable of type 2 if for every nonzero subsemimodule N of M, there exists a nonzero endomorphism φ such that $Im\varphi \subseteq N$.
iii M is said to be quasi-retractable of type 1 if for every nonzero $I \leqslant_S S$, there exists nonzero φ such that $\varphi(M) \subseteq r_M(I))$.
Equivalently : $\forall 0 \neq I \leqslant_S S, [r_M(I) \neq 0] \Rightarrow [r_S(I) \neq 0]$.
iv M is said to be quasi-retractable of type 2 if for every nonzero $I \leqslant_S S$, there exists a nonzero φ such that $Im_M(\varphi) \subseteq r_M(I)$.

Remark 8.2 It is clear that $ii \Rightarrow i$, and if M is subtractive, then $ii \Leftrightarrow i$.
It is obvious that $iii \Leftrightarrow iv$ since an annihilator is subtractive.
$iv \Rightarrow iii$.

Lemma 8.3 *Let M be an R-semimodule:*

i If M is retractable, then M is quasi-retractable.
ii If M is Baer, then M is quasi-retractable.

Proof

i Let us prove that: $[\forall I \leqslant S]\ [r_M(I) \neq 0 \Rightarrow r_S(I) \neq 0]$. Since M is retractable and $r_M(I)$ is a nonzero subsemimodule of M, there exists nonzero $\varphi \in S$, such that $Im_M(\varphi) \subseteq r_M(I)$. Since $[\varphi(M) \subseteq Im_M(\varphi) \Rightarrow I(\varphi(M)) \subseteq I(Im_M(\varphi))] \Rightarrow [I\varphi(M) \subseteq 0]$, then $I\varphi = 0$. Hence $0 \neq \varphi \in r_S(I)$, as desired.
ii Let us prove that $r_S(I) \neq 0$ is nonzero for every nonzero subset $I \subseteq S$. Since M is a Baer semimodule, then $r_M(I) = Im_M(e)$ for some nonzero idempotent e of S. We pretend that $r_S(I)$ contains e. We have $r_M(I) = Im_M(e) \Rightarrow I(Im_M(e)) = 0$, and since $eM \subseteq Im_M(e)$, then $I(eM) = I(eM) = 0$. Therefore $Ie = 0 \Leftrightarrow e \in r_S(I)$, so $r_S(I) \neq 0$ is nonzero, as sought.

Note that quasi-retractability generalizes retractability.

Let M be a R-semimodule.

We introduce a condition on the intersection of kernels of every subset of $S = End(M)$, denoted (CK), and a condition on the intersection of annihilators in $End(M)$, called (CA). These two conditions, along with quasi-retractability of M, are needed to characterize some class of Baer semimodules by their endomorphism semiring.

Definition 8.10 Let M a R-semimodule:

a M has (CK) condition means every intersection of kernels of $\varphi \in End(M)$ is nonzero.
Equivalently: $\forall I \subseteq S, r_M(I) \neq 0$, in other words $\bigcap_{\varphi \in I} \ker(\varphi) \neq 0$.
b M has (CA) condition means: For $I \subseteq S$,
if there exists an idempotent $e \in S$ such that $r_S(I) = Im_S(e)$, then $r_S(I) \bigcap r_S(e) = 0$.

It is remarkable that $(CK) \Rightarrow non(CA)$ is equivalent to the quasi-retractability of the semimodule in question. If M satisfies (CK), while (CA) is true, then M is not quasi-retractable.

For examples of semimodules M with (CA) condition, take Baer semimodules, which are always quasi-retractable. Every semimodule M does not have (CK) condition. Take $I = \{p, p^{\perp}\}$, and then $\ker(p) \bigcap \ker(p^{\perp}) = 0$, where p and $p^{\perp}$ are orthogonal complements endomorphisms of a semimodule.

Now we can fully characterize Baer semimodules with (CA) condition, using the notion of quasi-retraceability.

Theorem 8.2 *Let M be a semimodule with (CA) condition. Then the following statements are equivalent:*

i M is an i-Baer semimodule.
ii $S = End_R(M)$ is an i-Baer semiring and M is quasi-retractable.

Proof $i, \Rightarrow. ii$ The first part is achieved by Theorem 8.1 and Lemma 8.3.

i $\Leftarrow$ ii Let $I \subseteq End(M)$ a subset.

We wish to show : $r_M(I)$ is generated by an i-idempotent e of S.

We know $r_S(I) = Im_S(e)$ for some i-idempotent of e of S, since S is an i-Baer semiring.

Let us show : $r_M(I) = Im_M(e) = eM$. First we show $Im_M(e) = eM \subseteq r_M(I)$.

From $eS \subseteq Im_S(e)$ and $r_M(I) = Im_S(e) = eS \Rightarrow I(eS) = 0$, we deduce that $Ie(S) = I(eS) = 0$, and thus $Ie = 0$. Then $Ie = 0 \Rightarrow \forall \varphi \in I, \varphi e(M) = \varphi(eM) = \varphi(Im_M(e)) = 0$, and thus $Im_M(e) \subseteq r_M(I)$.

Conversely, let us show that $r_M(I) = Im_M(e)$. Assuming the opposite means that $Im_M(e) \subsetneq r_M(I)$, which implies that there exists nonzero $x \in r_M(I)$, in other words, M has (CK) condition by the alleged assumption. But by hypothesis M already has (CA) condition, and thus M is not quasi-retractable, contradicting the hypothesis of quasi-retractability of M. Hence we must admit that there does not exist nonzero $x \in r_M(I)$ such that x is not in eM. Therefore $r_M(I) = eM = Im_M(e)$, so that M is an i-Baer semimodule.

Example 8.8 Let us prove that the Rickart semiring $(R = \{0, 1, 2\}, \text{gcd}, \text{lcm}, 0, 1)$ in Example 8.4 (in part c.) is a R-Baer semimodule R which has (CA) condition and satisfies the theorem.

The multiplication by the elements of R gives all the elements of $S = End(R)$, which is summarized in the first table. φ_x denotes the multiplication by the element x. The last two give the laws in $End(R)$.

	φ_0	φ_1	φ_2
0	0	0	0
1	0	1	2
2	0	2	2

$+$	φ_0	φ_1	φ_2
φ_0	φ_0	φ_1	φ_2
φ_1	φ_1	φ_1	φ_1
φ_2	φ_2	φ_1	φ_2

$\circ$	φ_0	φ_1	φ_2
φ_0	φ_0	φ_0	φ_0
φ_1	φ_0	φ_1	φ_2
φ_2	φ_0	φ_2	φ_2

The Rickart R-semimodule R is a Baer one, as $End(M)$ is a semidomain:

- Clearly from table 1, M has not (CK), as all kernels are trivial.
- $End(M)$ is Baer, as it is zero divisor free and M has (CA) condition for the same reason.

Example 8.9 Let us prove that the Rickart semimodule $M = (\{0, 1, 2\}, \max, 0)$ on the boolean $R = \{0, 1\}$ semiring endowed with operation as in Example 8.1 is a R-Baer semimodule without (CK) condition and satisfies the theorem. $S = End(R)$ is defined in the following table.

	φ_0	φ_1	φ_2	φ_3	φ_4	φ_5
0	0	0	0	0	0	0
1	0	1	1	2	0	0
2	0	2	1	2	1	2

Let us prove that M is a Baer R-semimodule R:

- Table 1 shows that $I = \{\varphi_4, \varphi_5\} \subseteq End_R(M)$ is the only subset such that $\cap_{\varphi\in I} \ker\varphi \neq \{0\}$. Indeed $\cap_{\varphi\in I} \ker(\varphi = \{0, 1\}$; otherwise $\cap_{\varphi\in I} \ker(\varphi) \neq \{0\}$.
- We can verify that $r_S(I) = \{\varphi_0, \varphi_2, \varphi_4\} = Im(\varphi_2)$, with φ_2 an idempotent and $r_S(\varphi_2) = \{\varphi_0\}$, so that $r_S(I) \cap r_S(e) = \{\varphi_0\}$. Then M is quasi-retractable. Therefore M is an i-Baer semimodule, by Theorem 8.2.

Links Between Rickart and Baer Properties

Rickart Semimodules Versus Baer Semimodules

We first generalize the notion of a direct summand and then introduce intersection properties.

Definition 8.11 Let R be a semiring and I an ideal of R:

- I is a left w-Id-ideal (for weakly idempotent ideal), if $I = \{a \in R,\ a \equiv_{eR} 0\} = Im_R(e)$, where e is an idempotent of R. If e is i-regular, then $I = e(R)$ is said to be a left i-Id ideal. Analogously we define right w-Id-ideal and right i-Id ideal.
- I is said to be a w-Id (i-Id)ideal, if I is both left and right w-Id (i-Id).

Definition 8.12 Let R be a semiring:

1. R has the w-Id IP property, if for every two idempotents $e_1, e_2 \in S$, there exists an idempotent $e \in S$ such that $Im(e_1) \cap Im(e_2) = Im(e)$.

   If the involved idempotents are i-regular, R has i-Id IP.

   Equivalently, the intersection of every finite subset of w-Id-ideals is a w-Id ideal.

   The strong version is:
2. R has the w-Id SIP property (i-Id SIP property), if the intersection of every subset of w-Id-SIP (i-Id-SIP) is a w-Id ideal (i-Id ideal).

The next example of semiring has w-Id-SIP property.

Example 8.10 Let $M = (\{0, 1, 2\}, \max, 0)$ and $R = (\mathbb{N}, +, \times, 1, 0)$, with scalar multiplication defined as follows: $\forall (m, r) \in M \times R\, mr = m$. Then M is a R-semimodule. The table defining $End_R(M)$ does show that M has w-Id IP property.

	φ_0	φ_1	φ_2	φ_3	φ_4	φ_5
0	0	0	0	0	0	0
1	0	1	1	2	0	0
2	0	2	1	2	1	2

The possible intersections of w-Id subsemimodules $\{0\}$, $\{0, 1\}$, and $\{0, 2\}$ are w-Id subsemimodules.

The next example of semiring has not the w-id IP property.

Definition 8.13 Let M be a R-semimodule:

1. A subsemimodule N of M such that $N = Im_M(e)$ for some idempotent $e \in End(M)$ is called a weak-Id subsemimodule (briefly denoted w-Id subsemimodule).
2. If $N = eM$, with $e^2 = e$, we say that N is an i-Id subsemimodule.

In both cases, e is a generating idempotent (or generator) of eM or $Im_M(e)$. It may not be unique. If e is i-regular, then w-Id subsemimodule satisfies both definitions.

Note that for a ring, the w-id ideals are the direct summands of the ring, and the w-id subsemimodules of a module coincide with the direct summands of the module.

If we assume that e is not i-regular, then w-Id ideal (subsemimodule) is not a direct summand as it is not subtractive by Remark 2.2. in [1].

Definition 8.14 Let M be a R-semimodule:

1. M has the w-Id IP property, if for every two idempotents $e_1, e_2 \in S$, there exists an idempotent $e \in S$ such that $Im(e_1) \cap Im(e_2) = Im(e)$.

   If the involved idempotents are i-regular, M has i-Id IP.

Equivalently, the intersection of every finite subset of w-Id subsemimodules is a w-Id subsemimodule.

The strong version is:

2. M has the w-Id SIP property (i-Id SIP property), if the intersection of every subset of w-Id subsemimodules (i-Id-SIP property) is a w-Id subsemimodule (i-Id SIP property).

Example 8.11 The semiring $(R = \begin{pmatrix} \mathbb{N} & \mathbb{N} \\ 0 & \mathbb{N} \end{pmatrix}, +, \times, 0_R, 1_R)$ of the Example 8.2 has not w-id IP. Indeed $e_{11}(R) \cap e_p(R) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mathbb{N} \\ 0 & 0 \end{pmatrix} \neq Im_R(e)$, and then R has not w-Id property. Meanwhile, the Rickart semimodule in Example 8.10 clearly has the property. Then these properties allow us to characterize w-Baer semimodules in terms of w-Rickart semimodules.

Remark 8.3 Example 8.2 provides a R_R semimodule that does not have the w-Id property.

We introduce conditions on the intersection of kernels and annihilators:

1. AIP $(SAIP)$, for annihilators (strong annihilators) intersection property
2. KIP $(SKIP)$, for kernels (strong kernels) intersection property

Definition 8.15 Let R be a semiring and M a R-semimodule, with $I \subseteq End(M)$:

– R has AIP, if $\bigcap_{r \in I} r_R(r)$ is a w-Id ideal of R, for every finite I.
  In others words $\bigcap_{r \in I} r_R(r) = Im_R(e)$, with $e^2 = e \in S$. If I is arbitrary, M has $SAIP$.
– M has KIP, if $\bigcap_{\varphi \in I} \ker(\varphi)$ is a w-Id subsemimodule of M, for every finite I.
  In others words $\bigcap_{\varphi \in I} \ker(\varphi) = Im_M(e)$, with $e^2 = e \in S$. If I is arbitrary, M has $SKIP$.

Remarkably, these two conditions help us to reformulate the Rickart (Baer)property. Let us first begin with a lemma.

Lemma 8.4 *If R is w-Rickart and has w-Id SIP, then R has $SAIP$.*

Proof Since R is w-Rickart, annihilator of every singleton of R is an w-Id ideal of R. Furthermore, by definition 8.12, Id-SIP assumption implies that R has SAIP.

Proposition 8.3 *The following assertions are equivalent:*

1. *R is a w-Baer semiring.*
2. *R has $SAIP$.*
3. *R is a w-Rickart semiring and R has w-Id SIP.*

   Similarly, for a semimodule, the following are equivalent:
4. *M is a w-Baer semimodule.*
5. *M has $SKIP$.*
6. *M is a w-Rickart semimodule and M has w-Id SIP.*

Proof 1) ⇔ 2) is obvious by the definitions. For 3) ⇒ 2 holds out of lemma 8.4, while the converse is obvious.

Similarly, we prove that 4), 5) and 6 are equivalent.

Rickart Semirings Versus Baer Semirings

Recall that L. Small has established a well-known result on rings (Theorem 7.55, [13]) which has been extended to modules theory in theorem 4.5 [15] by Rizvi. We generalized it to some semirings and semimodules as well. We give additional definitions and extend two lemmas similar to that of 4.3 and 4.4 in [15] by Lee and proposition 6.59 in [14] by Lam.

To generalize L. Small's result we need to reformulate a lemma of Lam, using the concepts of w-Id-ideals (not necessarily direct summands) and properties on them. We also assume that w-d Ideals are uniquely generated, which means if a and b are two different idempotents, then $Im_R(a) \neq Im_R(b)$.

We first define an orthogonally finite semiring and a condition based upon it.

Definition 8.16 Let R be a semiring:

R is said to be orthogonally finite, if R has no infinite set of nonzero orthogonal idempotents.
R is said to be having orthogonally finite property (briefly OFP) on w-Id ideals, if:
R is orthogonally finite implies that R satisfies DCC on w-Id ideals.

Lemma 8.5 *Let R be a Rickart semiring with OFP such that w-Id-ideals are uniquely generated. Then the following are equivalent:*

R satisfies ACC on right w-Id ideals.
R satisfies DCC on left w-Id ideals.
R is orthogonally finite.

Proof 1 ⇔ 2. Assume that $Im_R(e_1) \subsetneq Im_R(e_2) \subsetneq ... \subsetneq Im_R(e_n) \subsetneq ....$ is a nonstationary sequel of right w-Id-ideals, as all e_i are pairwise different idempotents of R. Since R is w-Rickart, then $r_R(e_i) = Im_R(f_i)$, where $f_i^2 = f_i$ is in R. Then left annihilators imply the nonstationary following sequel of left w-Id ideals $Im_R(f_1) \subsetneq Im_R(f_2) \subsetneq ... \subsetneq Im_R(f_n) \subsetneq .....$ All f_i are pairwise different idempotents of M too, and thus the inclusion is strict. Therefore the sequels are simultaneously stationary or nonstationary. Hence we get (1) and (2) are equivalent.

3 ⇒ 1 comes out of assumption that R has OFP condition.

1 ⇒ 3. Assume that R is not orthogonally finite, say $(e_1)_{i \in \mathcal{I}}$ is an infinite family of pairwise orthogonal idempotents of R. Take $c_n = e_1+e_2+...e_n$ for $n \geqslant 1$. We can verify that $(c_1)_{i \in \mathcal{I}}$ is a family of pairwise orthogonal idempotents of R. Moreover, $c_{n+1}c_n = (e_1 + e_2 + ... + e_n + e_{n+1})(e_1 + e_2 + ... + e_n) = c_n^2 = c_n$, which is equivalent to $c_n c_{n+1} = c_n \neq c_{n+1}$. Therefore hypothesis ensures that $c_n R \neq c_{n+1} R$ for all n, so (1) fails.

Note that we already have an obvious instance where w-Ricart semirings are w-Baer semirings. To obtain this result, we have assumed that R is w-Rickart semiring with the strong version of w-Id property, say w-Id SIP, unlike what is done in what follows.

Now we can fully characterize w-Baer semirings having OFP condition by w-Rickart semirings with AIP condition. In some way, the next theorem extends L. Small theorem to semirings with assumption that different idempotents of $S = End_R(M)$ generate different w-Id ideals of S.

Theorem 8.3 *Let R be an OFP semiring such that w-Id-ideals are uniquely generated. Then R is a w-Rickart semiring iff R is a w-Baer semiring.*

Proof From the definitions, a w-Baer semiring is always w-Rickart. It suffices to show that R is a w-Baer, using the definition. Let $I = (\varphi_i)_{i\in\mathcal{I}} \subseteq S = End(R)$. Let us show that $r_R(I) = Im_R(e)$ for some idempotent $e \in S$.

Indeed $r_R(I) = \bigcap_{i\in\mathcal{I}} \ker(\varphi_i)$, and by Rickart assumption M has AIP, then for $i \in \mathcal{I}$, there exists $e_i^2 = e_i$ in S such that $\ker(\varphi_i) = Im_R(e_i)$, and thus $r_R(I) = \bigcap_{i\in\mathcal{I}} Im_R(e_i)$. Since the idempotent orthogonally finiteness of S amounts to S has DCC by lemma 8.5, we pretend that $\bigcap_{i\in\mathcal{I}} Im_R(e_i)$ is a w-Id ideal. Let us assume the opposite and show that we can build a nonstationary sequel. Note that the w-Id ideal hypothesis ensures that every finite intersection of $Im_R(e_i)$ is a w-Id ideal one. Hence $Im_R(e_1) \supsetneq Im_R(e_1) \cap Im_R(e_2) \supsetneq Im_R(e_1) \supsetneq Im_R(e_2) \cap Im_R(e_3) \supsetneq \ldots$ is a nonstationary sequel which is a contradiction. Therefore the last sequel should end at some index k. Hence $r_R(I) = \bigcap_{i\in\mathcal{I}} \ker(\varphi_i) = Im_R(e_k)$.

Corollary 8.1 *Let R be an orthogonally finite semiring with w-Id IP property such that w-Id ideals are uniquely generated. Then the R-semimodule R is w-Rickart if and only R is w-Baer R-semimodule R.*

Example 8.12 Let $\mathbf{B}^\infty$ be as in Example 8.5. S is i-Rickart semiring.

Note that every sequence $a = (a_n)_{n=1}^\infty$ has a complement $a^\perp = (a_n^\perp)_{n=1}^\infty$ sequence, where $aa^\perp = a^\perp a = 0$ and $a^\perp + a = 1_A$.

Now let us prove that it is not i-Baer, using L. Small theorem. B^∞ has an infinite set of nonzero orthogonal idempotents. Take the infinite set C of all elements of B^∞ having 1 at i-th position and 0 elsewhere. Clearly $C \subseteq B^\infty$ and every two elements of C are orthogonal. Hence B^∞ is not an i-Baer though it is an i-Rickart one.

Example 8.13 Let $R = \{0, 1, a, b\}$ be the semiring equipped with commutative laws as follows: $\forall x \in R, x + x = x + 0 = x; x + 1 = 1$ and $ab = 0x = 0; 1x = x$. We can verify that

$\{0, a\}$ and $\{0, b\}$ are the only subtractive and nontrivial ideals of R. $r_R(a) = \overline{\{0, b\}} = \overline{bR} = \{0, a, b\} = Im_R(b)$ and $r_R(b) = \overline{\{0, a\}} = \overline{bR} = \{0, a, b\} = Im_R(a)$.

Hence R and $End(R)$ are w-Rickart semirings.

Weak-Rickart Semimodules

Let $\varphi \in \hom_R(M, N)$.

Using Bourne congruence relation on N, $\overline{\varphi(M)} = Im_N(\varphi) = \{y \in N \ : \ y \equiv_{\varphi M} 0_N\} = \{y \in N : \ y + \varphi(m_1) = \varphi(m_2)\}$ defines the subtractive closure of $\varphi(M)$.

References

1. Abuhlail, Jawad and Noegraha, Rangga Ganzar *On k-Noetherian and k-Artinian semirings*, journal=arXiv preprint arXiv:1907.06149, 2019;
2. Armendariz, Efraim P. *A note on extensions of Baer and PP-rings*. Journal of the Australian Mathematical Society,18,4,470–473,1974,Cambridge University Press;
3. Bergman, George M. *Hereditary commutative rings and centres of hereditary rings*, Proceedings of the London Mathematical Society,3,2,214–236,1971,Wiley Online Library;
4. Birkenmeier, Gary F and Park, Jae Keol and Rizvi, S Tariq. *Extensions of rings and modules*. 2013,Springer;
5. Clark, W Edwin and others. *Twisted matrix units semigroup algebras*. Duke Mathematical Journal,34,3,Duke University Press;
6. Endo, Shizuo, *Note on PP Rings:(A Supplement to Hattori's Paper)*, Nagoya Mathematical Journal,17,167–170,1960, Cambridge University Press
7. Evans, Michael, *On commutative PP rings* Pacific Journal of mathematics,41,3,687–697,1972,Mathematical Sciences Publishers;
8. Gupta, V and Chaudhari, JN and Kumar, P.: *STRONGLY BAER AND STRONGLY QUASI-BAER SEMIRINGS AND GENERALIZATIONS*. Asian-European Journal of Mathematics,2,04,579–591,2009,World Scientific;
9. Golan, Jonathan S. *Semirings and their Applications*. 2013,Springer Science & Business Media;
10. Hattori, Akira *A foundation of torsion theory for modules over general rings*, Nagoya Mathematical Journal,17,147–158,1960,Cambridge University Press;
11. Juber, Alaa Zakaria *PROPERTIES OF APP-RINGS* 2008, The University of Jordan;
12. Kaplansky, Irving, *Rings of operators*,1968,WA Benjamin;
13. Lam, TY. *A First course in noncommutative rings, GTM 131 (1991)*,Springer-Verlag;
14. Lam, Tsit-Yuen. *Lectures on modules and rings*. 189,2012,Springer Science & Business Media;
15. Lee, Gangyong and Tariq Rizvi, S and Roman, Cosmin S. *Rickart modules* Communications in Algebra,38,11,4005–4027,2010,Taylor & Francis;
16. Lee, Gangyong, *Theory of Rickart modules*, 2010,The Ohio State University;
17. Maeda, Shûichirô and others, *On a ring whose principal right ideals generated by idempotents form a lattice* Journal of Science of the Hiroshima University, Series A (Mathematics, Physics, Chemistry),24,3,509–525,1960,Hiroshima University, Department of Mathematics;
18. Rizvi, S Tariq and Roman, Cosmin S, *Baer and quasi-Baer modules*, Communications in Algebra,32,1,103–123,2004,Taylor & Francis;
19. Small, Lance W, *Semihereditary rings*, Bulletin of the American Mathematical Society,73,5,656–658,1967,American Mathematical Society;
20. Wisbauer, Robert, *Foundations of module and ring theory*, 2018,Routledge;
21. Zelmanowitz, J, *Regular modules*, Transactions of the American Mathematical Society,163,341–355,1972;
22. Khuri, SM, *Endomorphism rings of modules*, Khuri, SM,2004,manuscript;

Chapter 9
Completion Fractions Modules of Filtered Modules over Non-necessarily Commutative Filtered Rings

Abdoulaye Mane, Mohamed Ben Maaouia, and Mamadou Sanghare

Abstract In this chapter, $(A, (I_n)_{n\in\mathbb{N}})$ is a filtered noncommutative ring, S is a saturated multiplicative subset of A satisfying the left Ore conditions, and $(M, (M_n)_{n\in\mathbb{N}})$ is a filtered left A-module.

The main results in this chapter are the following theorems:

- $\widehat{S} = \left\{\widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0} \text{ and } \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S\right\}$, the set of classes of Cauchy sequences in A with values in S that do not converge to 0, is a saturated multiplicative subset of $\widehat{A}$ satisfying the left Ore conditions.
- $\widehat{S}^{-1}\widehat{A}$ is isomorphic to $\widehat{S^{-1}A}$.
- $\widehat{S}^{-1}\widehat{M}$ is a left $\widehat{S^{-1}A}$-module and is isomorphic to $\widehat{S^{-1}M}$.
- $\widehat{S^{-1}(M/N)}$ is isomorphic to $\widehat{S}^{-1}(\widehat{M})/\widehat{S}^{-1}(\widehat{N})$.
- $\widehat{S^{-1}(A/I)}$ is isomorphic to $\widehat{S}^{-1}(\widehat{A})/\widehat{S}^{-1}(\widehat{I})$.

Keywords Ring · Modules · Filtration · Completion ring and modules · Localization · $\cong$ means isomorphics · S-saturated · $\widehat{S}$-saturated · Cauchy sequence · Ore condition · Multiplicative set

---

A. Mane · M. B. Maaouia (✉)
Laboratory of Algebra, Codes And Cryptography Applications (LACCA), UFR-SAT, University Gaston Berger (UGB), Saint-Louis, Senegal
e-mail: mane.abdoulaye1@ugb.edu.sn; mohamed-ben.maaouia@ugb.edu.sn

M. Sanghare
Doctoral School of Mathematics-Computer – UCAD-Sénégal, University Cheikh Anta Diop of Dakar, Dakar, Senegal
e-mail: mamadou.sanghare@ucad.edu.sn

C. T. Gueye et al. (eds.), *Mathematics of Computer Science, Cybersecurity and Artificial Intelligence*, Springer Proceedings in Mathematics & Statistics 468,
https://doi.org/10.1007/978-3-031-66222-5_9

Introduction

In this chapter, the rings are unitary, associative, not necessarily commutative, and the modules are left modules unifiers.

Let G be a group and $\mathcal{T}$ a topology on G. We say that G is a topological group if the topology $\mathcal{T}$ is compatible with the structure of the group G; that is, the application $(x, y) \mapsto x + y$ defined from $G \times G$ to G and the application $x \mapsto -x$ defined from G to G are continuous. Any filtered group $(G, (G_n)_{n\in\mathbb{N}})$ is equipped with a topology compatible with its group structure where the family $\{G_n\}$ constitutes a system of neighborhoods of 0.

A filtered ring (resp., module) is a ring A (resp., module M) equipped with a filtration $(I_n)_{n\in\mathbb{N}}$ (resp., $(M_n)_{n\in\mathbb{N}}$) formed by left or right ideals (resp., submodules of M). Thus we consider that any filtered ring (resp., module) is equipped with the topology associated with its filtration as a group . We are interested in the localization of the completion rings of non-necessarily commutative filtered rings and completion modules of filtered modules on non-necessarily commutative filtered rings. To do this, we use a saturated multiplicative part S of a filtered ring $(A, (I_n)_{n\in\mathbb{N}})$ not necessarily commutative which verifies the conditions of Ore on the left. In this chapter, we mainly use the following references [1–4, 6], and [5].

Thus, we have obtained the results organized as follows:

In section "Definitions and Preliminary Results" we constructed:

- The filtered group $(\widehat{G}, (\widehat{G_n})_{n\in\mathbb{N}})$, which is the completion of the filtered group $(G, (G_n)_{n\in\mathbb{N}})$
- The filtered ring $(\widehat{A}, (\widehat{I_n})_{n\in\mathbb{N}})$, which is the completion of the filtered ring $(A, (I_n)_{n\in\mathbb{N}})$
- The filtered module $(\widehat{M}, (\widehat{M_n})_{n\in\mathbb{N}})$, which is the completion of the filtered module $(M, (M_n)_{n\in\mathbb{N}})$

In section "On the Localization of Completion Modules":

- We have constructed $\widehat{S} = \left\{\widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0} \text{ and } \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S\right\}$, the set of classes of Cauchy sequences in A with values in S that do not converge to 0. This set is a multiplicatively saturated subset of $\widehat{A}$ satisfying the left Ore conditions (see Theorem 9.3).
- We have shown that the ring $\widehat{S}^{-1}\widehat{A}$ is isomorphic to the ring $\widehat{S^{-1}A}$ the completion of the fraction ring $S^{-1}A$ (see Theorem 9.5).
- If $(M, (M_n)_{n\in\mathbb{N}})$ is a filtered left $(A, (I_n)_{n\in\mathbb{N}})$-module, we have shown that the left $\widehat{S}^{-1}\widehat{A}$-module $\widehat{S}^{-1}\widehat{M}$ is isomorphic to the left $\widehat{S^{-1}A}$-module $\widehat{S^{-1}M}$ (see Theorem 9.6).
- If $(M, (M_n)_{n\in\mathbb{N}})$ is a filtered left $(A, (I_n)_{n\in\mathbb{N}})$-module, we have shown that for any S-saturated submodule N of M, $\widehat{N}$ is an $\widehat{S}$-saturated submodule of $\widehat{M}$, and $\widehat{N} = (i^{\widehat{S}}_{\widehat{M}})^{-1}(N')$ and $\widehat{N} = (\phi \circ i^{\widehat{S}}_{\widehat{M}})^{-1}(N')$, where N' is a submodule of $\widehat{S}^{-1}\widehat{M}$.

- If $(M, (M_n)_{n\in\mathbb{N}})$ is a filtered left $(A, (I_n)_{n\in\mathbb{N}})$-module, we have shown that for any S-saturated submodule N of M, we have $\widehat{S^{-1}(M/N)} \cong \widehat{S^{-1}(M)/S^{-1}(N)} \cong \widehat{S}^{-1}(\widehat{M})/\widehat{S}^{-1}(\widehat{N})$ (see Theorem 9.10).
- If $(A, (I_n)_{n\in\mathbb{N}})$ is a filtered ring, we have shown that for any left ideal I of A, $\widehat{S^{-1}(A/I)} \cong \widehat{S^{-1}(A)/S^{-1}(I)} \cong \widehat{S}^{-1}(\widehat{A})/\widehat{S}^{-1}(\widehat{I})$ (see Theorem 9.10).

Definitions and Preliminary Results

Definition 9.1 Let S be a non-empty subset of a ring A. We say that S is multiplicative if:

1. $1 \in S$ and $0 \notin S$.
2. $\forall s, t \in S,\ st \in S$.

Definition 9.2 Let A be a ring, and S a multiplicative subset of A. We say that S is saturated if for all $a, b \in A$ such that $ab \in S$, we have $a \in S$ and $b \in S$.

Definition 9.3 Let A be a ring and S a subset of A. We say that S is invariant if, for every nonzero element $a \in A$, we have $aS = Sa$.

Definition 9.4 Let A be a ring, and S a saturated multiplicative subset of A. We say that S satisfies the left (resp., right) Ore conditions if:

1. S is left permutable (resp., right permutable): For every $(a, s) \in A \times S$, there is $(b, t) \in A \times S$ such that $ta = bs$ (resp., $at = sb$).
2. S is left reversible (resp., right reversible): For every $a \in A$, if there is $s \in S$ such that $as = 0$ (resp., $sa = 0$), then there is $t \in S$ such that $ta = 0$ (resp., $at = 0$).

Definition 9.5 Let N be a submodule of a left A-module M, and S a left-saturated multiplicative subset of A satisfying the left Ore conditions. We say that N is left saturated for S in M if, for all $s \in S$ and $x \in M$ such that $sx \in N$, then $x \in N$. It is also said that N is S-saturated in M.

Topological Group

Definition 9.6 Let G be a group. We call a descending (resp., ascending) filtration of group G any sequence $(G_n)_{n\in\mathbb{N}}$ of normal subgroups of G such that:

1. $G_{n+1} \subset G_n$ (resp., $G_n \subset G_{n+1}$) $\forall n \in \mathbb{N}$.
2. $\bigcup_{n\in\mathbb{N}} G_n = G$.

And we say that G is a filtered group by $(G_n)_{n\in\mathbb{N}}$ denoted by $(G, (G_n)_{n\in\mathbb{N}})$.

Theorem 9.1 *Let* $(G, (G_n)_{n\in\mathbb{N}})$ *be a filtered group. Then,* G *is equipped with a topological group structure whose open sets are the subsets* $K \subseteq G$ *such that for all* $x \in K$*, there is* $n \in \mathbb{N}$ *such that* $x + G_n \subseteq K$*.*

Proof Suppose that the filtration $(G_n)_{n\in\mathbb{N}}$ ~~is~~ ascending. Let $\mathcal{T} = \{K \subset G | \forall x \in K, \exists n \in \mathbb{N}, x + G_n \subset K\}$ be the set associated with the filtration $(G_n)_{n\in\mathbb{N}}$:

1. Let us show that $\mathcal{T}$ endows G with a topological structure:

   (a) G and $\emptyset \in \mathcal{T}$. Indeed, let $+ : \begin{array}{l} G \times G \longrightarrow G \\ (x, y) \mapsto x + y \end{array}$ be an application. Then for all $(x, n) \in G \times \mathbb{N}$, $+$ restricted to $\{x\} \times G_n$ is an application such that $+(\{x\} \times G_n) = \{x\} + G_n = \{x + y, \forall y \in G_n\}$. In particular, replacing $\{x\}$ by $\emptyset$, there is $n_0 \in \mathbb{N}$ such that
   $+(\emptyset \times G_{n_0}) = \emptyset + G_{n_0} = \emptyset \subset \emptyset$, and hence $\emptyset \in \mathcal{T}$.
   For G, we have for all $n \in \mathbb{N}$, $0 + G_n \subset G = \underset{n\in\mathbb{N}}{\cup} G_n \in \mathcal{T}$.

   (b) Let $\{U_i\}_{i\in I} \in \mathcal{T}$ be any family. Then, for all $x \in \underset{n\in I}{\cup} U_n$, there is $i_0 \in I$ such that $x \in U_{i_0}$. Now $U_{i_0} \in \mathcal{T}$, so there is $n_0 \in \mathbb{N}$ such that $x + G_{n_0} \subset U_{i_0} \subset \underset{n\in I}{\cup} U_i$, which proves that $\underset{n\in I}{\cup} U_i \in \mathcal{T}$.

   (c) Let $U_1, \cdots, U_N \in \mathcal{T}$, for any given $N \in \mathbb{N}$. Then, for all $x \in \bigcap_{i=1}^{n} U_i$ implies $x \in U_i, i \in \{1, \cdots, N\}$. Now $U_i \in \mathcal{T}$, so there is $n \in \mathbb{N}$ such that $x + G_n \subset U_i, i \in \{1, \cdots, N\}$; hence, $x + G_n \subset \bigcap_{i=1}^{n} U_i$, which proves that $\bigcap_{i=1}^{n} U_i \in \mathcal{T}$.

2. Let us show that $\mathcal{T}$ is compatible with the group structure of G.
   Just show that the applications defined by
   $- : G \longrightarrow G$ such that $x \longmapsto -x$ and $+ : G \times G \longrightarrow G$ such that $(x, y) \longmapsto x + y$,
   where $G \times G$ is endowed with the product topology, are continuous:

   (a) Let $x \in G$ such that there is $U \in \mathcal{T}$ such that $x \in -^{-1}U$. Then, $x \in -^{-1}U$ implies that $-x \in U$, so there is $n \in \mathbb{N}$ such that $-x + G_n \subset U$, which implies $-(x - G_n) \subset U$. Furthermore, G_n is a subgroup, so $-G_n = G_n$. Thus, $-(x - G_n) = -(x + G_n) \subset U$. It follows that $x + G_n \subset -^{-1}(U)$, and therefore the map $-$ is continuous.

   (b) Consider $G \times G$ equipped with the product topology defined by:

   $$\mathcal{T} \times \mathcal{T} = \{U \times V, \forall\, U, V \in \mathcal{T}\}.$$

   Let $(x, y) \in G \times G$ such that there is $U \in \mathcal{T}$ such that $(x, y) \in +^{-1}(U)$.
   Then, $(x, y) \in +^{-1}(U) \Rightarrow +(x, y) = x + y \in U$
   $\Rightarrow \exists n \in \mathbb{N}, x + y + G_n \subset U$

$\Rightarrow \exists n \in \mathbb{N}$, $(x + G_n) + (y + G_n) \subset U$, since G_n is a normal subgroup. Then, $(x + G_n) \times (y + G_n) \subset G \times G$ is a neighborhood of (x, y) for the product topology, and $+((x+G_n)\times(y+G_n)) = (x+G_n)+(y+G_n) \subset U$. Therefore, $(x + G_n) \times (y + G_n) \subset +^{-1}(U)$, showing that the map $+$ is continuous.

In duality, if $(G_n)_{n\in\mathbb{N}}$ is decreasing, it is shown in the same manner that $\mathcal{T}$ forms a group topology.

Corollary 9.1 *Let $(G, (G_n)_{n\in\mathbb{N}})$ be a filtered topological group. Then, for every $n \in \mathbb{N}$, G_n is an open set in G.*

Completion Modules of Filtered Rings and Filtered Modules

Let $S(G)$ be the set of all sequences with values in G such that $\mathcal{S}(G)$, equipped with the operation induced by G, is a group.

Definition 9.7 Let $(G, (G_n)_{n\in\mathbb{N}})$ be a filtered group equipped with the topology associated with the filtration $(G_n)_{n\in\mathbb{N}}$. We say that a sequence $(x_n) \in \mathcal{S}(G)$ converges if there is $x \in G$ such that for all G_n, there is $n_0 \in \mathbb{N}$ such that for all $m \geq n_0$, we have $x - x_m \in G_n$.

Definition and Notation 1 Let $(G, (G_n)_{n\in\mathbb{N}})$ be a filtered group equipped with the topology associated with the filtration $(G_n)_{n\in\mathbb{N}}$. A Cauchy sequence is any sequence $(x_n) \in \mathcal{S}(G)$ such that: $\forall G_n, \exists n_0 \in \mathbb{N}, \forall p, q \geq n_0 \Rightarrow x_p - x_q \in G_n$. We denote by $\mathcal{C}(G)$ the set of Cauchy sequences.

Remark 9.1 According to the definitions above, every convergent sequence is Cauchy.

Proposition 9.1 *Let $(G, (G_n)_{n\in\mathbb{N}})$ be a filtered group equipped with the topology associated with the filtration $(G_n)_{n\in\mathbb{N}}$. Then, the set of Cauchy sequences $\mathcal{C}(G)$ forms a subgroup of $\mathcal{S}(G)$.*

Definition 9.8 Let $(G, (G_n)_{n\in\mathbb{N}})$ be a filtered group (resp., $(A, (I_n)_{n\in\mathbb{N}})$, a filtered ring (resp., $(M, (M_n)_{n\in\mathbb{N}})$ a filtered module), B a non-empty subset of G (resp., A, resp., M), and (u_n) a sequence with values in G (resp., A, resp., M). We say that (u_n) has support in B if all the terms of (u_n) are elements of B except, perhaps, for a finite number of terms (that is, $\exists n_0 \in \mathbb{N}, \forall n \geq n_0, u_n \in B$).

Theorem and Notations 1 *Let $(G, (G_n)_{n\in\mathbb{N}})$ be a filtered group equipped with the topology associated with $(G_n)_{n\in\mathbb{N}}$. Then, the binary relation $\mathcal{R}$ on $\mathcal{C}(G)$ defined by*

$$\forall (x_n), (y_n) \in \mathcal{C}(G), (x_n)\mathcal{R}(y_n) \Leftrightarrow (x_n) - (y_n) = (x_n - y_n) \to 0$$

is an equivalence relation. The quotient set $\mathcal{C}(G)/\mathcal{R} = \left\{\widehat{(x_n)} \mid (x_n) \in \mathcal{C}(G)\right\}$ is denoted $\widehat{G}$, where $\widehat{(x_n)}$ is the equivalence class of (x_n).

Proposition 9.2 *Let $(G, (G_n)_{n\in\mathbb{N}})$ be a filtered group equipped with the topology associated with the filtration $(G_n)_{n\in\mathbb{N}}$. For any sequence $(x_n) \in \mathcal{C}(G)$ converging to $x \in A$, then $\widehat{(x_n)} = \widehat{x}$, where $\widehat{x} = (x, \cdots, x)$ is the class of the constant sequence (x_n), for all $n \in \mathbb{N}$, such that $x_n = x$.*

Theorem and Definition 1 *Let $(G, (G_n)_{n\in\mathbb{N}})$ be a filtered group equipped with the topology associated with the filtration $(G_n)_{n\in\mathbb{N}}$, and let $\widehat{+}$ be the map defined by*

$$\begin{array}{rccc} \widehat{+}: & \widehat{G}\times\widehat{G} & \to & \widehat{G} \\ & (\widehat{(x_n)}, \widehat{(y_n)}) & \mapsto & \widehat{(x_n)}\widehat{+}\widehat{(y_n)} = \widehat{(x_n + y_n)} \end{array}.$$

Then, $(\widehat{G}, \widehat{+})$ is a group called the completion of G.

Proof Let us show that $\widehat{+}$ is well defined.

Let $(\widehat{(x_n)}, \widehat{(y_n)})$ and $(\widehat{(x'_n)}, \widehat{(y'_n)}) \in \widehat{G} \times \widehat{G}$ such that $(\widehat{(x_n)}, \widehat{(y_n)}) = (\widehat{(x'_n)}, \widehat{(y'_n)})$. We want to demonstrate that $\widehat{(x_n)}\widehat{+}\widehat{(y_n)} = \widehat{(x'_n)}\widehat{+}\widehat{(y'_n)}$.

By assumption, $(\widehat{(x_n)}, \widehat{(y_n)}) = (\widehat{(x'_n)}, \widehat{(y'_n)})$. Therefore: $\begin{cases} \widehat{(x_n)} = \widehat{(x'_n)} \\ \widehat{(y_n)} = \widehat{(y'_n)} \end{cases} \Rightarrow$
$\begin{cases} (x_n) - (x'_n) = (x_n - x'_n) \to 0 \\ (y_n) - (y'_n) = (y_n - y'_n) \to 0 \end{cases} \Rightarrow$
$(x_n) + (y_n) - (x'_n) - (y'_n) \to 0 \Rightarrow \widehat{(x_n)}\widehat{+}\widehat{(y_n)} = \widehat{(x'_n)}\widehat{+}\widehat{(y'_n)}$. Therefore, $\widehat{+}$ is an internal composition law on $\widehat{G}$.

The properties of the operation that define the group are evident. Therefore, $(\widehat{G}, \widehat{+})$ is a group.

Lemma 9.1 *Let $(G, (G_n)_{n\in\mathbb{N}})$ be a filtered group. Then, every subgroup H of G is a filtered group.*

Proof Let $H_n = H \cap G_n$ for all $n \in \mathbb{N}$. Then, for each $n \in \mathbb{N}$, the sets H_n are subgroups of H as groups. Since $(G_n)_{n\in\mathbb{N}}$ is a filtration, $(H, (H_n)_{n\in\mathbb{N}})$ forms a filtered subgroup H.

Definitions 9.1 *Let $(G, (G_n)_{n\in\mathbb{N}})$ be a filtered group, and let H be a subgroup of G. Then, $(H_n)_{n\in\mathbb{N}}$ is called the induced filtration of the filtration $(G_n)_{n\in\mathbb{N}}$ of G on H, where $H_n = H \cap G_n$ for all $n \in \mathbb{N}$.*

Proposition 9.3 *Let $(G, (G_n)_{n\in\mathbb{N}})$ be a filtered group, and let H be a subgroup of G. Then, $\widehat{H}$, the completion of H equipped with the topology associated with the induced filtration $(H_n)_{n\in\mathbb{N}}$ from the filtration $(G_n)_{n\in\mathbb{N}}$ of G, is isomorphic to a subgroup of $\widehat{G}$.*

Proof Let $\mathcal{H} = \left\{\widehat{(x_n)} \in \widehat{G} \mid x_n \in H, \forall n \in \mathbb{N}\right\}$:

1. Let us show that $\mathcal{H}$ is a subgroup of $\widehat{G}$:

(a) We have: $\mathcal{H} \subset \widehat{G}$ by definition, and $\mathcal{H} \neq \emptyset$. Indeed, $\widehat{0} \in \mathcal{H}$ because $0 \in H$.
(b) Let $\widehat{(x_n)}, \widehat{(y_n)} \in \mathcal{H}$, and then $x_n, y_n \in H, \forall n \in \mathbb{N} \Rightarrow x_n + y_n \in H, \forall n \in \mathbb{N} \Rightarrow \widehat{(x_n)}\widehat{+}\widehat{(y_n)} \in \mathcal{H}$.
(c) Let $\widehat{(x_n)} \in \mathcal{H}$, and then $x_n \in H, \forall n \in \mathbb{N} \Rightarrow -x_n \in H, \forall n \in \mathbb{N} \Rightarrow \widehat{-(x_n)} \in \mathcal{H}$.

Therefore, $\mathcal{H}$ is a subgroup of $\widehat{G}$.

2. Let ψ be the natural correspondence defined by:

$$\begin{array}{rcl} \psi \ \widehat{H} & \rightarrow & \mathcal{H} \\ \widehat{(x_n)} & \mapsto & \psi(\widehat{(x_n)}) = \widehat{(x_n)}. \end{array}$$

(a) ψ is indeed a function.
(b) ψ is a group homomorphism.
(c) ψ is bijective. Indeed:
$*$ For all $\widehat{(x_n)} \in \mathcal{H}$, we have: $(x_n) \in \mathcal{C}(H) \Rightarrow -(x_n) \in \mathcal{C}(H) \Rightarrow (x_n) - (x_n) \rightarrow 0 \Rightarrow \widehat{(x_n)} = \widehat{-(x_n)} \in \widehat{H}$ according to Theorem 1. Thus, for all $\widehat{(x_n)} \in \mathcal{H}$, there is $\widehat{(y_n)} \in \widehat{H}$ (by taking $\widehat{(y_n)} = \widehat{-(x_n)}$) such that $\psi(\widehat{(y_n)}) = \widehat{(x_n)}$, hence the surjectivity of ψ.
$*$ Let $\widehat{(x_n)} \in Ker(\psi)$; then $\psi(\widehat{(x_n)}) = \widehat{(x_n)} = \widehat{0}$. Thus, $\text{Ker}(\psi) = \left\{\widehat{0}_{\widehat{H}}\right\}$, indicating that ψ is injective.

We conclude that ψ is bijective.

Remark 9.2 In the following, we consider that $\widehat{H}$ is a subgroup of $\widehat{G}$.

Proposition 9.4 *Let $(G, (G_n)_{n\in\mathbb{N}})$ be a filtered group and $\widehat{G}$ its completion. Then, $(\widehat{G_n})_{n\in\mathbb{N}}$ is a filtration of the group $\widehat{G}$.*

Proof See 9.1 and 9.3.

Corollary 9.2 *Let $(G, (G_n)_{n\in\mathbb{N}})$ be a filtered group and $\widehat{G}$ its completion. Then, $\widehat{G}$ is equipped with a topological group structure associated with the filtration $(\widehat{G_n})_{n\in\mathbb{N}}$.*

Theorem and Definition 2 *Let $(A, (I_n)_{n\in\mathbb{N}})$ be a filtered ring where $(I_n)_{n\in\mathbb{N}}$ is a sequence of left (resp., right) ideals, and let $\widehat{\times}$ be the correspondence defined by:*

$$\begin{array}{rrcl} \widehat{\times}: & \widehat{A} \times \widehat{A} & \rightarrow & \widehat{A} \\ & (\widehat{(x_n)}, \widehat{(y_n)}) & \mapsto & \widehat{(x_n)}\widehat{\times}\widehat{(y_n)} = \widehat{(x_n \times y_n)} \end{array}.$$

Then, $(\widehat{A}, \widehat{+}, \widehat{\times})$ is a ring called the completion of A:

Proof

1. According to Theorem 1, $(\widehat{A}, \widehat{+})$ is a group, and it is abelian because $(A, +)$ is abelian.
2. Let $\widehat{(a_n)}, \widehat{(a'_n)}, \widehat{(b_n)}, \widehat{(b'_n)} \in \widehat{A}$, and suppose that $(\widehat{(a_n)}, \widehat{(b_n)}) = (\widehat{(a'_n)}, \widehat{(b'_n)})$. Let us show that $\widehat{(a_n)}\widehat{\times}\widehat{(b_n)} = \widehat{(a'_n)}\widehat{\times}\widehat{(b'_n)}$.

   We have : $(\widehat{(a_n)}, \widehat{(b_n)}) = (\widehat{(a'_n)}, \widehat{(b'_n)}) \Rightarrow \begin{cases} \widehat{(a_n)} = \widehat{(a'_n)} \\ \widehat{(b_n)} = \widehat{(b'_n)} \end{cases} \Rightarrow \begin{cases} (a_n - a'_n) \to 0 \\ (b_n - b'_n) \to 0 \end{cases}$

   $\Rightarrow (a_n - a'_n)(b_n - b'_n) \to 0.$

   But $a_n - a'_n, b_n - b'_n \in A$ for all $n \in \mathbb{N}$; thus, $(a_n - a'_n)(b_n - b'_n) = a_n b_n - a'_n b'_n - a_n(b_n - b'_n) - (a_n - a'_n)b'_n$ for all $n \in \mathbb{N}$. Consequently, $(a_n b_n - a'_n b'_n) = (a_n - a'_n)(b_n - b'_n) + (a_n)(b_n - b'_n) + (a_n - a'_n)(b'_n)$. Since $(a_n - a'_n)(b_n - b'_n) \to 0$, $(a_n - a'_n) \to 0$ and $(b_n - b'_n) \to 0$, we have $(a_n b_n - a'_n b'_n) \to 0$. Thus, $\widehat{(a_n b_n)} = \widehat{(a'_n b'_n)} \Leftrightarrow \widehat{(a_n)}\widehat{\times}\widehat{(b_n)} = \widehat{(a'_n)}\widehat{\times}\widehat{(b'_n)}$. Therefore, $\widehat{\times}$ is an internal composition law.

The properties that define the ring $\widehat{A}$ are evident.

Proposition 9.5 *Let A be a ring and I a left (resp., right) ideal of A. Then, $\widehat{I}$ is an ideal of $\widehat{A}$.*

Proof

1. According to Proposition 9.3, $\widehat{I}$ is a subgroup of $(\widehat{A}, \widehat{+})$.
2. Suppose that I is a left ideal.

   Let $\widehat{(a_n)} \in \widehat{A}$ and $\widehat{(x_n)} \in \widehat{I}$. Then, $\widehat{(a_n)}\widehat{\times}\widehat{(x_n)} = \widehat{(a_n x_n)}$. Since $a_n x_n \in I$ for all $n \in \mathbb{N}$, it follows that $\widehat{(a_n)}\widehat{\times}\widehat{(x_n)} = \widehat{(a_n x_n)} \in \widehat{I}$. Thus, $\widehat{I}$ is a left ideal.
3. If I is a right ideal, we can show similarly by taking $\widehat{(a_n)} \in \widehat{A}$ and $\widehat{(x_n)} \in \widehat{I}$, and $\widehat{(x_n)}\widehat{\times}\widehat{(a_n)} = \widehat{(x_n a_n)} \Rightarrow x_n a_n \in I_n$ for all $n \in \mathbb{N}$. Therefore, $\widehat{(x_n a_n)} \in \widehat{I} \Rightarrow \widehat{(x_n)}\widehat{\times}\widehat{(a_n)} \in \widehat{I}$.

Therefore, $\widehat{I}$ is a left (resp., right) ideal of $\widehat{A}$.

Theorem 9.2 *Let $(A, (I_n)_{n\in\mathbb{N}})$ be a filtered ring where $(I_n)_{n\in\mathbb{N}}$ is a sequence of left ideals, and $(M, (M_n)_{n\in\mathbb{N}})$ be a left-filtered A-module where $(M_n)_{n\in\mathbb{N}}$ is a sequence of submodules of M. Let $\widehat{\cdot}$ be the correspondence defined by:*

$$\widehat{\cdot} : \begin{array}{ccc} \widehat{A} \times \widehat{M} & \to & \widehat{M} \\ (\widehat{(a_n)}, \widehat{(m_n)}) & \mapsto & \widehat{(a_n)}\widehat{\cdot}\widehat{(m_n)} = \widehat{(a_n \cdot m_n)} \end{array}.$$

Then, $(\widehat{M}, \widehat{+}, \widehat{\cdot})$ is a left $\widehat{A}$-module.

Proof

1. Let $\widehat{(a_n)}, \widehat{(a'_n)} \in \widehat{A}$ and $\widehat{(m_n)}, \widehat{(m'_n)} \in \widehat{M}$. Suppose $(\widehat{(a_n)}, \widehat{(m_n)}) = (\widehat{(a'_n)}, \widehat{(m'_n)})$, and let us show that $\widehat{(a_n)}\widehat{\cdot}\widehat{(m_n)} = \widehat{(a'_n)}\widehat{\cdot}\widehat{(m'_n)}$.

We have $(\widehat{(a_n)}, \widehat{(m_n)}) = (\widehat{(a'_n)}, \widehat{(m'_n)})$, which means

$$\begin{cases} (a_n - a'_n) \to 0 \\ (m_n - m'_n) \to 0 \end{cases} \Rightarrow (a_n - a'_n)(m_n - m'_n) \to 0.$$

Now, for all $n \in \mathbb{N}$, consider the expression

$$(a_n - a'_n)(m_n - m'_n) = a_n m_n - a_n m'_n - a'_n m_n + a'_n m'_n.$$

This simplifies to

$$a_n m_n - a'_n m'_n = (a_n - a'_n)(m_n - m'_n) + (a_n - a'_n)m'_n + a'_n(m_n - m'_n)$$

$$\Rightarrow (a_n m_n - a'_n m'_n) = (a_n - a'_n)(m_n - m'_n) + (a_n - a'_n)(m'_n) + (a'_n)(m_n - m'_n).$$

Now, since $(a_n - a'_n) \to 0$, $(m_n - m'_n) \to 0$, and $(a_n - a'_n)(m_n - m'_n) \to 0$, we conclude that $(a_n m_n - a'_n m'_n) \to 0$. This implies $\widehat{(a_n)}\widehat{\cdot}\widehat{(m_n)} - \widehat{(a'_n)}\widehat{\cdot}\widehat{(m'_n)} = \widehat{0}$. Thus, $\widehat{(a_n)}\widehat{\cdot}\widehat{(m_n)} = \widehat{(a'_n)}\widehat{\cdot}\widehat{(m'_n)}$, showing that $\widehat{\cdot}$ is an external composition law.

The properties that define the $\widehat{A}$-module $\widehat{M}$ are evident.

Proposition 9.6 *Let $(A, (I_n)_{n\in\mathbb{N}})$ be a filtered ring where $(I_n)_{n\in\mathbb{N}}$ is a sequence of left (resp., right) ideals, and $(M, (M_n)_{n\in\mathbb{N}})$ be a filtered left A-module where $(M_n)_{n\in\mathbb{N}}$ is a sequence of submodules of M and $\widehat{M}$ be its completion. Then, for any submodule N of M, its completion $\widehat{N}$ is a submodule of $\widehat{M}$.*

Proof According to Proposition 9.3, $\widehat{N}$ is a subgroup of $(\widehat{M}, \widehat{+})$. Let $\widehat{(a_n)} \in \widehat{A}$ and $\widehat{(m_n)} \in \widehat{N}$; we want to show that $\widehat{(a_n)}\widehat{\cdot}\widehat{(m_n)} \in \widehat{N}$.

Since $\widehat{(a_n)} \in \widehat{A}$ and $\widehat{(m_n)} \in \widehat{N}$, it follows that $(a_n) \in \mathcal{C}(A)$ and $(m_n) \in \mathcal{C}(N)$. This implies $a_n \in A$ and $m_n \in N$ for all $n \in \mathbb{N}$. Therefore, $(a_n) \cdot (m_n) \in \mathcal{C}(N)$, and thus, $\widehat{(a_n)}\widehat{\cdot}\widehat{(m_n)} = \widehat{(a_n) \cdot (m_n)} \in \widehat{N}$. Consequently, $\widehat{N}$ is a submodule of $\widehat{M}$.

On the Localization of Completion Modules

On the Localization of Completion Rings and Modules

Theorem 9.3 *Let $(A, (I_n)_{n\in\mathbb{N}})$ be a filtered ring where $(I_n)_{n\in\mathbb{N}}$ ~~is~~ a sequence of left (resp., right) ideals, and $\widehat{A}$ be its completion. Let S be a saturated multiplicative subset of A that satisfies left Ore conditions, and let*

$$\widehat{S} = \left\{ \widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0} \; et \; \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S \right\}$$

be the set of classes of Cauchy sequences in A with values in S that do not converge to 0*. Then,* $\widehat{S}$ *is a saturated multiplicative subset of* $\widehat{A}$ *that satisfies the left Ore conditions.*

Proof

1. We have $\widehat{0} \notin \widehat{S}$. Indeed, suppose that $\widehat{0} \in \widehat{S}$, then $0 \in S$, which is absurd.
2. Let the constant sequence $(x_n) = (1)$, it is a convergent sequence, and thus, $\widehat{1} \in \widehat{S}$.
3. Let $\widehat{(x_n)}, \widehat{(y_n)} \in \widehat{S}$. We have $\widehat{(x_n)} \widehat{\times} \widehat{(y_n)} = \widehat{(x_n y_n)}$. Since S is multiplicative, $x_n y_n \in S$ for all $n \in \mathbb{N}$, and thus $\widehat{(x_n y_n)} \in \widehat{S}$. Therefore, $\widehat{(x_n)} \widehat{\times} \widehat{(y_n)} \in \widehat{S}$. This shows that $\widehat{S}$ is multiplicative.
4. Let us show that $\widehat{S}$ is saturated.

   Let $\widehat{(x_n)}, \widehat{(y_n)} \in \widehat{A}$ such that $\widehat{(x_n)} \widehat{\times} \widehat{(y_n)} \in \widehat{S}$. We have $\widehat{(x_n)} \widehat{\times} \widehat{(y_n)} = \widehat{(x_n y_n)} \in \widehat{S}$, implying that $x_n y_n \in S$ for all $n \in \mathbb{N}$. Since S is saturated, it follows that x_n and $y_n \in S$ for all $n \in \mathbb{N}$. Therefore, $\widehat{(x_n)}$ and $\widehat{(y_n)} \in \widehat{S}$. Then $\widehat{S}$ is saturated.
5. Let us show that $\widehat{S}$ satisfies the left Ore conditions:

   a. Let $\widehat{(a_n)} \in \widehat{A}$ and $\widehat{(s_n)} \in \widehat{S}$. Let us demonstrate that there are $\widehat{(b_n)} \in \widehat{A}$ and $\widehat{(t_n)} \in \widehat{S}$ such that $\widehat{(t_n)} \widehat{\times} \widehat{(a_n)} = \widehat{(b_n)} \widehat{\times} \widehat{(s_n)}$.

   We have $\widehat{(a_n)} \in \widehat{A}$ and $\widehat{(s_n)} \in \widehat{S}$, so $a_n \in A$ and $s_n \in S, \forall n \in \mathbb{N}$. Since S is left permutable, there are sequences $(b_n) \in A$ and $(t_n) \in S$ such that $t_n a_n = b_n s_n, \forall n \in \mathbb{N}$. Thus, $\widehat{(t_n a_n)} = \widehat{(b_n s_n)} \Leftrightarrow \widehat{(t_n)} \widehat{\times} \widehat{(a_n)} = \widehat{(b_n)} \widehat{\times} \widehat{(s_n)}$. Therefore, for any $\widehat{(a_n)} \in \widehat{A}$ and $\widehat{(s_n)} \in \widehat{S}$, there are $\widehat{(b_n)}$ and $\widehat{(t_n)}$ such that $\widehat{(t_n)} \widehat{\times} \widehat{(a_n)} = \widehat{(b_n)} \widehat{\times} \widehat{(s_n)}$.

   b. Let $\widehat{(a_n)} \in \widehat{A}$. Suppose there is $\widehat{(s_n)} \in \widehat{S}$ such that $\widehat{(a_n)} \widehat{\times} \widehat{(s_n)} = \widehat{0}$. Let us show that there is $\widehat{(t_n)} \in \widehat{S}$ such that $\widehat{(t_n)} \widehat{\times} \widehat{(a_n)} = \widehat{0}$.

   We have $\widehat{(a_n)} \widehat{\times} \widehat{(s_n)} = \widehat{0} \Rightarrow (a_n) \times (s_n) \to 0 \Rightarrow (a_n s_n) \to 0 \Rightarrow \left(\frac{a_n s_n}{1}\right) \to 0_{S^{-1}A} \Rightarrow \left(\frac{a_n}{1} \times \frac{s_n}{1}\right) = \left(\frac{a_n}{1}\right) \times \left(\frac{s_n}{1}\right) \to 0_{S^{-1}A}$.

   But $\left(\frac{s_n}{1} \times \frac{1}{s_n}\right) = \left(\frac{s_n}{1}\right) \times \left(\frac{1}{s_n}\right) = 1$, and then $\left[\left(\frac{a_n}{1}\right) \times \left(\frac{s_n}{1}\right)\right] \times \left(\frac{1}{s_n}\right) \to 0_{S^{-1}A} \times \left(\frac{1}{s_n}\right) = 0_{S^{-1}A}$
   $\Rightarrow \left(\frac{a_n}{1}\right) \times \left[\left(\frac{s_n}{1}\right) \times \left(\frac{1}{s_n}\right)\right] \to 0_{S^{-1}A} \Rightarrow \left(\frac{a_n}{1}\right) \to 0_{S^{-1}A} \Rightarrow (a_n) \to 0 \Rightarrow \widehat{(a_n)} = \widehat{0}$.

   Therefore, there is $\widehat{(t_n)} \in \widehat{S}$ (by taking $\widehat{(t_n)} = \widehat{(s_n)}$) such that $\widehat{(t_n)} \widehat{\times} \widehat{(a_n)} = \widehat{0}$.

Theorem 9.4 *Let* $(A, (I_n)_{n\in\mathbb{N}})$ *be a filtered ring where* $(I_n)_{n\in\mathbb{N}}$ *is a sequence of left (resp., right) ideals, and* $\widehat{A}$ *be its completion. Let S be a saturated multiplicative subset invariant of A and let*

$$\widehat{S} = \left\{\widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0}\ et\ \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S\right\}$$

be the set of classes of Cauchy sequences in A with values in S that do not converge to 0*. Then,* $\widehat{S}$ *is a saturated multiplicative subset of* $\widehat{A}$ *that satisfies the left Ore conditions.*

Proof According to Theorem 9.3, $\widehat{S}$ is multiplicative, saturated, and left permutable. Thus, it suffices to show that it is left reversible.

Let $\widehat{(a_n)} \in \widehat{A}$. Suppose there is $\widehat{(s_n)} \in \widehat{S}$ such that $\widehat{(a_n)}\widehat{\times}\widehat{(s_n)} = \widehat{0}$, and let us show that there is $\widehat{(t_n)} \in \widehat{S}$ such that $\widehat{(t_n)}\widehat{\times}\widehat{(a_n)} = \widehat{0}$.

We have $\widehat{(a_n)}\widehat{\times}\widehat{(s_n)} = \widehat{0} \Rightarrow (a_n) \times (s_n) = (a_n s_n) \to 0$. Since $a_n s_n \in a_n S = S a_n$ for all $n \in \mathbb{N}$ (as S is invariant), for every $n \in \mathbb{N}$, there is $t_n \in S$ such that $a_n s_n = t_n a_n$ for all $n \in \mathbb{N}$. Let (t_n) be the sequence with terms t_n, and then $(a_n s_n) = (t_n a_n) \to 0$, implying $\widehat{(t_n)}\widehat{\times}\widehat{(a_n)} = \widehat{0}$. Therefore, $\widehat{S}$ is left reversible.

Corollary 9.3 *Let* $(A, (I_n)_{n\in\mathbb{N}})$ *be a filtered ring where* $(I_n)_{n\in\mathbb{N}}$ *is a sequence of left (resp., right) ideals, and* $\widehat{A}$ *be its completion. Let S be a central subset of A and let*

$$\widehat{S} = \left\{\widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0}\ et\ \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S\right\}$$

be the set of classes of Cauchy sequences in A with values in S that do not converge to 0*. Then,* $\widehat{S}$ *is a saturated multiplicative subset of* $\widehat{A}$ *that satisfies the left Ore conditions.*

Lemma 9.2 *Let* $(A, (I_n)_{n\in\mathbb{N}})$ *be a filtered ring where* $(I_n)_{n\in\mathbb{N}}$ *is a sequence of left (resp., right) ideals. Let S be a saturated multiplicative subset of A that satisfies left Ore conditions. Then,* $(S^{-1}I_n)_{n\in\mathbb{N}}$ *is a filtration of* $S^{-1}A$*.*

Theorem 9.5 *Let* $(A, (I_n)_{n\in\mathbb{N}})$ *be a filtered ring where* $(I_n)_{n\in\mathbb{N}}$ *is a sequence of left (resp., right) ideals, and* $\widehat{A}$ *be its completion. Let S be a saturated multiplicative subset of A that satisfies left Ore conditions (resp., invariant, resp., central), and*

$$\widehat{S} = \left\{\widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0}\ et\ \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S\right\}$$

the set of classes of Cauchy sequences in A with values in S that do not converge to 0*. Then, the correspondence* ψ *defined by*

$$\begin{aligned} \psi : \widehat{S^{-1}A} &\to \widehat{S}^{-1}\widehat{A} \\ \widehat{\left(\frac{a_n}{s_n}\right)} &\mapsto \frac{\widehat{(a_n)}}{\widehat{(s_n)}} \end{aligned}$$

is an isomorphism of rings.

Proof

1. Let us show that ψ is a function.

Let $\widehat{\left(\frac{a_n}{s_n}\right)}$ and $\widehat{\left(\frac{b_n}{t_n}\right)} \in \widehat{S^{-1}A}$ such that $\widehat{\left(\frac{a_n}{s_n}\right)} = \widehat{\left(\frac{b_n}{t_n}\right)}$. Let us show that $\psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\right) = \psi\left(\widehat{\left(\frac{b_n}{t_n}\right)}\right)$.

By $\widehat{\left(\frac{a_n}{s_n}\right)} = \widehat{\left(\frac{b_n}{t_n}\right)} \Rightarrow \left(\frac{a_n}{s_n}\right) - \left(\frac{b_n}{t_n}\right) = \left(\frac{a_n}{s_n} - \frac{b_n}{t_n}\right) \to 0 \Rightarrow \exists u_n, v_n \in S, \forall n \in \mathbb{N}$ such that $\left(\frac{a_n}{s_n}\right) - \left(\frac{b_n}{t_n}\right) = \left(\frac{u_n a_n - v_n b_n}{u_n s_n}\right) = \frac{(u_n a_n - v_n b_n)}{(u_n s_n)} \to 0$ with $u_n s_n = v_n t_n, \forall n \in \mathbb{N}$.

Let (u_n) and (v_n) be the sequences with general terms u_n and v_n, respectively, and since $u_n s_n = v_n t_n, \forall n \in \mathbb{N}$, then $(u_n s_n) = (v_n t_n) \Rightarrow (u_n) \times (s_n) = (v_n) \times (t_n)$, where $(u_n), (v_n) \in \mathcal{C}(S)$.

Then, $\frac{(u_n a_n - v_n b_n)}{(u_n s_n)} = \frac{(u_n)\times(a_n)-(v_n)\times(b_n)}{(u_n)\times(s_n)} \to 0 \Rightarrow (u_n)\times(a_n)-(v_n)\times(b_n) \to 0$ and

$(u_n) \times (s_n) \not\to 0$.

Thus, we have

$$\begin{cases} (u_n) \times (a_n) - (v_n) \times (b_n) \to 0 \\ (u_n) \times (s_n) \not\to 0 \end{cases} \text{ with } (u_n) \times (s_n) = (v_n) \times (t_n), \text{ where } (u_n), (v_n) \in \mathcal{C}(S)$$

$$\Rightarrow \begin{cases} \widehat{(u_n)}\widehat{\times}\widehat{(a_n)} = \widehat{(v_n)}\widehat{\times}\widehat{(b_n)} \\ \widehat{(u_n)}\widehat{\times}\widehat{(s_n)} \neq \widehat{0} \end{cases} \text{ with } \widehat{(u_n)}\widehat{\times}\widehat{(s_n)} = \widehat{(v_n)}\widehat{\times}\widehat{(t_n)}, \text{ where } \widehat{(u_n)}, \widehat{(v_n)} \in \widehat{S}$$

$$\Rightarrow \frac{\widehat{(u_n)}\widehat{\times}\widehat{(a_n)}}{\widehat{(u_n)}\widehat{\times}\widehat{(s_n)}} = \frac{\widehat{(v_n)}\widehat{\times}\widehat{(b_n)}}{\widehat{(u_n)}\widehat{\times}\widehat{(s_n)}} \Rightarrow \frac{\widehat{(u_n)}\widehat{\times}\widehat{(a_n)}}{\widehat{(u_n)}\widehat{\times}\widehat{(s_n)}} = \frac{\widehat{(v_n)}\widehat{\times}\widehat{(b_n)}}{\widehat{(v_n)}\widehat{\times}\widehat{(t_n)}} \Rightarrow \frac{\widehat{(a_n)}}{\widehat{(s_n)}} = \frac{\widehat{(b_n)}}{\widehat{(t_n)}}.$$

Then $\psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\right) = \psi\left(\widehat{\left(\frac{b_n}{t_n}\right)}\right)$.

2. Let us show that ψ is a ring homomorphism.

   Let $\frac{\widehat{(a_n)}}{\widehat{(s_n)}}, \frac{\widehat{(b_n)}}{\widehat{(t_n)}} \in \widehat{S}^{-1}\widehat{A}$. We have:

   a. Let $\widehat{\left(\frac{a_n}{s_n}\right)}, \widehat{\left(\frac{b_n}{t_n}\right)} \in \widehat{S^{-1}A}$, and let us show that $\psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\widehat{+}\widehat{\left(\frac{b_n}{t_n}\right)}\right) = \psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\right)\widehat{+}\psi\left(\widehat{\left(\frac{b_n}{t_n}\right)}\right)$.

   $\psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\widehat{+}\widehat{\left(\frac{b_n}{t_n}\right)}\right) = \psi\left(\widehat{\left(\frac{a_n}{s_n}\right) + \left(\frac{b_n}{t_n}\right)}\right) = \psi\left(\widehat{\left(\frac{a_n}{s_n} + \frac{b_n}{t_n}\right)}\right) = \psi\left(\widehat{\left(\frac{u_n a_n + v_n b_n}{u_n s_n}\right)}\right)$ with $u_n s_n = v_n t_n$, where $u_n, v_n \in S, \forall n \in \mathbb{N}$.

   Let (u_n) and (v_n) be the sequences with general terms u_n and v_n, respectively. Since $u_n s_n = v_n t_n, \forall n \in \mathbb{N}$, we have $(u_n s_n) = (v_n t_n) \Rightarrow (u_n) \times (s_n) = (v_n) \times (t_n)$, where $(u_n), (v_n) \in \mathcal{C}(S)$, and hence $\psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\widehat{+}\widehat{\left(\frac{b_n}{t_n}\right)}\right) = \psi\left(\widehat{\left(\frac{u_n a_n + v_n b_n}{u_n s_n}\right)}\right) = \frac{\widehat{(u_n a_n + v_n b_n)}}{\widehat{(u_n s_n)}}$ with $\widehat{(u_n)}\widehat{\times}\widehat{(s_n)} = \widehat{(v_n)}\widehat{\times}\widehat{(t_n)}$, where $\widehat{(u_n)}, \widehat{(v_n)} \in \widehat{S}$ because $(u_n s_n) = (v_n t_n) \Rightarrow (u_n) \times (s_n) = (v_n) \times (t_n)$, where $(u_n), (v_n) \in \mathcal{C}(S)$.

$\Rightarrow /\ \psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\widehat{+}\widehat{\left(\frac{b_n}{t_n}\right)}\right) = /\ \psi\left(\widehat{\left(\frac{u_n a_n + v_n b_n}{u_n s_n}\right)}\right) = /\ \frac{\widehat{(u_n a_n + v_n b_n)}}{\widehat{(u_n s_n)}} = /$
$\frac{\widehat{(u_n)}\widehat{\times}\widehat{(a_n)}}{\widehat{(u_n)}\widehat{\times}\widehat{(s_n)}} \widehat{+} \frac{\widehat{(v_n)}\widehat{\times}\widehat{(b_n)}}{\widehat{(u_n)}\widehat{\times}\widehat{(s_n)}} = \frac{\widehat{(u_n)}\widehat{\times}\widehat{(a_n)}}{\widehat{(u_n)}\widehat{\times}\widehat{(s_n)}} \widehat{+} \frac{\widehat{(v_n)}\widehat{\times}\widehat{(b_n)}}{\widehat{(v_n)}\widehat{\times}\widehat{(t_n)}} = \frac{\widehat{(a_n)}}{\widehat{(s_n)}} \widehat{+} \frac{\widehat{(b_n)}}{\widehat{(s_n)}}$
$= \psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\right) \widehat{+} \psi\left(\widehat{\left(\frac{b_n}{t_n}\right)}\right).$

Then, $\psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\widehat{+}\widehat{\left(\frac{b_n}{t_n}\right)}\right) = \psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\right) \widehat{+} \psi\left(\widehat{\left(\frac{b_n}{t_n}\right)}\right).$

b. Let us show that $\psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\widehat{\times}\widehat{\left(\frac{b_n}{t_n}\right)}\right) = \psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\right) \widehat{\times} \psi\left(\widehat{\left(\frac{b_n}{t_n}\right)}\right).$

We have:
$\psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\widehat{\times}\widehat{\left(\frac{b_n}{t_n}\right)}\right) = \psi\left(\widehat{\left(\frac{a_n}{s_n}\right) \times \left(\frac{b_n}{t_n}\right)}\right) = \psi\left(\widehat{\left(\frac{a_n}{s_n} \times \frac{b_n}{t_n}\right)}\right) = \psi\left(\widehat{\left(\frac{z_n b_n}{w_n s_n}\right)}\right)$
with $w_n a_n = z_n t_n$, where $w_n \in S, z_n \in A, \forall n \in \mathbb{N}$.

Let us define (w_n) and (z_n) as the sequences of general terms w_n and z_n, respectively. Then, $\psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\widehat{\times}\widehat{\left(\frac{b_n}{t_n}\right)}\right) = \psi\left(\widehat{\left(\frac{z_n b_n}{w_n s_n}\right)}\right) = \frac{\widehat{(z_n b_n)}}{\widehat{(w_n s_n)}}$

$$= \frac{\widehat{(z_n)}\widehat{\times}\widehat{(b_n)}}{\widehat{(w_n)}\widehat{\times}\widehat{(s_n)}} = \frac{\widehat{(a_n)}}{\widehat{(s_n)}} \widehat{\times} \frac{\widehat{(b_n)}}{\widehat{(t_n)}} = \psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\right) \widehat{\times} \psi\left(\widehat{\left(\frac{b_n}{t_n}\right)}\right)$$

with $\widehat{(w_n)}\widehat{\times}\widehat{(a_n)} = \widehat{(z_n)}\widehat{\times}\widehat{(t_n)}$, where $\widehat{(w_n)} \in \widehat{S}, \widehat{(z_n)} \in \widehat{A}$ because $w_n a_n = /\ z_n t_n$, where $w_n \in /\ S, z_n \in /\ A, \forall n \in /\ \mathbb{N}$. Then, we have $\psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\widehat{\times}\widehat{\left(\frac{b_n}{t_n}\right)}\right) = \psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\right) \widehat{\times} \psi\left(\widehat{\left(\frac{b_n}{t_n}\right)}\right).$

c. We have:
$1_{\widehat{S^{-1}A}} = \widehat{\left(\frac{1}{1}\right)}$, and then $\psi(1_{\widehat{S^{-1}A}}) = \psi\left(\widehat{\left(\frac{1}{1}\right)}\right) = \left(\frac{\widehat{1}}{\widehat{1}}\right) = 1_{\widehat{S}^{-1}\widehat{A}}$.

Then, ψ is a ring homomorphism.

3. Let us consider the correspondence:

$$\varphi : \widehat{S}^{-1}\widehat{A} \to \widehat{S^{-1}A}$$
$$\frac{\widehat{(a_n)}}{\widehat{(s_n)}} \mapsto \widehat{\left(\frac{a_n}{s_n}\right)}.$$

Let us show that φ is a map.

Let $\frac{\widehat{(a_n)}}{\widehat{(s_n)}}, \frac{\widehat{(b_n)}}{\widehat{(t_n)}} \in \widehat{S}^{-1}\widehat{A}$ such that $\frac{\widehat{(a_n)}}{\widehat{(s_n)}} = \frac{\widehat{(b_n)}}{\widehat{(t_n)}}$, and let us show that $\varphi\left(\frac{\widehat{(a_n)}}{\widehat{(s_n)}}\right) = /$ $\varphi\left(\frac{\widehat{(b_n)}}{\widehat{(t_n)}}\right)$.

We have: $\varphi\left(\frac{\widehat{(a_n)}}{\widehat{(s_n)}}\right) = \widehat{\left(\frac{a_n}{s_n}\right)}$.

But $\frac{\widehat{(a_n)}}{\widehat{(s_n)}} = /\ \frac{\widehat{(b_n)}}{\widehat{(t_n)}} \Rightarrow /\ \frac{\widehat{(a_n)}}{\widehat{(s_n)}} \widehat{+} (-\frac{\widehat{(b_n)}}{\widehat{(t_n)}}) = \widehat{0} \Rightarrow /\ \frac{\widehat{(u_n)}\widehat{\times}\widehat{(a_n)} - \widehat{(v_n)}\widehat{\times}\widehat{(b_n)}}{\widehat{(u_n)}\widehat{\times}\widehat{(s_n)}} = \widehat{0} \Rightarrow /$
$\frac{\widehat{(u_n a_n - v_n b_n)}}{\widehat{(u_n s_n)}} = \widehat{0}$

$\Rightarrow (u_n a_n - v_n b_n) \to 0$ et $(u_n s_n) \not\to 0$, where $\widehat{(u_n)}\widehat{\times}\widehat{(s_n)} = \widehat{(v_n)}\widehat{\times}\widehat{(t_n)}$, where $\widehat{(u_n)}, \widehat{(v_n)} \in \widehat{S}$.

Then, $\frac{(u_n a_n - v_n b_n)}{(u_n s_n)} \to 0 \Rightarrow \frac{(u_n a_n)}{(u_n s_n)} - \frac{(v_n b_n)}{(u_n s_n)} \to 0 \Rightarrow \widehat{\left(\frac{u_n a_n}{u_n s_n}\right)} = \widehat{\left(\frac{v_n b_n}{u_n s_n}\right)} \Rightarrow$
$\varphi\left(\frac{\widehat{(a_n)}}{\widehat{(s_n)}}\right) = \varphi\left(\frac{\widehat{(v_n b_n)}}{\widehat{(u_n s_n)}}\right) = \varphi\left(\frac{\widehat{(u_n)}\widehat{\times}\widehat{(b_n)}}{\widehat{(u_n)}\widehat{\times}\widehat{(s_n)}}\right) = \varphi\left(\frac{\widehat{(u_n)}\widehat{\times}\widehat{(b_n)}}{\widehat{(v_n)}\widehat{\times}\widehat{(t_n)}}\right) = \varphi\left(\frac{\widehat{(b_n)}}{\widehat{(t_n)}}\right)$, where $\widehat{(u_n)}\widehat{\times}\widehat{(s_n)} = \widehat{(v_n)}\widehat{\times}\widehat{(t_n)}$. Thus $\varphi\left(\frac{\widehat{(a_n)}}{\widehat{(s_n)}}\right) = \varphi\left(\frac{\widehat{(b_n)}}{\widehat{(t_n)}}\right)$.

Moreover, we have:

a. $\psi \circ \varphi\left(\frac{\widehat{(a_n)}}{\widehat{(s_n)}}\right) = \psi\left(\varphi\left(\frac{\widehat{(a_n)}}{\widehat{(s_n)}}\right)\right) = \psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\right) = \frac{\widehat{(a_n)}}{\widehat{(s_n)}}$, for all $\frac{\widehat{(a_n)}}{\widehat{(s_n)}} \in \widehat{S}^{-1}\widehat{A}$.
Then, $\psi \circ \varphi = Id_{\widehat{S}^{-1}\widehat{A}}$.

b. $\varphi \circ \psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\right) = \varphi\left(\psi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\right)\right) = \varphi\left(\frac{\widehat{(a_n)}}{\widehat{(s_n)}}\right) = \widehat{\left(\frac{a_n}{s_n}\right)}$, for all $\widehat{\left(\frac{a_n}{s_n}\right)} \in \widehat{S^{-1}A}$.
Then, $\varphi \circ \psi = Id_{\widehat{S^{-1}A}}$.

Thus, φ and ψ are bijective, and each one is the inverse of the other ($\varphi^{-1} = \psi$ and $\psi^{-1} = \varphi$).

Corollary 9.4 *Let $(A, (I_n)_{n\in\mathbb{N}})$ be a filtered ring where $(I_n)_{n\in\mathbb{N}}$ is a sequence of left ideals, and $\widehat{A}$ its completion. Let S be a saturated multiplicative subset of A that satisfies the left Ore conditions (resp., invariant, resp., central). Let $(M, (M_n)_{n\in\mathbb{N}})$ be a left A-module filtered, where $(M_n)_{n\in\mathbb{N}}$ is a sequence of submodules of M, and*

$$\widehat{S} = \left\{\widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0} \textit{ and } \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S\right\},$$

the set of classes of Cauchy sequences in A with values in S that do not converge to 0*. Then:*

1. *The module $\widehat{S}^{-1}\widehat{M}$ has a left $\widehat{S^{-1}A}$-module structure.*
2. *The module $\widehat{S^{-1}M}$ has a left $\widehat{S}^{-1}\widehat{A}$-module structure.*

Proof It suffices to set:

1. $\widehat{\left(\frac{a_n}{s_n}\right)} \,\widehat{\bullet}\, \frac{\widehat{(m_n)}}{\widehat{(t_n)}} = \psi\left(\frac{\widehat{(a_n)}}{\widehat{(s_n)}}\right) \widehat{\bullet} \frac{\widehat{(m_n)}}{\widehat{(t_n)}}$.
2. $\frac{\widehat{(a_n)}}{\widehat{(s_n)}} \,\widehat{\bullet}\, \widehat{\left(\frac{m_n}{t_n}\right)} = \psi^{-1}\left(\frac{\widehat{(a_n)}}{\widehat{(s_n)}}\right) \widehat{\bullet} \left(\frac{\widehat{(m_n)}}{\widehat{(t_n)}}\right)$.

Theorem 9.6 *Let $(A, (I_n)_{n\in\mathbb{N}})$ be a filtered ring where $(I_n)_{n\in\mathbb{N}}$ is a sequence of left ideals, $\widehat{A}$ its completion, S a saturated multiplicative subset of A that satisfies the left Ore conditions (resp., invariant, resp., central), $(M, (M_n)_{n\in\mathbb{N}})$ a left A-module, where $(M_n)_{n\in\mathbb{N}}$ is a sequence of submodules of M, and*

$$\widehat{S} = \left\{\widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0} \textit{ et } \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S\right\}$$

the set of classes of Cauchy sequences of A with values in S that do not converge to 0*. Then, the left* $\widehat{S}^{-1}\widehat{A}$*-module* $\widehat{S}^{-1}\widehat{M}$ *and the left* $\widehat{S^{-1}A}$*-module* $\widehat{S^{-1}M}$ *are isomorphic.*

Proof Let us consider the correspondence:

$$\begin{aligned} \vartheta : \widehat{S^{-1}M} &\to \widehat{S}^{-1}\widehat{M} \\ \widehat{\left(\frac{m_n}{s_n}\right)} &\mapsto \frac{\widehat{(m_n)}}{\widehat{(s_n)}} \end{aligned}.$$

1. Let us show that ϑ is a map.

   Let $\widehat{\left(\frac{m_n}{s_n}\right)}, \widehat{\left(\frac{m'_n}{s'_n}\right)} \in \widehat{S^{-1}M}$ such that $\widehat{\left(\frac{m_n}{s_n}\right)} = \widehat{\left(\frac{m'_n}{s'_n}\right)}$, and let us show that $\vartheta\left(\widehat{\left(\frac{m_n}{s_n}\right)}\right) = \vartheta\left(\widehat{\left(\frac{m'_n}{s'_n}\right)}\right)$.

   We have:
   $\widehat{\left(\frac{m_n}{s_n}\right)} = \widehat{\left(\frac{m'_n}{s'_n}\right)} \Rightarrow \left(\frac{m_n}{s_n}\right) - \left(\frac{m'_n}{s'_n}\right) = \left(\frac{m_n}{s_n} - \frac{m'_n}{s'_n}\right) = \left(\frac{x_n m_n - y_n m'_n}{x_n s_n}\right) \to 0$ with $x_n s_n = y_n s'_n$, where $x_n, y_n \in S, \forall n \in \mathbb{N}$.

   Let (x_n) and (y_n) be the sequences of general terms x_n and y_n, respectively, and as $x_n s_n = y_n s'_n$ where $x_n, y_n \in S, \forall n \in \mathbb{N}$, then $(x_n) \times (s_n) = (y_n) \times (s'_n)$ where $(x_n), (y_n) \in \mathcal{C}(S)$. And then $\left(\frac{x_n m_n - y_n m'_n}{x_n s_n}\right) = \frac{(x_n m_n - y_n m'_n)}{(x_n s_n)} \to 0 \Rightarrow (x_n m_n - y_n m'_n) \to 0$ et $(x_n s_n) \not\to 0$, and then we have:

   $\begin{cases} (x_n m_n - y_n m'_n) \to 0 \\ (x_n s_n) \not\to 0 \end{cases}$ with $(x_n) \times (s_n) = (y_n) \times (s'_n)$ where $(x_n), (y_n) \in \mathcal{C}(S)$

   $\Rightarrow \begin{cases} (x_n) \cdot (m_n) - (y_n) \cdot (m'_n) \to 0 \\ (x_n) \times (s_n) \not\to 0 \end{cases}$ with $(x_n) \times (s_n) = (y_n) \times (s'_n)$ where $(x_n), (y_n) \in \mathcal{C}(S)$

   $\Rightarrow \begin{cases} \widehat{(x_n)} \widehat{\cdot} \widehat{(m_n)} = \widehat{(y_n)} \widehat{\cdot} \widehat{(m'_n)} \\ \widehat{(x_n)} \widehat{\times} \widehat{(s_n)} \neq \widehat{0} \end{cases}$ with $\widehat{(x_n)} \widehat{\times} \widehat{(s_n)} = \widehat{(y_n)} \widehat{\times} \widehat{(s'_n)}$ where $\widehat{(x_n)}, \widehat{(y_n)} \in \widehat{S}$. $\Rightarrow \frac{\widehat{(x_n)} \widehat{\cdot} \widehat{(m_n)}}{\widehat{(x_n)} \widehat{\times} \widehat{(s_n)}} = \frac{\widehat{(y_n)} \widehat{\cdot} \widehat{(m'_n)}}{\widehat{(y_n)} \widehat{\times} \widehat{(s'_n)}} \Rightarrow \frac{\widehat{(m_n)}}{\widehat{(s_n)}} = \frac{\widehat{(m'_n)}}{\widehat{(s'_n)}} \Rightarrow \vartheta\left(\widehat{\left(\frac{m_n}{s_n}\right)}\right) = \vartheta\left(\widehat{\left(\frac{m'_n}{s'_n}\right)}\right)$.

2. Let us show that ϑ is a module homomorphism:

   a. Let $\widehat{\left(\frac{m_n}{s_n}\right)}, \widehat{\left(\frac{m'_n}{s'_n}\right)} \in \widehat{S^{-1}M}$, and let us show that $\vartheta\left(\widehat{\left(\frac{m_n}{s_n}\right)} \widehat{+} \widehat{\left(\frac{m'_n}{s'_n}\right)}\right) = \vartheta\left(\widehat{\left(\frac{m_n}{s_n}\right)}\right) \widehat{+} \vartheta\left(\widehat{\left(\frac{m'_n}{s'_n}\right)}\right)$.

      We have:
      $\vartheta\left(\widehat{\left(\frac{m_n}{s_n}\right)} \widehat{+} \widehat{\left(\frac{m'_n}{s'_n}\right)}\right) = \vartheta\left(\widehat{\left(\frac{m_n}{s_n} + \frac{m'_n}{s'_n}\right)}\right) = \vartheta\left(\widehat{\left(\frac{x_n m_n + y_n m'_n}{x_n s_n}\right)}\right)$ with $x_n s_n = y_n s'_n$ where $x_n, y_n \in S, \forall n \in \mathbb{N}$.

Let (x_n) and (y_n) be the sequences of general terms x_n and y_n, respectively, and since $x_n s_n =/ y_n s'_n$ where $x_n, y_n \in/ S, \forall n \in/ \mathbb{N}$, then $(x_n) \times /(s_n) =/ (y_n) \times /(s'_n)$ where $(x_n), (y_n) \in/ \mathcal{C}(S)$. Therefore, $\widehat{(x_n)}\widehat{\times}\widehat{(s_n)} =/ \widehat{(y_n)}\widehat{\times}\widehat{(s'_n)}$ where $\widehat{(x_n)}, \widehat{(y_n)} \in \widehat{S}$.

Then, $\vartheta\left(\widehat{\left(\frac{m_n}{s_n}\right)}\widehat{+}\widehat{\left(\frac{m'_n}{s'_n}\right)}\right) = \vartheta\left(\widehat{\left(\frac{x_n m_n + y_n m'_n}{x_n s_n}\right)}\right) = /\frac{\widehat{(x_n m_n + y_n m'_n)}}{\widehat{(x_n s_n)}}$

$= /\frac{\left(\widehat{(x_n)}\widehat{\times}\widehat{(m_n)}\widehat{+}\widehat{(y_n)}\widehat{\times}\widehat{(m'_n)}\right)}{\widehat{(x_n)}\widehat{\times}\widehat{(s_n)}} = /\frac{\widehat{(x_n)}\widehat{\times}\widehat{(m_n)}}{\widehat{(x_n)}\widehat{\times}\widehat{(s_n)}}\widehat{+}\frac{\widehat{(y_n)}\widehat{\times}\widehat{(m'_n)}}{\widehat{(x_n)}\widehat{\times}\widehat{(s_n)}}$

$= /\frac{\widehat{(x_n)}\widehat{\times}\widehat{(m_n)}}{\widehat{(x_n)}\widehat{\times}\widehat{(s_n)}}\widehat{+}\frac{\widehat{(y_n)}\widehat{\times}\widehat{(m'_n)}}{\widehat{(y_n)}\widehat{\times}\widehat{(s'_n)}} = \frac{\widehat{(m_n)}}{\widehat{(s_n)}}\widehat{+}\frac{\widehat{(m'_n)}}{\widehat{(s'_n)}} = /\vartheta\left(\widehat{\left(\frac{m_n}{s_n}\right)}\right)\widehat{+}\vartheta\left(\widehat{\left(\frac{m'_n}{s'_n}\right)}\right).$

b. Let $\widehat{\left(\frac{a_n}{s_n}\right)} \in \widehat{S^{-1}A}$, $\widehat{\left(\frac{m_n}{t_n}\right)} \in \widehat{S^{-1}M}$, and let us show

$\vartheta\left(\widehat{\left(\frac{a_n}{s_n}\right)}\widehat{\bullet}\widehat{\left(\frac{m_n}{t_n}\right)}\right) = \widehat{\left(\frac{a_n}{s_n}\right)}\widehat{\bullet}\vartheta\left(\widehat{\left(\frac{m_n}{t_n}\right)}\right).$

We have:

$\vartheta\left(\widehat{\left(\frac{a_n}{s_n}\right)}\widehat{\bullet}\widehat{\left(\frac{m_n}{t_n}\right)}\right) =/ \vartheta\left(\widehat{\left(\frac{a_n}{s_n}\bullet /\frac{m_n}{t_n}\right)}\right) =/ \vartheta\left(\widehat{\left(\frac{z_n \cdot m_n}{w_n \times s_n}\right)}\right)$ with $w_n a_n =/ z_n t_n$ where $w_n \in /S, z_n \in /S, \forall n \in /\mathbb{N}$.

Let (w_n) and (z_n) be the sequences of general terms w_n and z_n, respectively, and since $w_n a_n =/ z_n t_n$ where $w_n \in/ S, z_n \in/ S, \forall n \in/ \mathbb{N}$, then $(w_n) \times /(a_n) = /(z_n) \times /(t_n)$ where $(w_n) \in \mathcal{C}(S), (z_n) \in \mathcal{C}(A) \Rightarrow /\widehat{(w_n)}\widehat{\times}\widehat{(a_n)} = \widehat{(z_n)}\widehat{\times}\widehat{(t_n)}$ where $\widehat{(w_n)} \in \widehat{S}$, $\widehat{(z_n)} \in \widehat{A}$.

Then:

$\vartheta\left(\widehat{\left(\frac{a_n}{s_n}\right)}\widehat{\bullet}\widehat{\left(\frac{m_n}{t_n}\right)}\right) =/ \vartheta\left(\widehat{\left(\frac{a_n}{s_n}\bullet /\frac{m_n}{t_n}\right)}\right) =/ \vartheta\left(\widehat{\left(\frac{z_n \cdot m_n}{w_n \times s_n}\right)}\right) = \frac{\widehat{(z_n m_n)}}{\widehat{(w_n s_n)}} =/ \frac{\widehat{(z_n)}\widehat{\cdot}\widehat{(m_n)}}{\widehat{(w_n)}\widehat{\times}\widehat{(s_n)}} = /\frac{\widehat{(a_n)}}{\widehat{(s_n)}}\widehat{\bullet}\frac{\widehat{(m_n)}}{\widehat{(t_n)}}$ by $\widehat{(w_n)}\widehat{\times}\widehat{(a_n)} = \widehat{(z_n)}\widehat{\times}\widehat{(t_n)}$ where $\widehat{(w_n)} \in \widehat{S}$, $\widehat{(z_n)} \in \widehat{A}$.

Since $\widehat{S^{-1}A}$ and $\widehat{S}^{-1}\widehat{A}$ are isomorphic, then let us set $\frac{\widehat{(a_n)}}{\widehat{(s_n)}} \simeq \widehat{\left(\frac{a_n}{s_n}\right)}$, and then

$\vartheta\left(\widehat{\left(\frac{a_n}{s_n}\right)}\widehat{\bullet}\widehat{\left(\frac{m_n}{t_n}\right)}\right) = /\frac{\widehat{(a_n)}}{\widehat{(s_n)}}\widehat{\bullet}\frac{\widehat{(m_n)}}{\widehat{(t_n)}} = \widehat{\left(\frac{a_n}{s_n}\right)}\widehat{\bullet}\frac{\widehat{(m_n)}}{\widehat{(t_n)}} = \widehat{\left(\frac{a_n}{s_n}\right)}\widehat{\bullet}\vartheta\left(\widehat{\left(\frac{m_n}{t_n}\right)}\right).$

Then ϑ is a module homomorphism.

3. Let us show that ϑ is bijective.

To do this, consider the correspondence:

$$\begin{aligned} \phi : \widehat{S}^{-1}\widehat{M} &\to \widehat{S^{-1}M} \\ \frac{\widehat{(m_n)}}{\widehat{(s_n)}} &\mapsto \widehat{\left(\frac{m_n}{s_n}\right)} \end{aligned}.$$

Let us show that ϕ is a function.

Let $\frac{\widehat{(m_n)}}{\widehat{(s_n)}}, \frac{\widehat{(m'_n)}}{\widehat{(t_n)}} \in \widehat{S}^{-1}\widehat{A}$ such that $\frac{\widehat{(m_n)}}{\widehat{(s_n)}} = \frac{\widehat{(m'_n)}}{\widehat{(t_n)}}$, and let us show that $\phi\left(\frac{\widehat{(m_n)}}{\widehat{(s_n)}}\right) = /\phi\left(\frac{\widehat{(m'_n)}}{\widehat{(t_n)}}\right)$.

We have: $\phi\left(\frac{\widehat{(m_n)}}{\widehat{(s_n)}}\right) = \widehat{\left(\frac{m_n}{s_n}\right)}$.

But $\frac{\widehat{(m_n)}}{\widehat{(s_n)}} = \frac{\widehat{(m'_n)}}{\widehat{(t_n)}} \Rightarrow \frac{\widehat{(m_n)}}{\widehat{(s_n)}} \widehat{+} (-\frac{\widehat{(m'_n)}}{\widehat{(t_n)}}) = \widehat{0} \Rightarrow / \frac{\widehat{(u_n)}\widehat{\times}\widehat{(m_n)} - \widehat{(v_n)}\widehat{\times}\widehat{(m'_n)}}{\widehat{(u_n)}\widehat{\times}\widehat{(s_n)}} = \widehat{0} \Rightarrow /$
$\frac{\widehat{(u_n m_n - v_n m'_n)}}{\widehat{(u_n s_n)}} = \widehat{0}$

$\Rightarrow \not(u_n m_n - v_n m'_n) \to \not 0$ and $(u_n s_n) \not\to \not 0$ with $\widehat{(u_n)}\widehat{\times}\widehat{(s_n)} = \not\widehat{(v_n)}\widehat{\times}\widehat{(t_n)}$ where $\widehat{(u_n)}, \widehat{(v_n)} \in \widehat{S}$.

Then, $\frac{(u_n m_n - v_n m'_n)}{(u_n s_n)} \to \not 0 \Rightarrow \not\frac{(u_n m_n)}{(u_n s_n)} - \not\frac{(v_n m'_n)}{(u_n s_n)} \to \not 0 \Rightarrow \widehat{\left(\frac{u_n m_n}{u_n s_n}\right)} = \widehat{\left(\frac{v_n m'_n}{u_n s_n}\right)} \Rightarrow /$
$\phi\left(\frac{\widehat{(m_n)}}{\widehat{(s_n)}}\right) = \not\phi\left(\frac{\widehat{(v_n m'_n)}}{\widehat{(u_n s_n)}}\right) = \not\phi\left(\frac{\widehat{(u_n)}\widehat{\times}\widehat{(m'_n)}}{\widehat{(u_n)}\widehat{\times}\widehat{(s_n)}}\right) = \not\phi\left(\frac{\widehat{(u_n)}\widehat{\times}\widehat{(m'_n)}}{\widehat{(v_n)}\widehat{\times}\widehat{(t_n)}}\right) = \not\phi\left(\frac{\widehat{(m'_n)}}{\widehat{(t_n)}}\right)$ by $\widehat{(u_n)}\widehat{\times}\widehat{(s_n)} = \not\widehat{(v_n)}\widehat{\times}\widehat{(t_n)}$. Then $\phi\left(\frac{\widehat{(m_n)}}{\widehat{(s_n)}}\right) = \not\phi\left(\frac{\widehat{(m'_n)}}{\widehat{(t_n)}}\right)$. Moreover, we have:

(a) $\psi \circ \phi\left(\frac{\widehat{(m_n)}}{\widehat{(s_n)}}\right) = \psi\left(\phi\left(\frac{\widehat{(m_n)}}{\widehat{(s_n)}}\right)\right) = \not\psi\left(\widehat{\left(\frac{m_n}{s_n}\right)}\right) = \frac{\widehat{(m_n)}}{\widehat{(s_n)}}$, for all $\frac{\widehat{(m_n)}}{\widehat{(s_n)}} \in \widehat{S}^{-1}\widehat{M}$. Then, $\psi \circ \not\phi = \not id_{\widehat{S}^{-1}\widehat{M}}$.

(b) $\phi \circ \not\psi\left(\widehat{\left(\frac{m_n}{s_n}\right)}\right) = \not\phi\left(\psi\left(\widehat{\left(\frac{m_n}{s_n}\right)}\right)\right) = \not\phi\left(\frac{\widehat{(m_n)}}{\widehat{(s_n)}}\right) = \widehat{\left(\frac{m_n}{s_n}\right)}$, for all $\widehat{\left(\frac{m_n}{s_n}\right)} \in \widehat{S^{-1}M}$. Then, $\phi \circ \not\psi = id_{\widehat{S^{-1}M}}$.

Thus, ϕ and ϑ are bijective, and each is the inverse of the other ($\phi^{-1} = \not\vartheta$ and $\vartheta^{-1} = \not\psi$).

We thus conclude that ϑ is a left $\widehat{S^{-1}A}$-module isomorphism.

$\widehat{S}$-Saturated Submodules of Completion Modules

Lemma 9.3 *Let $(A, (I_n)_{n\in\mathbb{N}})$ be a filtered ring where $(I_n)_{n\in\mathbb{N}}$ ~~is a~~ sequence of left ideals, $\widehat{A}$ its completion, S a saturated multiplicative subset of A that satisfies the left Ore conditions (resp., invariant, resp., central), $(M, (M_n)_{n\in\mathbb{N}})$ a left A-module filtered, where $(M_n)_{n\in\mathbb{N}}$ ~~is a~~ sequence of submodules of M, and*

$$\widehat{S} = \left\{\widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0} \; et \; \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S\right\}$$

the set of classes of Cauchy sequences of A with values in S that do not converge to 0*. Then, the correspondence defined by*

$$\begin{aligned} \phi : \widehat{S}^{-1}\widehat{M} &\to \widehat{S^{-1}M} \\ \frac{\widehat{(m_n)}}{\widehat{(s_n)}} &\mapsto \widehat{\left(\frac{m_n}{s_n}\right)} \end{aligned}.$$

is a left $\widehat{S}^{-1}\widehat{A}$-module isomorphism.

Proof See Theorem 9.6.

Definitions 9.2 *Let N be a submodule of a left A-module M and S a saturated multiplicative subset of A that satisfies the left Ore conditions. We say that N is left S-saturated in M if for all $s \in S$, $x \in M$ such that $sx \in N$, then $x \in N$.*

Proposition 9.7 *Let $(A, (I_n)_{n\in\mathbb{N}})$ be a filtered ring where $(I_n)_{n\in\mathbb{N}}$ ~~is a~~ sequence of left ideals, S a saturated multiplicative subset of A that satisfies the left Ore conditions (resp., invariant, resp., central), N an S-saturated submodule of a filtered left A-module $(M, (M_n)_{n\in\mathbb{N}})$, where $(M_n)_{n\in\mathbb{N}}$ ~~is a~~ sequence of submodules of M, and*

$$\widehat{S} = \left\{\widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0} \; et \; \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S\right\}$$

the set of classes of Cauchy sequences of A with values in S that do not converge to 0. Then, the completion of N, $\widehat{N}$, is $\widehat{S}$-saturated in $\widehat{M}$.

Proof By Proposition 9.6, $\widehat{N}$ is a submodule of $\widehat{M}$. Let $\widehat{(s_n)} \in \widehat{S}$, $\widehat{(x_n)} \in \widehat{M}$ such that $\widehat{(s_n)}\widehat{\cdot}\widehat{(x_n)} \in \widehat{N}$, and let us show that $\widehat{(x_n)} \in \widehat{N}$. We have $\widehat{(s_n)}\widehat{\cdot}\widehat{(x_n)} = \widehat{(s_n x_n)} \in \widehat{N}$, so $(s_n x_n) \in \mathcal{C}(N)$, which implies that $s_n x_n \in N$, $\forall n \in \mathbb{N}$. Since N is S-saturated in M, then $x_n \in N$, $\forall n \in \mathbb{N}$. Thus, $(x_n) \in \mathcal{C}(N)$, so $\widehat{(x_n)} \in \widehat{N}$. Therefore, we conclude that $\widehat{N}$ is $\widehat{S}$-saturated in $\widehat{M}$.

Theorem 9.7 *Let $(A, (I_n)_{n\in\mathbb{N}})$ be a filtered ring where $(I_n)_{n\in\mathbb{N}}$ ~~is a~~ sequence of left ideals, S a saturated multiplicative subset of A that satisfies the left Ore conditions (resp., invariant, resp., central), N an S-saturated submodule of a filtered left A-module $(M, (M_n)_{n\in\mathbb{N}})$, where $(M_n)_{n\in\mathbb{N}}$ ~~is a~~ sequence of submodules of M, and*

$$\widehat{S} = \left\{\widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0} \; et \; \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S\right\}$$

the set of classes of Cauchy sequences of A with values in S that do not converge to 0. Then, $\widehat{N}$ is $\widehat{S}$-saturated in $\widehat{M}$ if and only if there is a submodule N' of the $\widehat{S}^{-1}\widehat{A}$-module $\widehat{S}^{-1}\widehat{M}$ such that $\widehat{N} = (i^{\widehat{S}}_{\widehat{M}})^{-1}(N')$.

Proof It should be noted that the correspondence

$$\begin{aligned} i^{\widehat{S}}_{\widehat{M}} : \quad \widehat{M} &\to \widehat{S}^{-1}\widehat{M} \\ \widehat{(m_n)} &\mapsto \frac{\widehat{(m_n)}}{\widehat{(s_n)}} \end{aligned}$$

is the canonical morphism:

1. Suppose that $\widehat{N}$ is a $\widehat{S}$-saturated submodule of $\widehat{M}$. We will show that there is a submodule $\widehat{N'}$ of the left $\widehat{S}^{-1}\widehat{A}$-module $\widehat{S}^{-1}\widehat{M}$ such that $\widehat{N} = (i^{\widehat{S}}_{\widehat{M}})^{-1}(N')$. It suffices to show that $\widehat{N} = (i^{\widehat{S}}_{\widehat{M}})^{-1}(\widehat{S}^{-1}\widehat{N})$.

We have $\widehat{N} \subset /(i^{\widehat{S}}_{\widehat{M}})^{-1}(\widehat{S}^{-1}\widehat{N})$ by $i^{\widehat{S}}_{\widehat{M}}(\widehat{N}) \subseteq \widehat{S}^{-1}\widehat{N}$.

Let $\widehat{(m_n)} \in /(i^{\widehat{S}}_{\widehat{M}})^{-1}(\widehat{S}^{-1}\widehat{N}) \Rightarrow /i^{\widehat{S}}_{\widehat{M}}(\widehat{(m_n)}) = \frac{\widehat{(m_n)}}{\widehat{1}} \in \widehat{S}^{-1}\widehat{N}$, and then there is $\widehat{(a_n)} \in /\widehat{N}$, $\widehat{(s_n)} \in \widehat{S}$ such that $\frac{\widehat{(m_n)}}{\widehat{1}} = /\frac{\widehat{(a_n)}}{\widehat{(s_n)}} \Rightarrow /\frac{\widehat{(u_n)}\widehat{\times}\widehat{(m_n)}}{\widehat{(u_n)}} = /\frac{\widehat{(v_n)}\widehat{\times}\widehat{(a_n)}}{\widehat{(v_n)}\widehat{\times}\widehat{(s_n)}}$ with $\widehat{(u_n)} = /\widehat{(v_n)}\widehat{\times}\widehat{(s_n)}$ and $\widehat{(u_n)}\widehat{\times}\widehat{(m_n)} = /\widehat{(v_n)}\widehat{\times}\widehat{(a_n)}$ where $(\widehat{(u_n)}, \widehat{(v_n)}) \in \widehat{A} \times \widehat{S}$, since $\widehat{(u_n)}\widehat{\times}\widehat{(m_n)} \in /\widehat{N} \Rightarrow /\widehat{(m_n)} \in /\widehat{N}$ because $\widehat{N}$ is $\widehat{S}$-saturated. Therefore, $(i^{\widehat{S}}_{\widehat{M}})^{-1}(\widehat{S}^{-1}\widehat{N}) \subset /\widehat{N}$.

We deduce that $\widehat{N} = (i^{\widehat{S}}_{\widehat{M}})^{-1}(\widehat{S}^{-1}\widehat{N})$ and let $N' = \widehat{S}^{-1}\widehat{N}$, hence the existence of N' such that $\widehat{N} = /(i^{\widehat{S}}_{\widehat{M}})^{-1}(N')$.

2. Suppose that there is a submodule N' of the left $\widehat{S}^{-1}\widehat{A}$-module $\widehat{S}^{-1}\widehat{M}$ such that $\widehat{N} = (i^{\widehat{S}}_{\widehat{M}})^{-1}(N')$. We will show that $\widehat{N}$ is $\widehat{S}$-saturated in $\widehat{M}$.

   Let $\widehat{(m_n)} \in /\widehat{M}$ and $\widehat{(s_n)} \in \widehat{S}$ such that $\widehat{(s_n)}\widehat{\times}\widehat{(m_n)} \in /\widehat{N}$, and let us show that $\widehat{(m_n)} \in /\widehat{N}$.

   We have $\widehat{(s_n)}\widehat{\times}\widehat{(m_n)} \in /\widehat{N}$, and then $i^{\widehat{S}}_{\widehat{M}}(\widehat{(s_n m_n)}) = \frac{\widehat{(s_n m_n)}}{\widehat{1}} \in /N'$ by the assumption. Thus, $\frac{\widehat{(s_n)}\widehat{\times}\widehat{(m_n)}}{\widehat{1}} \in /N'$. But $\widehat{(s_n)} \in \widehat{S}$ implies that $\frac{\widehat{1}}{\widehat{(s_n)}} \in \widehat{S^{-1}A}$, and thus $\frac{\widehat{1}}{\widehat{(s_n)}} \widehat{\bullet} \frac{\widehat{(s_n)}\widehat{\times}\widehat{(m_n)}}{\widehat{1}} = \frac{\widehat{(m_n)}}{\widehat{1}} \in /N'$. Therefore, $\widehat{(m_n)} \in /\widehat{N}$. We conclude that $\widehat{N}$ is $\widehat{S}$-saturated in $\widehat{M}$.

Corollary 9.5 *Let $(A, (I_n)_{n\in\mathbb{N}})$ be a filtered ring where $(I_n)_{n\in\mathbb{N}}$ ~~is a~~ sequence of left ideals, S a multiplicative subset of A that satisfies the left Ore conditions (resp., invariant, resp., central), and N an S-saturated submodule of a filtered left A-module $(M, (M_n)_{n\in\mathbb{N}})$, where $(M_n)_{n\in\mathbb{N}}$ ~~is a~~ sequence of submodules of M, and*

$$\widehat{S} = \left\{\widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0} \; et \; \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S\right\}$$

the set of classes of Cauchy sequences in A with values in S that do not converge to 0. Then, $\widehat{N}$ is $\widehat{S}$-saturated in $\widehat{M}$ if and only if there is a submodule N' of the $\widehat{S^{-1}A}$-module $\widehat{S^{-1}M}$ such that $\widehat{N} = /(\phi \circ /i^{\widehat{S}}_{\widehat{M}})^{-1}(N')$.

Proof Consider the exact sequence: $\widehat{M} \xrightarrow{i^{\widehat{S}}_{\widehat{M}}} \widehat{S}^{-1}\widehat{M} \xrightarrow{\phi} \widehat{S^{-1}M}$ and then, $\phi \circ i^{\widehat{S}}_{\widehat{M}}$ is well defined, and thus, there is a submodule N' of the $\widehat{S^{-1}A}$-module $\widehat{S^{-1}M}$ such that $\widehat{N} = /(\phi \circ /i^{\widehat{S}}_{\widehat{M}})^{-1}(N')$.

Proposition 9.8 *Let $(A, (I_n)_{n\in\mathbb{N}})$ be a filtered ring where $(I_n)_{n\in\mathbb{N}}$ ~~is a~~ sequence of left ideals, S be a saturated multiplicative subset of A that satisfies the left Ore conditions (resp., invariant, resp., central), and $(M, (M_n)_{n\in\mathbb{N}})$ be a filtered left A-*

module, where $(M_n)_{n\in\mathbb{N}}$ ~~*is a*~~ *sequence of submodules of* M*, and*

$$\widehat{S} = \left\{\widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0}\ et\ \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S\right\}$$

the set of classes of Cauchy sequences of A *with values in* S *that do not converge to* 0*. Then, for any submodule* N *of the* $\widehat{S}^{-1}\widehat{A}$*-module* $\widehat{S}^{-1}\widehat{M}$ *(resp.,* $\widehat{S^{-1}A}$*-module* $\widehat{S^{-1}M}$*), we have* $\widehat{S}^{-1}((i^{\widehat{S}}_{\widehat{M}})^{-1}(N)) = N$ *(resp.,* $\widehat{S}^{-1}((\phi \circ i^{\widehat{S}}_{\widehat{M}})^{-1}(N)) = N$*).*

Proof For all submodules N of the $\widehat{S}^{-1}\widehat{A}$-module $\widehat{S}^{-1}\widehat{M}$, let us show that $\widehat{S}^{-1}((i^{\widehat{S}}_{\widehat{M}})^{-1}(N)) = N$:

1. Let us show that $\widehat{S}^{-1}\left((i^{\widehat{S}}_{\widehat{M}})^{-1}(N)\right) \subset N$.
   Let $\frac{\widehat{(m_n)}}{\widehat{(s_n)}} \in \widehat{S}^{-1}\left((i^{\widehat{S}}_{\widehat{M}})^{-1}(N)\right) \Rightarrow \widehat{(m_n)} \in (i^{\widehat{S}}_{\widehat{M}})^{-1}(N)$ and $\widehat{(s_n)} \in \widehat{S} \Rightarrow \frac{\widehat{(m_n)}}{\widehat{1}} \in N$ and $\frac{\widehat{1}}{\widehat{(s_n)}} \in \widehat{S^{-1}A}$, and thus $\frac{\widehat{1}}{\widehat{(s_n)}} \widehat{\bullet} \frac{\widehat{(m_n)}}{\widehat{1}} = \frac{\widehat{(m_n)}}{\widehat{(s_n)}} \in N$. Then, we have $\widehat{S}^{-1}((i^{\widehat{S}}_{\widehat{M}})^{-1}(N)) \subset N$.
2. Inversely, let us show that $N \subset \widehat{S}^{-1}\left((i^{\widehat{S}}_{\widehat{M}})^{-1}(N)\right)$.
   Let $\frac{\widehat{(m_n)}}{\widehat{(s_n)}} \in N$, and then $\frac{\widehat{(m_n)}}{\widehat{(s_n)}} \in \widehat{S}^{-1}\widehat{M} \Rightarrow \widehat{(m_n)} \in \widehat{M}$. But, $\frac{\widehat{(m_n)}}{\widehat{1}} = \frac{\widehat{(s_n)}}{\widehat{1}} \widehat{\bullet} \frac{\widehat{(m_n)}}{\widehat{(s_n)}} \in N$, and then $i^{\widehat{S}}_{\widehat{M}}(\widehat{(m_n)}) \in N \Rightarrow \widehat{(m_n)} \in (i^{\widehat{S}}_{\widehat{M}})^{-1}(N) \Rightarrow \frac{\widehat{(m_n)}}{\widehat{(s_n)}} \in \widehat{S}^{-1}\left((i^{\widehat{S}}_{\widehat{M}})^{-1}(N)\right)$. Thus,
   $N \subset \widehat{S}^{-1}\left((i^{\widehat{S}}_{\widehat{M}})^{-1}(N)\right)$.

We conclude that for any submodule N of the $\widehat{S}^{-1}\widehat{A}$-module $\widehat{S}^{-1}\widehat{M}$, we have

$$\widehat{S}^{-1}((i^{\widehat{S}}_{\widehat{M}})^{-1}(N)) = N.$$

Similarly, we can show that $\widehat{S}^{-1}((\phi \circ i^{\widehat{S}}_{\widehat{M}})^{-1}(N)) = N$ using Corollary 9.5.

Corollary 9.6 *Let* $(A, (I_n)_{n\in\mathbb{N}})$ *be a filtered ring where* $(I_n)_{n\in\mathbb{N}}$ ~~*is a*~~ *sequence of left ideals, and let* S *be a saturated multiplicative subset of* A *that satisfies the left Ore conditions (resp., invariant, resp., central), and*

$$\widehat{S} = \left\{\widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0}\ et\ \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S\right\}$$

the set of classes of Cauchy sequences in A *with values in* S *that do not converge to* 0*. Then, for any ideal* I *of* $\widehat{S}^{-1}\widehat{A}$ *(resp.,* $\widehat{S^{-1}A}$*), we have* $\widehat{S}^{-1}((i^{\widehat{S}}_{\widehat{M}})^{-1}(I)) = I$ *(resp.,* $\widehat{S}^{-1}((\phi \circ i^{\widehat{S}}_{\widehat{M}})^{-1}(I)) = I$*).*

Theorem 9.8 *Let $(A, (I_n)_{n\in\mathbb{N}})$ be a filtered ring where $(I_n)_{n\in\mathbb{N}}$ ~~is~~ a sequence of left ideals, S a saturated multiplicative subset of A that satisfies the left Ore conditions (resp., invariant, resp., central), $(M, (M_n)_{n\in\mathbb{N}})$ a filtered left A-module, where $(M_n)_{n\in\mathbb{N}}$ ~~is a~~ sequence of submodules of M, and*

$$\widehat{S} = \left\{\widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0} \; et \; \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S\right\}$$

the set of classes of Cauchy sequences of A with values in S that do not converge to 0. *Then, there is an increasing bijection (for inclusion) from the set of submodules of the $\widehat{S}^{-1}\widehat{A}$-module $\widehat{S}^{-1}\widehat{M}$ to the set of submodules of the $\widehat{A}$-module $\widehat{M}$ saturated for $\widehat{S}$.*

Proof Let $\mathcal{E}_{\widehat{S}^{-1}\widehat{M}}$ be the set of submodules of $\widehat{S}^{-1}\widehat{M}$, $\mathcal{E}_{\widehat{M}}$ be the set of submodules of $\widehat{M}$ that are $\widehat{S}$-saturated, and ψ be the correspondence defined by

$$\begin{array}{rcl}\psi : \mathcal{E}_{\widehat{S}^{-1}\widehat{M}} & \to & \mathcal{E}_{\widehat{M}} \\ N & \mapsto & (i^{\widehat{S}}_{\widehat{M}})^{-1}(N)\end{array}.$$

ψ is a map by definition:

1. Show that ψ is surjective.

   Let $N' \in \mathcal{E}_{\widehat{M}}$. By definition of the localization morphism $i^{\widehat{S}}_{\widehat{M}}$, there is $N \in \mathcal{E}_{\widehat{S}^{-1}\widehat{M}}$ such that $i^{\widehat{S}}_{\widehat{M}}(N) = N'$. Thus, $N = (i^{\widehat{S}}_{\widehat{M}})^{-1}(N') = \psi(N')$. Hence, ψ is surjective.
2. Let us show that ψ is injective.

   Suppose $N_1, N_2 \in \mathcal{E}_{\widehat{S}^{-1}\widehat{M}}$ such that $\psi(N_1) = \psi(N_2)$. We will show that $N_1 = N_2$.

   Since $\psi(N_1) = \psi(N_2)$, we have $(i^{\widehat{S}}_{\widehat{M}})^{-1}(N_1) = (i^{\widehat{S}}_{\widehat{M}})^{-1}(N_2) \Rightarrow \widehat{S}^{-1}((i^{\widehat{S}}_{\widehat{M}})^{-1}(N_1)) = \widehat{S}^{-1}((i^{\widehat{S}}_{\widehat{M}})^{-1}(N_2))$
   $\Rightarrow N_1 = N_2$ by Proposition 9.8. Thus, ψ is injective.

Therefore, ψ is bijective:

3. Let us show that ψ is increasing.

   Let $N_1, N_2 \in \mathcal{E}_{\widehat{S}^{-1}\widehat{M}}$, and let us show that $\psi(N_1) \subset \psi(N_2)$.

   Suppose $\widehat{(m_n)} \in \psi(N_1)$, then $\psi(\widehat{(m_n)}) \in N_1 \subset N_2 \Rightarrow \widehat{(m_n)} \in (i^{\widehat{S}}_{\widehat{M}})^{-1}(N_2) = \psi(N_2)$. Therefore, ψ is increasing.

Therefore, there is an increasing bijection (for inclusion) from the set of submodules of the $\widehat{S}^{-1}\widehat{A}$-module $\widehat{S}^{-1}\widehat{M}$ to the set of submodules of the $\widehat{A}$-module $\widehat{M}$ saturated for $\widehat{S}$.

Corollary 9.7 *Let $(A, (I_n)_{n\in\mathbb{N}})$ be a filtered ring where $(I_n)_{n\in\mathbb{N}}$ ~~is~~ a sequence of left ideals, S a saturated multiplicative subset of A that satisfies the left Ore* *conditions (resp., invariant, resp., central),* $(M, (M_n)_{n\in\mathbb{N}})$ *a filtered left* A*-module, where* $(M_n)_{n\in\mathbb{N}}$ ~~*is a*~~ *sequence of submodules of* M*, and*

$$\widehat{S} = \left\{\widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0} \; et \; \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S\right\}$$

the set of classes of Cauchy sequences of A *with values in* S *that do not converge to* 0*. Then, there is an increasing bijection (for inclusion) from the set of submodules of the* $\widehat{S^{-1}A}$*-module* $\widehat{S^{-1}M}$ *to the set of submodules of the* $\widehat{A}$*-module* $\widehat{M}$ *saturated for* $\widehat{S}$*.*

Corollary 9.8 *Let* $(A, (I_n)_{n\in\mathbb{N}})$ *be a filtered ring where* $(I_n)_{n\in\mathbb{N}}$ ~~*is*~~ *a sequence of left ideals,* S *a saturated multiplicative subset of* A *that satisfies the left Ore conditions (resp., invariant, resp., central),* $(M, (M_n)_{n\in\mathbb{N}})$ *a filtered left* A*-module, where* $(M_n)_{n\in\mathbb{N}}$ ~~*is a*~~ *sequence of submodules of* M*, and*

$$\widehat{S} = \left\{\widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0} \; et \; \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S\right\}$$

the set of classes of Cauchy sequences of A *with values in* S *that do not converge to* 0*. If* M *is Noetherian (resp., Artinian), then* $\widehat{S}^{-1}\widehat{M}$ *and* $\widehat{S^{-1}M}$ *are also Noetherian (resp. Artinian).*

Theorem 9.9 *Let* $(A, (I_n)_{n\in\mathbb{N}})$ *be a filtered ring where* $(I_n)_{n\in\mathbb{N}}$ ~~*is a*~~ *sequence of left ideals,* S *a saturated multiplicative subset of* A *that satisfies the left Ore conditions (resp., invariant, resp., central),* N *a* S*-saturated submodule in a filtered left* A*-module* $(M, (M_n)_{n\in\mathbb{N}})$*, where* $(M_n)_{n\in\mathbb{N}}$ ~~*is a*~~ *sequence of submodules of* M*, and*

$$\widehat{S} = \left\{\widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0} \; et \; \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S\right\}$$

the set of classes of Cauchy sequences of A *with values in* S *that do not converge to* 0*. Then, the* $\widehat{S}^{-1}\widehat{A}$*-module* $\widehat{S}^{-1}\left(\widehat{M}/\widehat{N}\right)$ *is isomorphic to* $\widehat{S}^{-1}(\widehat{M})/\widehat{S}^{-1}(\widehat{N})$*.*

Proof Let ψ be the match defined by

$$\begin{array}{rcl} \psi : \widehat{S}^{-1}\left(\widehat{M}/\widehat{N}\right) & \to & \widehat{S}^{-1}(\widehat{M})/\widehat{S}^{-1}(\widehat{N}) \\ \overline{\dfrac{\widehat{((m_n))}}{\widehat{(s_n)}}} & \mapsto & \overline{\left(\dfrac{\widehat{(m_n)}}{\widehat{(s_n)}}\right)} \end{array}.$$

1. Let us show that ψ is a map.

   Let $\overline{\frac{\widehat{((m_n))}}{\widehat{(s_n)}}}, \overline{\frac{\widehat{((m'_n))}}{\widehat{(s'_n)}}} \in \widehat{S}^{-1}(M/N)$ such that $\overline{\frac{\widehat{((m_n))}}{\widehat{(s_n)}}} = \overline{\frac{\widehat{((m'_n))}}{\widehat{(s'_n)}}}$, and let us show that $\psi\left(\overline{\frac{\widehat{((m_n))}}{\widehat{(s_n)}}}\right) = \psi\left(\overline{\frac{\widehat{((m'_n))}}{\widehat{(s'_n)}}}\right)$.

   We have

$\overline{\frac{\widehat{((m_n))}}{\widehat{(s_n)}}} = \not{/} \overline{\frac{\widehat{((m'_n))}}{\widehat{(s'_n)}}}$, and then there is $\widehat{(x_n)}, \widehat{(y_n)} \in \widehat{S}$ such that:
$\begin{cases} \widehat{(x_n)} \centerdot \overline{\left(\widehat{(m_n)}\right)} = \widehat{(y_n)} \centerdot \overline{\left(\widehat{(m'_n)}\right)} \\ \widehat{(x_n)} \widehat{\times} \widehat{(s_n)} = \widehat{(y_n)} \widehat{\times} \widehat{(t_n)} \end{cases} \Rightarrow \not{/} \begin{cases} \widehat{(x_n)} \widehat{\cdot} \widehat{(m_n)} = \not{/} \widehat{(y_n)} \widehat{\cdot} \widehat{(m'_n)} \\ \widehat{(x_n)} \widehat{\times} \widehat{(s_n)} = \not{/} \widehat{(y_n)} \widehat{\times} \widehat{(t_n)} \end{cases} \Rightarrow \not{/} \frac{\widehat{(x_n)} \widehat{\cdot} \widehat{(m_n)}}{\widehat{(x_n)} \widehat{\times} \widehat{(s_n)}} = \not{/} \frac{\widehat{(y_n)} \widehat{\cdot} \widehat{(m'_n)}}{\widehat{(y_n)} \widehat{\times} \widehat{(t_n)}}$
$\frac{\widehat{(m_n)}}{\widehat{(s_n)}} = \frac{\widehat{(m'_n)}}{\widehat{(t_n)}} \Rightarrow \not{/} \overline{\left(\frac{\widehat{(m_n)}}{\widehat{(s_n)}}\right)} = \not{/} \overline{\left(\frac{\widehat{(m'_n)}}{\widehat{(t_n)}}\right)}$. Then, $\psi \left(\overline{\frac{\widehat{((m_n))}}{\widehat{(s_n)}}} \right) = \not{/}\psi \left(\overline{\frac{\widehat{((m'_n))}}{\widehat{(s'_n)}}} \right)$.

2. Let us show that ψ is a morphism of $\widehat{S}^{-1}\widehat{A}$-modules:

a. Let $\overline{\frac{\widehat{((m_n))}}{\widehat{(s_n)}}}, \overline{\frac{\widehat{((m'_n))}}{\widehat{(s'_n)}}} \in \widehat{S}^{-1}(M/N)$, and show that
$\psi \left(\overline{\frac{\widehat{((m_n))}}{\widehat{(s_n)}}} \bar{+} \overline{\frac{\widehat{((m'_n))}}{\widehat{(s'_n)}}} \right) = \not{/}\psi \left(\overline{\frac{\widehat{((m_n))}}{\widehat{(s_n)}}} \right) \bar{\widehat{+}} \psi \left(\overline{\frac{\widehat{((m'_n))}}{\widehat{(s'_n)}}} \right)$.
We have
$\psi \left(\overline{\frac{\widehat{((m_n))}}{\widehat{(s_n)}}} \bar{+} \overline{\frac{\widehat{((m'_n))}}{\widehat{(s'_n)}}} \right) = \not{/}\psi \left(\frac{\widehat{(x_n)} \centerdot \overline{\widehat{((m_n))}} \bar{+} \widehat{(y_n)} \centerdot \overline{\widehat{((m'_n))}}}{\widehat{(y_n)} \widehat{\times} \widehat{(s'_n)}} \right)$ with
$\widehat{(x_n)} \widehat{\times} \widehat{(s_n)} = \not{/} \widehat{(y_n)} \widehat{\times} \widehat{(s'_n)}$, where $\widehat{(x_n)}, \widehat{(y_n)} \in \widehat{S}$
$= \not{/} \overline{\left(\frac{\widehat{(x_n)} \widehat{\times} \widehat{((m_n))} \widehat{+} \widehat{(y_n)} \widehat{\times} \widehat{((m'_n))}}{\widehat{(y_n)} \widehat{\times} \widehat{(s'_n)}} \right)}$
$= \not{/} \overline{\left(\frac{\widehat{(x_n)} \widehat{\times} \widehat{((m_n))}}{\widehat{(y_n)} \widehat{\times} \widehat{(s'_n)}} \right)} \bar{\widehat{+}} \not{/} \overline{\left(\frac{\widehat{(y_n)} \widehat{\times} \widehat{((m'_n))}}{\widehat{(y_n)} \widehat{\times} \widehat{(s'_n)}} \right)}$
$= \not{/} \overline{\left(\frac{\widehat{(x_n)} \widehat{\times} \widehat{((m_n))}}{\widehat{(x_n)} \widehat{\times} \widehat{(s_n)}} \right)} \bar{\widehat{+}} \not{/} \overline{\left(\frac{\widehat{(y_n)} \widehat{\times} \widehat{((m'_n))}}{\widehat{(y_n)} \widehat{\times} \widehat{(s'_n)}} \right)}$ by $\widehat{(x_n)} \widehat{\times} \widehat{(s_n)} = \not{/} \widehat{(y_n)} \widehat{\times} \widehat{(s'_n)}$
$= \not{/} \overline{\left(\frac{\widehat{(m_n)}}{\widehat{(s_n)}} \right)} \bar{\widehat{+}} \not{/} \overline{\left(\frac{\widehat{(m'_n)}}{\widehat{(s'_n)}} \right)} = \not{/}\psi \left(\overline{\frac{\widehat{((m_n))}}{\widehat{(s_n)}}} \right) \bar{\widehat{+}} \psi \left(\overline{\frac{\widehat{((m'_n))}}{\widehat{(s'_n)}}} \right)$.
Then, $\psi \left(\overline{\frac{\widehat{((m_n))}}{\widehat{(s_n)}}} \bar{+} \overline{\frac{\widehat{((m'_n))}}{\widehat{(s'_n)}}} \right) = \not{/}\psi \left(\overline{\frac{\widehat{((m_n))}}{\widehat{(s_n)}}} \right) \bar{\widehat{+}} \psi \left(\overline{\frac{\widehat{((m'_n))}}{\widehat{(s'_n)}}} \right)$.

b. Let $\frac{\widehat{(a_n)}}{\widehat{(t_n)}} \in \widehat{A}$, $\overline{\frac{\widehat{((m_n))}}{\widehat{(s_n)}}} \in \widehat{S}^{-1}(M/N)$, and show that
$\psi \left(\frac{\widehat{(a_n)}}{\widehat{(t_n)}} \centerdot \overline{\frac{\widehat{((m'_n))}}{\widehat{(s_n)}}} \right) = \frac{\widehat{(m_n)}}{\widehat{(t_n)}} \centerdot \psi \left(\overline{\frac{\widehat{((m_n))}}{\widehat{(s_n)}}} \right)$.
We have :
$\psi \left(\frac{\widehat{(a_n)}}{\widehat{(t_n)}} \centerdot \overline{\frac{\widehat{((m_n))}}{\widehat{(s_n)}}} \right) = \not{/}\psi \left(\frac{\widehat{(y_n)} \centerdot \overline{\widehat{((m'_n))}}}{\widehat{(x_n)} \widehat{\times} \widehat{(t_n)}} \right)$ with $\widehat{(x_n)} \widehat{\times} \widehat{(a_n)} = \widehat{(y_n)} \widehat{\times} \widehat{(s_n)}$ where
$(\widehat{(x_n)}, \widehat{(y_n)}) \in \widehat{S} \times \widehat{A}$
$= \not{/}\psi \left(\overline{\frac{\widehat{((y_n)} \widehat{\times} \widehat{(m'_n))}}{\widehat{(x_n)} \widehat{\times} \widehat{(t_n)}}} \right) = \not{/} \left(\overline{\frac{\widehat{((y_n)} \widehat{\times} \widehat{(m'_n))}}{\widehat{(x_n)} \widehat{\times} \widehat{(t_n)}}} \right)$
$= \not{/} \left(\overline{\frac{\widehat{((y_n)} \widehat{\times} \widehat{(m'_n))}}{\widehat{(x_n)} \widehat{\times} \widehat{(t_n)}}} \right) = \not{/} \overline{\left(\frac{\widehat{(a_n)}}{\widehat{(t_n)}} \centerdot \frac{\widehat{((m_n))}}{\widehat{(s_n)}} \right)} = \not{/} \frac{\widehat{(a_n)}}{\widehat{(t_n)}} \centerdot \overline{\left(\frac{\widehat{((m_n))}}{\widehat{(s_n)}} \right)} = \not{/} \frac{\widehat{(a_n)}}{\widehat{(t_n)}} \centerdot \psi \left(\overline{\frac{\widehat{((m_n))}}{\widehat{(s_n)}}} \right)$.

Thus, $\psi\left(\frac{\widehat{(a_n)}}{\widehat{(t_n)}} \cdot \frac{\overline{\widehat{((m'_n))}}}{\widehat{(s_n)}}\right) = \frac{\widehat{(m_n)}}{\widehat{(t_n)}} \cdot \psi\left(\frac{\overline{\widehat{((m_n))}}}{\widehat{(s_n)}}\right)$.

Then ψ is as morphism of modules.

3. Let us show that ψ is bijective:

   a. Let us show that ψ is injective.

   Let $\frac{\overline{\widehat{((m_n))}}}{\widehat{(s_n)}} \in Ker(\psi)$, then $\psi\left(\frac{\overline{\widehat{((m_n))}}}{\widehat{(s_n)}}\right) = \overline{\left(\frac{\widehat{((m_n)}}{\widehat{(s_n)}}\right)} = \bar{\bar{\widehat{0}}}_{\widehat{S}^{-1}(\widehat{M})/\widehat{S}^{-1}(\widehat{N})} \Rightarrow$ $\frac{\widehat{(m_n)}}{\widehat{(s_n)}} = \widehat{0} \Rightarrow \widehat{(m_n)} = \widehat{0}$ by $\widehat{(s_n)} \neq \widehat{0} \Rightarrow \frac{\overline{\widehat{((m_n))}}}{\widehat{(s_n)}} = \frac{\bar{\widehat{0}}}{\widehat{(s_n)}} = \bar{\bar{\widehat{0}}}_{\widehat{S}^{-1}\left(\widehat{M}/\widehat{N}\right)}$, and thus $Ker(\psi) = \left\{\bar{\bar{\widehat{0}}}_{\widehat{S}^{-1}\left(\widehat{M}/\widehat{N}\right)}\right\}$. Then, ψ is injective.

   b. Let us show that ψ is surjective.

   It is evident that ψ is surjective because for all $\overline{\left(\frac{\widehat{((m_n)}}{\widehat{(s_n)}}\right)} \in \widehat{S}^{-1}(\widehat{M})/\widehat{S}^{-1}(\widehat{N})$, there is $\frac{\overline{\widehat{((m_n))}}}{\widehat{(s_n)}} \in \widehat{S}^{-1}\left(\widehat{M}/\widehat{N}\right)$ such that $\psi\left(\frac{\overline{\widehat{((m_n))}}}{\widehat{(s_n)}}\right) = \overline{\left(\frac{\widehat{((m_n)}}{\widehat{(s_n)}}\right)}$.

   Then, ψ is bijective.

We conclude that the $\widehat{S}^{-1}\widehat{A}$-module $\widehat{S}^{-1}\left(\widehat{M}/\widehat{N}\right)$ is isomorphic to $\widehat{S}^{-1}(\widehat{M})/\widehat{S}^{-1}(\widehat{N})$.

Corollary 9.9 *Let $(A, (I_n)_{n\in\mathbb{N}})$ be a filtered ring where $(I_n)_{n\in\mathbb{N}}$ is a sequence of left ideals, S a saturated multiplicative subset of A that satisfies the left Ore conditions (resp. invariant, resp. central) and*

$$\widehat{S} = \left\{\widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0} \; et \; \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S\right\}$$

the set of classes of Cauchy sequences of A with values in S that do not converge to 0. *Then, the $\widehat{S}^{-1}\widehat{A}$-module $\widehat{S}^{-1}\left(\widehat{A}/\widehat{I}\right)$ is isomorphic to $\widehat{S}^{-1}(\widehat{A})/\widehat{S}^{-1}(\widehat{I})$.*

Theorem 9.10 *Let $(A, (I_n)_{n\in\mathbb{N}})$ be a filtered ring where $(I_n)_{n\in\mathbb{N}}$ is a sequence of left ideals, S a saturated multiplicative subset of A that satisfies the left Ore conditions (resp., invariant, resp., central), and*

$$\widehat{S} = \left\{\widehat{(x_n)} \in \widehat{A} \mid \widehat{(x_n)} \neq \widehat{0} \; et \; \exists n_0 \in \mathbb{N}, n \geq n_0, x_n \in S\right\}$$

the set of classes of Cauchy sequences of A with values in S that do not converge to 0. *And N an S-saturated submodule in a left A-module $(M, (M_n)_{n\in\mathbb{N}})$, and then:*

1. $\widehat{S^{-1}(M/N)} \cong \widehat{S^{-1}(M)/S^{-1}(N)} \cong \widehat{S}^{-1}(\widehat{M})/\widehat{S}^{-1}(\widehat{N})$.
2. $\widehat{S^{-1}(A/I)} \cong \widehat{S^{-1}(A)/S^{-1}(I)} \cong \widehat{S}^{-1}(\widehat{A})/\widehat{S}^{-1}(\widehat{I})$.

Proof

1. Let us consider the correspondence:

$$\begin{array}{rcl} \varphi : \widehat{S^{-1}(M/N)} & \to & \widehat{S}^{-1}(\widehat{M})/\widehat{S}^{-1}(\widehat{N}) \\ \widehat{\left(\frac{\overline{m_n}}{s_n}\right)} & \mapsto & \overline{\left(\frac{\widehat{(m_n)}}{\widehat{(s_n)}}\right)} \end{array}.$$

a. Let us show φ is a map.

Let $\widehat{\left(\frac{\overline{m_n}}{s_n}\right)}, \widehat{\left(\frac{\overline{m'_n}}{s'_n}\right)} \in \widehat{S^{-1}(M/N)}$ such that $\widehat{\left(\frac{\overline{m_n}}{s_n}\right)} = \widehat{\left(\frac{\overline{m'_n}}{s'_n}\right)}$ and show that $\varphi\left(\widehat{\left(\frac{\overline{m_n}}{s_n}\right)}\right) = \varphi\left(\widehat{\left(\frac{\overline{m'_n}}{s'_n}\right)}\right)$.

We have:
$\widehat{\left(\frac{\overline{m_n}}{s_n}\right)} = \widehat{\left(\frac{\overline{m'_n}}{s'_n}\right)} \Rightarrow \left(\frac{\overline{m_n}}{s_n} - \frac{\overline{m'_n}}{s'_n}\right) \to 0$, and then $\forall n \in \mathbb{N}$ there is $x_n, y_n \in S$ such that $\left(\frac{x_n\overline{m_n} - y_n\overline{m'_n}}{x_n s_n}\right) \to 0$ with $x_n s_n = y_n s'_n$.

But $\left(\frac{1}{x_n s_n}\right) \times \left(\frac{x_n s_n}{1}\right) = 1_{S^{-1}A}$, and thus $\left(\frac{1}{x_n s_n}\right) \times \left(\frac{x_n\overline{m_n} - y_n\overline{m'_n}}{x_n s_n}\right) \to 0 \Rightarrow \left(\frac{x_n\overline{m_n} - y_n\overline{m'_n}}{1}\right) \to 0 \Rightarrow (x_n\overline{m_n} - y_n\overline{m'_n}) \to 0 \Rightarrow (x_n m_n - y_n m'_n) \to 0$.

Let $(x_n), (y_n)$ be the sequences of general terms x_n, y_n, respectively, and since $x_n s_n = y_n t_n$, where $x_n, y_n \in S$, thus $(x_n) \times (s_n) = (y_n) \times (s'_n)$, where $(x_n), (y_n) \in \mathcal{C}(S)$.

Then, we have

$$\begin{cases} (x_n m_n - y_n m'_n) \to 0 \\ (x_n) \times (s_n) = (y_n) \times (s'_n) \end{cases} \text{ with } (x_n), (y_n) \in \mathcal{C}(S)$$

$$\Rightarrow \begin{cases} (x_n) \times (m_n) - (y_n) \times (m'_n) \to 0 \\ (x_n) \times (s_n) = (y_n) \times (s'_n) \end{cases} \text{ with } (x_n), (y_n) \in \mathcal{C}(S)$$

$$\Rightarrow \begin{cases} \widehat{(x_n)}\widehat{\times}\widehat{(m_n)} = \widehat{(y_n)}\widehat{\times}\widehat{(m'_n)} \\ \widehat{(x_n)}\widehat{\times}\widehat{(s_n)} = \widehat{(y_n)}\widehat{\times}\widehat{(s'_n)} \end{cases} \text{ with } \widehat{(x_n)}, \widehat{(y_n)} \in \widehat{S}$$

$$\Rightarrow \frac{\widehat{(x_n)}\widehat{\times}\widehat{(m_n)}}{\widehat{(x_n)}\widehat{\times}\widehat{(s_n)}} = \frac{\widehat{(y_n)}\widehat{\times}\widehat{(m'_n)}}{\widehat{(y_n)}\widehat{\times}\widehat{(s'_n)}} \Rightarrow \frac{\widehat{(m_n)}}{\widehat{(s_n)}} = \frac{\widehat{(m'_n)}}{\widehat{(s'_n)}} \Rightarrow \overline{\left(\frac{\widehat{(m_n)}}{\widehat{(s_n)}}\right)} = \overline{\left(\frac{\widehat{(m'_n)}}{\widehat{(s'_n)}}\right)} \Rightarrow$$

$$\varphi\left(\widehat{\left(\frac{\overline{m_n}}{s_n}\right)}\right) = \varphi\left(\widehat{\left(\frac{\overline{m'_n}}{s'_n}\right)}\right).$$

b. Let us show that φ is a module morphism:

i. Let $\widehat{\left(\frac{\overline{m_n}}{s_n}\right)}, \widehat{\left(\frac{\overline{m'_n}}{s'_n}\right)} \in \widehat{S^{-1}(M/N)}$, and show that

$$\varphi\left(\widehat{\left(\frac{\overline{m_n}}{s_n}\right)}\,\widehat{\bar{+}}\,\widehat{\left(\frac{\overline{m'_n}}{s'_n}\right)}\right) = \varphi\left(\widehat{\left(\frac{\overline{m_n}}{s_n}\right)}\right)\widehat{\bar{+}}\,\varphi\left(\widehat{\left(\frac{\overline{m'_n}}{s'_n}\right)}\right).$$

We have:
$\varphi\left(\widehat{\left(\frac{\overline{m_n}}{s_n}\right)}\,\widehat{\bar{+}}\,\widehat{\left(\frac{\overline{m'_n}}{s'_n}\right)}\right) =/ \varphi\left(\widehat{\left(\frac{\overline{m_n}}{s_n} + /\frac{\overline{m'_n}}{s'_n}\right)}\right)$, and then $\forall n \in/ \mathbb{N}$, there is $x_n, y_n \in /S$ such that
$\varphi\left(\widehat{\left(\frac{\overline{m_n}}{s_n}\right)}\,\widehat{\bar{+}}\,\widehat{\left(\frac{\overline{m'_n}}{s'_n}\right)}\right) =/ \varphi\left(\widehat{\left(\frac{\overline{m_n}}{s_n} + /\frac{\overline{m'_n}}{s'_n}\right)}\right) =/ \varphi\left(\widehat{\left(\frac{x_n\overline{m_n}+y_n\overline{m'_n}}{x_n s_n}\right)}\right)$ with $x_n s_n =/ y_n s'_n$.

Let (x_n), (y_n) be the sequences of general terms x_n, y_n, respectively, and since $x_n s_n =/ y_n s'_n$, where $x_n, y_n \in S$, thus $(x_n)\times(s_n) = (y_n)\times(s'_n)$, where $(x_n), (y_n) \in \mathcal{C}(P) \Rightarrow \widehat{(x_n)}\widehat{\times}\widehat{(s_n)} =/ \widehat{(y_n)}\widehat{\times}\widehat{(s'_n)}$, where $\widehat{(x_n)}, \widehat{(y_n)} \in \widehat{S}$.

Then, we have:
$\varphi\left(\widehat{\left(\frac{\overline{m_n}}{s_n}\right)}\,\widehat{\bar{+}}\,\widehat{\left(\frac{\overline{m'_n}}{s'_n}\right)}\right) =/ \varphi\left(\widehat{\left(\frac{x_n\overline{m_n}+y_n\overline{m'_n}}{x_n s_n}\right)}\right) =/ \overline{\left(\frac{\widehat{(x_n m_n+y_n m'_n)}}{\widehat{(x_n s_n)}}\right)} =/ \overline{\left(\frac{\widehat{(x_n)}\widehat{\cdot}\widehat{(m_n)}\widehat{+}\widehat{(y_n)}\widehat{\cdot}\widehat{(m'_n)}}{\widehat{(x_n)}\widehat{\times}\widehat{(s_n)}}\right)}$
$=/ \overline{\left(\frac{\widehat{(x_n)}\widehat{\cdot}\widehat{(m_n)}}{\widehat{(x_n)}\widehat{\times}\widehat{(s_n)}}\widehat{+}\frac{\widehat{(y_n)}\widehat{\cdot}\widehat{(m'_n)}}{\widehat{(x_n)}\widehat{\times}\widehat{(s_n)}}\right)} =/ \overline{\left(\frac{\widehat{(x_n)}\widehat{\cdot}\widehat{(m_n)}}{\widehat{(x_n)}\widehat{\times}\widehat{(s_n)}}\widehat{+}\frac{\widehat{(y_n)}\widehat{\cdot}\widehat{(m'_n)}}{\widehat{(y_n)}\widehat{\times}\widehat{(s'_n)}}\right)} =/ \overline{\left(\frac{\widehat{(m_n)}}{\widehat{(s_n)}}\widehat{+}\frac{\widehat{(m'_n)}}{\widehat{(s'_n)}}\right)} =/ \overline{\left(\frac{\widehat{(m_n)}}{\widehat{(s_n)}}\right)}\,\widehat{\bar{+}}\,/\overline{\left(\frac{\widehat{(m'_n)}}{\widehat{(s'_n)}}\right)}$
$=/ \varphi\left(\widehat{\left(\frac{\overline{m_n}}{s_n}\right)}\right)\widehat{\bar{+}}\varphi\left(\widehat{\left(\frac{\overline{m'_n}}{s'_n}\right)}\right)$.

ii. Let $\widehat{\left(\frac{a_n}{s_n}\right)} \in \widehat{S^{-1}A}$, $\widehat{\left(\frac{\overline{m_n}}{t_n}\right)} \in /\widehat{S^{-1}(M/N)}$, and show that
$\varphi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\widehat{\centerdot}\widehat{\left(\frac{\overline{m_n}}{t_n}\right)}\right) = \widehat{\left(\frac{a_n}{s_n}\right)}\widehat{\bar{\centerdot}}\varphi\left(\widehat{\left(\frac{\overline{m_n}}{t_n}\right)}\right)$.

We have:
$\varphi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\widehat{\centerdot}\widehat{\left(\frac{\overline{m_n}}{t_n}\right)}\right) =/ \varphi\left(\widehat{\left(\frac{a_n}{s_n}\centerdot\frac{\overline{m_n}}{t_n}\right)}\right)$, and then $\forall n \in /\mathbb{N}$, there is $z_n, w_n \in /S$ such that
$\varphi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\widehat{\centerdot}\widehat{\left(\frac{\overline{m_n}}{t_n}\right)}\right) =/ \varphi\left(\widehat{\left(\frac{a_n}{s_n}\centerdot\frac{\overline{m_n}}{t_n}\right)}\right) =/ \varphi\left(\widehat{\left(\frac{z_n\overline{m_n}}{w_n s_n}\right)}\right)$ with $w_n a_n =/ z_n t_n$.

Let (z_n), (w_n) be the sequences of general terms z_n, w_n, respectively, and since $w_n a_n =/ z_n t_n$, where $w_n \in/ S$, $z_n \in/ A$, $\forall n \in /\mathbb{N}$, thus $(w_n) \times/ (a_n) =/ (z_n) \times/ (t_n)$, where $(w_n) \in /\mathcal{C}(S)$, $(z_n) \in/ \mathcal{C}(A) \Rightarrow/ \widehat{(w_n)}\widehat{\times}\widehat{(a_n)} = \widehat{(z_n)}\widehat{\times}\widehat{(t_n)}$, where $\widehat{(w_n)} \in \widehat{S}$, $\widehat{(z_n)} \in \widehat{A}$.

Therefore, we have:
$\varphi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\widehat{\centerdot}\widehat{\left(\frac{\overline{m_n}}{t_n}\right)}\right) =/ \varphi\left(\widehat{\left(\frac{z_n\overline{m_n}}{w_n s_n}\right)}\right) =/ \overline{\left(\frac{\widehat{(z_n)}\widehat{\cdot}\widehat{(m_n}}{\widehat{(w_n)}\widehat{\times}\widehat{(s_n)}}\right)}$ with $\widehat{(w_n)}\widehat{\times}\widehat{(a_n)} = \widehat{(z_n)}\widehat{\times}\widehat{(t_n)}$ where $\widehat{(w_n)} \in \widehat{S}$, $\widehat{(z_n)} \in \widehat{A}$.

$\Rightarrow \varphi\left(\widehat{\left(\frac{a_n}{s_n}\right)}\hat{\cdot}\widehat{\left(\frac{\overline{m_n}}{t_n}\right)}\right) = \overline{\left(\frac{\widehat{(z_n)}\cdot\widehat{(m_n}}{\widehat{(w_n)}\hat{\times}\widehat{(s_n)}}\right)} = \overline{\left(\frac{\widehat{(a_n)}}{\widehat{(s_n)}}\hat{\bullet}\frac{\widehat{(m_n)}}{\widehat{(t_n)}}\right)} = \overline{\left(\frac{\widehat{(a_n)}}{\widehat{(s_n)}}\right)}\bar{\hat{\bullet}}\overline{\left(\frac{\widehat{(m_n)}}{\widehat{(t_n)}}\right)} =$
$\widehat{\left(\frac{a_n}{s_n}\right)}\bar{\hat{\bullet}}\overline{\left(\frac{\widehat{(m_n)}}{\widehat{(t_n)}}\right)}$
$= \widehat{\left(\frac{a_n}{s_n}\right)}\bar{\hat{\bullet}}\overline{\left(\frac{\widehat{(m_n)}}{\widehat{(t_n)}}\right)} = \widehat{\left(\frac{a_n}{s_n}\right)}\bar{\hat{\bullet}}\varphi\left(\widehat{\left(\frac{\overline{m_n}}{t_n}\right)}\right).$

iii. We have $1_{\widehat{S^{-1}(M/N)}} = \widehat{\left(\frac{\bar{1}}{1}\right)}$, and thus

$$\varphi(1_{\widehat{S^{-1}(M/N)}}) = \varphi\left(\widehat{\left(\frac{\bar{1}}{1}\right)}\right) = \overline{\left(\frac{\hat{1}}{\hat{1}}\right)} = \overline{\left(1_{\widehat{S}^{-1}(\widehat{M})}\right)} = 1_{\widehat{S}^{-1}(\widehat{M})/\widehat{S}^{-1}(\widehat{N})}.$$

c. Let us show that φ is bijective:

i. Let us show that φ is injective.

Let $\widehat{\left(\frac{\overline{m_n}}{s_n}\right)} \in Ker(\varphi)$, and then $\varphi\left(\widehat{\left(\frac{\overline{m_n}}{s_n}\right)}\right) = \bar{\hat{0}} \Rightarrow \overline{\left(\frac{\widehat{(m_n)}}{\widehat{(s_n)}}\right)} = \bar{\hat{0}} \Rightarrow$
$\frac{\widehat{(m_n)}}{\widehat{(s_n)}} \in \widehat{S}^{-1}(\widehat{N}) \Rightarrow \widehat{(m_n)} \in \widehat{N}, \widehat{(s_n)} \in \widehat{S} \Rightarrow m_n \in N, s_n \in S, \forall n \in \mathbb{N} \Rightarrow$
$\overline{m_n} = \bar{0}, s_n \in S, \forall n \in \mathbb{N}.$

As $\left(\frac{\overline{m_n}}{s_n}\right)$ is the sequence of general term $\frac{\overline{m_n}}{s_n}$ and $\overline{m_n} = \bar{0}, \forall n \in \mathbb{N}$, then $\widehat{\left(\frac{\overline{m_n}}{s_n}\right)} = \widehat{\left(\frac{\bar{0}}{s_n}\right)} = \hat{\bar{0}} \Rightarrow Ker(\varphi) = \left\{\hat{\bar{0}}\right\}$. Therefore, we deduce that φ is injective.

ii. Let us show that φ is surjective.

Let $\overline{\left(\frac{\widehat{(m_n)}}{\widehat{(s_n)}}\right)} \in \widehat{S}^{-1}(\widehat{M})/\widehat{S}^{-1}(\widehat{N}) \Rightarrow \frac{\widehat{(m_n)}}{\widehat{(s_n)}} \in \widehat{S}^{-1}(\widehat{M}) \Rightarrow$
$\widehat{(m_n)} \in \widehat{M}, \widehat{(s_n)} \in \widehat{S} \Rightarrow$
$m_n \in M, s_n \in S, \forall n \in \mathbb{N} \Rightarrow \overline{m_n} \in M/N, s_n \in S, \forall n \in \mathbb{N}.$

Let $\left(\frac{\overline{m_n}}{s_n}\right)$ be the sequence with general term $\frac{\overline{m_n}}{s_n} \in S^{-1}(M/N)$, and then $\widehat{\left(\frac{\overline{m_n}}{s_n}\right)} \in \widehat{S^{-1}(M/N)}$. This implies that for any $\overline{\left(\frac{\widehat{(m_n)}}{\widehat{(s_n)}}\right)} \in \widehat{S}^{-1}(\widehat{M})/\widehat{S}^{-1}(\widehat{N})$, there is $\widehat{\left(\frac{\overline{m_n}}{s_n}\right)} \in \widehat{S^{-1}(M/N)}$ such that $\varphi\left(\widehat{\left(\frac{\overline{m_n}}{s_n}\right)}\right) = \overline{\left(\frac{\widehat{(m_n)}}{\widehat{(s_n)}}\right)}$, demonstrating the surjectivity of φ.

We conclude that φ is an isomorphism of modules. Therefore, we have $\widehat{S^{-1}(M/N)} \cong \widehat{S}^{-1}(\widehat{M})/\widehat{S}^{-1}(\widehat{N})$.

d. On the other hand, we have: $\widehat{S^{-1}(M/N)} \cong \widehat{S^{-1}(M)/S^{-1}(N)}$, and indeed, let us consider the correspondence

$$\begin{array}{rcl} \Psi : S^{-1}(M/N) & \rightarrow & S^{-1}(M)/S^{-1}(N) \\ \frac{m}{s} & \mapsto & \frac{\bar{m}}{s} \end{array}$$

is an isomorphism of modules according to [2].

Thus, we have $S^{-1}(M/N) \simeq S^{-1}(M)/S^{-1}(N) \Rightarrow / \widehat{S^{-1}(M/N)} \simeq \widehat{S^{-1}(M)/S^{-1}(N)}$.

Finally, we have $\widehat{S^{-1}(M/N)} \cong \widehat{S^{-1}(M)/S^{-1}(N)} \cong \widehat{S}^{-1}(\widehat{M})/\widehat{S}^{-1}(\widehat{N})$.

2. To show that $\widehat{S^{-1}(A/I)} \cong \widehat{S^{-1}(A)/S^{-1}(I)} \cong \widehat{S}^{-1}(\widehat{A})/\widehat{S}^{-1}(\widehat{I})$, it suffices to consider the ring A as an A-module.

References

1. M.B. Maaouia. *Thèse de doctorat de 3éme cycle sur Localisation et enveloppe plate dans un duo-anneaux*. Faculté des sciences et techniques, UCAD, Dakar, Juillet, 2003.
2. M.B. Maaouia. *Thèse d'état*. Faculté des sciences et techniques, UCAD, Dakar, Avril, 2011.
3. M.B. Maaouia and M. Sanghare. *Enveloppe plate sur un duo-anneau*. International Journal of Contemporary Mathematical Sciences, Vol.6, 2011, no. 18, 869-886.
4. M.B. Maaouia and M. Sanghare. *Localisations dans les duo-anneaux*. Afrika Matematika, Série 3, Volume 20, pp. 163-179, 2009.
5. J.P. Lafon. *Algèbre commutative: langages géométrique et algébrique*. Hermann, 1977.
6. G. Renault. *Algèbre non commutative*. Gauthier-Villars, Paris-Bruxelles-Montréal, 1975.

Chapter 10
On S-Lifting Semimodules over Semirings

Moussa Sall, Landing Fall, and Djiby Sow

Abstract The notion of lifting module is well studied in rings and module theory. Recently, many concepts in rings and modules were introduced in semirings and semimodules such us radical of semiring, projective covers of semimodules, and superfluous subsemimodules. In this chapter, we introduce the notion of s-lifting semimodules, and we study their properties.

Keywords Subtractive semimodule · S-Lifting semimodule · Local direct summand · Small subsemimodule

Introduction

Extending semimodules are generalization of injective semimodules and, dually, lifting semimodules generalize projective supplemented semimodules. Every module has an injective hull, but it does not necessarily have a projective cover. This creates a certain asymmetry in the duality between extending modules and lifting modules because in any module M there exists complements for any submodule N, but, by contrast, supplements for N in M need not exist. The terms extending and lifting were coined by Oshiro [2]. Several results concerning lifting modules have appeared in the literature in recent years. In addition, progress in other branches of module theory has also had a flow-on effect, providing further enrichment to lifting module theory.

Moreover, the notion of lifting modules is largely developed by J. Clark, C. Lomp, N. Vanaja, and R. Wisbauer [1]; thus, the study of their results in theory semimodules is very interesting.

M. Sall · L. Fall (✉) · D. Sow
Département de Mathématiques et Informatique, Faculté des Sciences et Techniques, Université Cheikh Anta Diop, Dakar, Sénégal
e-mail: moussa18.sall@ucad.edu.sn; landing.fall@ucad.edu.sn; djiby.sow@ucad.edu.sn

C. T. Gueye et al. (eds.), *Mathematics of Computer Science, Cybersecurity and Artificial Intelligence*, Springer Proceedings in Mathematics & Statistics 468,
https://doi.org/10.1007/978-3-031-66222-5_10

Let M be a semimodule over a semiring R. An equivalence relation ρ on M is an R-congruence relation if and only if

$[m\rho m^{'}$ and $n\rho n^{'}] \Rightarrow [(m+n)\rho(m^{'}+n^{'})$ and $(rm)\rho(rm^{'})]$ $\forall m, m^{'}, n, n^{'} \in M$ and $r \in R$.

So R-congruence relation ρ is trivial if $m\rho m^{'} \Leftrightarrow m = m^{'}$.

Consider the subsemimodule N of M. Then N induces on M an R-congruence relation $\equiv_N$, known as the Bourne relation defined by $\forall m, m^{'} \in M; m \equiv_N m^{'} \Leftrightarrow \exists n, n^{'} \in N$ such that $m + n = m^{'} + n^{'}$.

The set of all the equivalences classes modulo "$\equiv_N$" denoted by M/N is such that $(M/N, +, .)$ is an R-semimodule which is called quotient semimodule where the operations are defined by $''+''$: $\overline{m} + \overline{m'} = \overline{m + m'}$ and $''.''$: $r\overline{m} = \overline{rm}$.

Let $M_1,\ M_2$ be two subsemimodules of M.

- M is a direct weak sum of M_1 and M_2 (denoted $M = M_1\overline{\oplus}M_2$) if $M = M_1 + M_2$ and $M_1 \cap M_2 = \{0\}$.
- M is a direct strong sum of M_1 and M_2 (denoted by $M = M_1 \oplus M_2$) if and only if $M = M_1 + M_2$ and the restriction "$\equiv_{M_1}$" to M_2 and the restriction "$\equiv_{M_2}$" to M_1 are trivial.
- A subsemimodule K of M is called a direct summand if there exists $K' \leq M$ such that $M = K \oplus K'$.

  A subsemimodule N of M is subtractive (= k-subsemimodule) if $\forall\ x,\ y \in M$, $(x + y \in N,\ y \in N) \Rightarrow x \in N$, and in addition, M is called subtractive if every subsemimodule of M is subtractive.
- A subsemimodule N of M is called a fully invariant subsemimodule of M if for every endomorphism φ of M, $\varphi(N) \subseteq N$.

In this chapter we study some properties of lifting semimodules over semirings. Indeed, we prove that an R-semimodule M is s-lifting if and only if for every subtractive subsemimodule N of M, there is a strong direct sum $M = M_1 \oplus M_2$ such that $M_1 \leq N$ and $N \cap M_2 \ll M$ (i.e., $N \cap M_2$ is small in M).

In addition we prove that if M is subtractive and $M = M_1 \oplus M_2 \oplus \ldots \oplus M_n$ is a finite strong direct sum, such that every supplement subsemimodule of M is fully invariant, then M is s-lifting if and only if it is amply supplemented, and $\forall i,\ M_i$ is s-lifting. And to finish, considering the R-semimodule $M = \bigoplus_{i \in I} M_i$ which is a subtractive and amply supplemented, we prove that if for every coclosed subsemimodule K of M, with $M = K + M_i$ or $M = K + \bigoplus_{i \neq j \in I} M_j$, is a direct summand of M, then for every supplement $K^{'}$ of M_i or $\bigoplus_{i \neq j \in I} M_j$, in M, $M/K^{'}$ is s-lifting and $K^{'}$ is a direct summand of M. This chapter is organized as follows:

- Section "Introduction" is about the basic notions, where more notions are defined.
- In section "Basic Notions", we study the notions of s-lifting semimodule.
- In section "S-Lifting Semimodules", we study the direct sums of s-lifting semimodules.

In the following, R is always an associative, commutative semiring with unit and $1_R \neq 0_R$, and the direct summands are strong direct summands.

Basic Notions

Let M be an R-semimodule and N, H, L three subsemimodules of M. We have the following definitions.

- A proper subsemimodule S of M is called a small subsemimodule of M if for all subsemimodules T of M, $S + T = M$ implies that $T = M$. It is indicated by the notation $S \ll M$ and $Rad(N) = \sum_{K \ll N} K$. If S is not small in M, we denote $S \not\ll M$.

  The semimodule M is called hollow if every proper subsemimodule of M is small in M.
- The subsemimodule N of M is called a supplement of L in M if $N + L = M$ and $N \cap L \ll N$. In addition, if N is subtractive, it is trivial to see that N is a supplement of L in M if and only if it is minimal with the property of $N+L = M$.

  The subsemimodule N of an R-semimodule M is called a weak supplement of L if $N + L = M$ and $N \cap L \ll M$ (see [1]).

  If every subsemimodule of M has a supplement (respectively, a weak supplement), then M is called a supplemented semimodule (respectively, weakly supplemented semimodule).

  M is amply supplemented if $M = L + N$ implies there exists a supplement K of L such that $K \leq N$.
- If $H \leq N$ and $N/H \ll M/H$, then H is called a coessential subsemimodule of N in M, and it is denoted by $H \leq^{ce} N$.

  The subsemimodule N of M is coclosed in M (denoted by $N \leq^{cc} M$) if N has no proper coessential subsemimodule in M.

  H is called an s-closure of N in M if H is coessential subsemimodule of N and H is coclosed in M.
- The R-semimodule M is called simple if it has no nontrivial subsemimodules.

  In addition, if M is a direct sum of simple semimodules, we say that it is semisimple.
- An internal direct sum $\bigoplus_I A_i$ of subsemimodules of M is called a local direct sum of M if, given any finite subset F of the index set I, the direct sum $\bigoplus_{i \in F} A_i$ is a direct summand of M.

In the following we give an example of subtractive subsemimodule, weakly direct sum, and strong direct sum.

Example 10.1 Let $R = \{0; 1\}$ be the Boole semiring and the set $M = \{0; 1; a; b\}$. Define on M the operations " $+$ " and " $\times$ " as follows:

$0 + 0 = 0;\ 1 + 1 = 1 + a = a + 1 = 1 + b = b + 1 = a + b = b + a = 1.$
$a + 0 = 0 + a = a + a = a;\ b + 0 = 0 + b = b + b = b.$

$0 \times 0 = 0 \times 1 = 1 \times 0 = 0 \times a = 0 \times b = 0$; $1 \times 1 = 1$; $1 \times a = a$; $1 \times b = b$.

Then $(M, +, \times)$ is a commutative left R-semimodule. In addition, we have:

- Subtractivity: $\{0; a\}$ is a subtractive subsemimodule of M, but $\{0; 1; a\}$ is not subtractive (because $1 + b = 1 \in \{0; 1; a\}$, $1 \in \{0; 1; a\}$ and $b \notin \{0; 1; a\}$).
- Weakly direct sum: $M = \{0; a\} + \{0; 1; b\}$, $\{0; a\} \cap \{0; 1; b\} = \{0\}$ and $1 = 0 + 1 = a + b$. Since $a \neq 0$ and $b \neq 1$, the decomposition of 1 is not unique, and hence $M = \{0; a\}\overline{\oplus}\{0; 1; b\}$.
- Strong direct sum: $M = \{0; a\}+\{0; b\}$, there does not exist $x, y \in \{0; a\} \mid 0+x = b + y$, therefore $m \equiv_{\{0;a\}} m' \Leftrightarrow m = m'$, $\forall\, m, m' \in \{0; b\}$, and hence the restriction of $\equiv_{\{0;a\}}$ to $\{0; b\}$ is trivial. Similarly, the restriction of $\equiv_{\{0;b\}}$ to $\{0; a\}$ is trivial.

  Thus $M = \{0; a\}\oplus\{0; b\}$.

S-Lifting Semimodules

Definition 10.1 Let M be an R-semimodule. M is called lifting if for every subsemimodule N of M there exists a direct summand K of M such that $K \leq N$ and $N/K \ll M/K$.

Definition 10.2 An R-semimodule M is called s-lifting if for every subtractive subsemimodule N of M there exists a direct summand K of M such that $K \leq N$ and $N/K \ll M/K$.

Example 10.2 Let $R = \{0; 1\}$ be the Boole semiring and the set $M = \{0; 1; a; b\}$.

(1) Define on M the same operations " $+$ " and " $\times$ " as in Example 10.1.

    Then $(M, +, \times)$ is a lifting R-semimodule.

    Indeed, the only subsemimodules of M are $\{0\}$, $\{0, 1\}$, $\{0, a\}$, $\{0, b\}$, $\{0, 1, a\}$, $\{0, 1, b\}$, and M. Clearly, $\{0, 1\} \ll M$, $\{0, 1, a\}/\{0, a\} \ll M/\{0, a\}$, $\{0, 1, b\}/\{0, b\} \ll M/\{0, b\}$, and since $\{0\}$, $\{0, a\}$, and $\{0, b\}$ are the direct summands of M (see Example 10.1), then $(M, +, \times)$ is lifting.

(2) Define on M the operations " $*/$" and "." as follows:

$$0 * 1 = 1 * 0 = 1 * 1 = 1 * a =/a * 1 =/1 * b = b * 1 = 1.$$
$$a * a = a * 0 = 0 * a = a;\ \ b */b = b * 0 = 0 * b =/b;\ \ a * b = b * a = 0.$$
$$a.0 = 0.a = b.0 = 0.b = 0.0 = 0.1 =/1.0 = 0;\ \ 1.1 =/1;\ \ 1.a =/a;\ \ 1.b =/b.$$

Then $(M, *, .)$ is an R-semimodule which is not s-lifting.

Indeed, consider the subsemimodule $N = \{0, a, b\}$ of M. Clearly, N is subtractive and the only direct summand of M contained in N is $\{0\}$. Since $N + \{0, 1\} = M$ and $\{0, 1\} \neq M$, $N \not\ll M$, therefore $N/\{0\} \not\ll M/\{0\}$, and hence $(M, *, .)$ is not s-lifting.

(3) Define on M the operations as follows: $0_R = 0_M$, $1_R = 1_M = 1$, $1+1 = a+1 = 1+a = 1+b = b+1 = a+b = b+a = a+a = b+b = 1$; $b+0 = 0+b = b$; $a+0 = 0+a = a$; $a.0 = 0.a = b.0 = 0.b = 0$; $1.a = a.1 = a$;
$1.b = b.1 = b$.
Then $(M, +, .)$ is an R-semimodule S-lifting, but it is not lifting.
Indeed, clearly, the only subtractive subsemimodule of M is $\{0\}$, and, then M is s-lifting. The only direct summand of M contained in $\{0; 1; a\}$ is $\{0\}$. Since $M = \{0; 1; a\} + \{0; 1; b\}$ and $\{0; 1; b\} \neq M$, then $\{0; 1; a\}/\{0\} \not\ll M/\{0\}$; therefore, M is not lifting.

Similar as in [1], 3.6, by subtractivity, we can define a coclosed subsemimodule as follows:

Definition 10.3 A subsemimodule L of a subtractive R-semimodule M is coclosed if and only if for any proper subsemimodule $K \subseteq L$, there is a subsemimodule N of M such that $L + N = M$ and $N + K \neq M$.

Proposition 10.1 *Let M be a weakly supplemented subtractive R-semimodule and $N \leq M$. The following statements are equivalent:*

(1) *N is a supplement subsemimodule.*
(2) *N is coclosed in M.*
(3) *For all X, $X \leq N$ and $X \ll M$ implies that $X \ll N$.*

Proof (1) $\Rightarrow$ (2): Let N be a supplement subsemimodule of M, then there exists a subsemimodule L of M such that N is minimal of the propriety $N + L = M$ (because since M is subtractive, then N is subtractive). Let $K \leq N$ such that $N/K \ll M/K$.

$$\begin{aligned} N + L = M &\Rightarrow N + (K + L) = M \\ &\Rightarrow (N + (K + L))/K = N/K + (K + L)/K = M/K \\ &\Rightarrow (K + L)/K = M/K \text{ (because } N/K \ll M/K) \\ &\Rightarrow K + L = M. \end{aligned}$$

Since N is minimal with the propriety $N + L = M$, we conclude that $N = K$; therefore, N is coclosed.

(2) $\Rightarrow$ (3): Let $X \leq N$ such that $X \ll M$. Let $X' \leq N$ such that $X + X' = N$.
Assume that $X' \neq N$. By Definition 10.3, there is $K \leq M$ such that $M = N + K$ and $X' + K \neq M$. And hence $M = N + K = X + X' + K$; since $X \ll M$, $X' + K = M$, which is a contradiction; therefore, $X' = N$. Thus $X \ll N$.

(3) $\Rightarrow$ (1): Since M is a weakly supplemented, there exists a subsemimodule H of M such that $N + H = M$ and $N \cap H \ll M$.

Since $N \cap H \leq N$ and $N \cap H \ll M$, from (3), we conclude that $N \cap H \ll N$.

Now we have $N + H = M$ and $N \cap H \ll N$; therefore N is a supplement of H in M. Thus N is a supplement subsemimodule, and hence these statements are equivalent. □

Remark 10.1 Every supplement subsemimodule of a subtractive semimodule M is coclosed in M.

Lemma 10.1 *A subtractive R-semimodule M is amply supplemented if and only if it is weakly supplemented and every subsemimodule has an s-coclosure in M.*

Proof $\Rightarrow$) Let $K \leq M$. Since M is amply supplemented, M is weakly supplemented and supplemented. Let L be a supplement of K in M. Then $M = K + L$ with $K \cap L \ll L$ and so $K \cap L \ll M$. Since M is amply supplemented, there exists a supplement N of L such that $N \subseteq K$. Then $M = N + L$ and $N \cap L \ll N$. We claim that N is an s-coclosure of K in M. As a supplement subsemimodule, N is coclosed in M (from Remark 10.1). Since M is subtractive, K is subtractive, and hence $K = N + (K \cap L)$ (by modularity condition).

Let us prove that $K/N \ll M/N$. Let H be a subsemimodule of M containing N such that $K/N + H/N = M/N$. Then $(K + H)/N = M/N$; since M is subtractive, $K + H = M$ and hence $(K \cap L) + H = M$ (because $K = N + (K \cap L)$ and $N \subseteq H$); therefore, $H = M$ (because $K \cap L \ll M$) and so $H/N = M/N$. Thus $K/N \ll M/N$.

Therefore N is an s-coclosure of K in M.

$\Leftarrow$) Assume that M is weakly supplemented, and every subsemimodule of M has an s-coclosure in M. Let $K, L \leq M$ with $M = K + L$. Since M is weakly supplemented, there is a weak supplement T of K such that $T \subseteq L$. Then $M = K + T$ and $K \cap T \ll M$. Suppose N is an s-coclosure of T in M, then $T/N \ll M/N$ and $N \leq^{cc} M$; since M is subtractive and $T/N + (K + N)/N = (K + T)/N = M/N$, then not only $K + N = M$, but also $N \cap K \ll M$, and from proposition 10.1, $N \cap K \ll N$ (because $N \leq^{cc} M$). Thus N is a supplement of K in M such that $N \subseteq L$.

Similarly, we verify that K contains a supplement of L in M, and hence M is amply supplemented. □

Proposition 10.2 (see [1], 22.3)

Let M be a subtractive R-semimodule. Then the following statements are equivalents:

(1) *M is s-lifting.*
(2) *M is lifting.*
(3) *M is amply supplemented, and every supplement (namely, coclosed) subsemimodule is a direct summand of M.*

Proof (1) $\Leftrightarrow$ (2) Trivial.

(2) $\Rightarrow$ (3) Let N be a supplement subsemimodule of L in M. Then $N + L = M$ and $N \cap L \ll N$. By assumption, there exists a direct summand K of M such that $K \leq N$ and $N/K \ll M/K$. By Remark 10.1, N is coclosed in M (because M is subtractive) and hence $N = K$; therefore, N is a strong direct summand of M.

Let L be a subsemimodule of M. Then there exists a strong direct summand H of M such that $H \leq L$ and $L/H \ll M/H$. So $M = H \oplus' H'$ (a strong direct sum) for some $H' \leq M$. We show that H' is a supplement of L in M. Clearly, $M =' L + H'$. We claim that $L \cap' H' \ll H'$.

Let $H'' \leq H'$ such that $H' = (L \cap H') + H''$. Then $M = L + H' = L + H''$(because $H' = (L \cap H') + H''$), and hence $M/H = L/H + (H + H'')/H$ and so $(H + H'')/H = M/H$(because $L/H \ll M/H$), therefore, $M = H + H''$. Since $H'' \leq H'$ and $M = H \oplus H'$, then if $h' = h +' h''$ with $h \in H, h' \in H'$, and $h'' \in H''$, we have $h' \equiv_H h''$, and then $h' = h'' \in H''$. Thus $H' = H''$ and so $L \cap H' \ll H'$. Thus L has a supplement, namely H' in M, and hence M is supplemented; therefore, it is weakly supplemented. By assumption, it is very easy to verify that every subsemimodule of M has an s-coclosure in M, and hence, from Lemma 10.1, M is amply supplemented.

(3) $\Rightarrow$ (2) Reciprocally, we suppose M is amply supplemented, and every supplement is a strong direct summand of M.

Let N be a subsemimodule of M. Since M is subtractive and amply supplemented, N has an s-closure in M (see Lemma 10.1); therefore, there exists $K \leq N$ such that $K \leq^{cc} M$ and $N/K \ll M/K$ so, by Lemma 10.1, K is supplement, and by assumption, K is a strong direct summand of M. Thus, for every subsemimodule N of M, there exists a strong direct summand K of M such that $K \leq N$ and $N/K \ll M/K$, and hence M is lifting. □

Lemma 10.2 *Let I, J be R-semimodules, $f :' I \longrightarrow' J$ be an isomorphism, and S a subsemimodule of I. Then $S \ll I$ if and only if $f(S) \ll J$.*

Proof Let $S \ll I$ and H be a subsemimodule of J such that $f(S) + H = J$.

Then $f^{-1}(f(S) + H) = f^{-1}(J) = I$, and hence $f^{-1}(f(S)) + f^{-1}(H) = S + f^{-1}(H) = I$. Since $S \ll I$, then $f^{-1}(H) = I = f^{-1}(J)$, and hence $H = J$, so $f(S) \ll J$.

Conversely, considering the isomorphism f^{-1} and from what precedes, if $f(S) \ll J$, then $S \ll I$. □

Theorem 10.1 *An R-semimodule M is s-lifting if and only if for every subtractive subsemimodule N of M, there is a decomposition $M = M_1 \oplus M_2$ such that $M_1 \subseteq N$ and $N \cap M_2 \ll M$.*

Proof Assume that M is s-lifting. Let N be a subtractive subsemimodule of M. Since M is s-lifting, there is a strong direct summand M_1 of M such that $M_1 \leq N$ and $N/M_1 \ll M/M_1$.

Since M_1 is a strong direct summand of M, there is $M_2 \leq M$ such that $M = M_1 \oplus M_2$.

Moreover, we want to verify that $N \cap M_2 \ll M$. Consider the obvious isomorphism

$$f : M/M_1 \longrightarrow M_2$$

$$\bar{x} \longmapsto x_2$$

with $x = x_1 + x_2$, where $x_1 \in M_1$ and $x_2 \in M_2$. It is very easy to verify that $f(N/M_1) = N \cap M_2$.

Indeed, let $x_2 \in f(N/M_1)$. Then there is $\bar{x} \in N/M_1$ such that $f(\bar{x}) = x_2$ with $x = x_1 + x_2,\ x_1 \in M_1,\ x_2 \in M_2$.

$\bar{x} \in N/M_1 \Rightarrow \exists\, x' \in N$ such that $\bar{x} = \overline{x'}$.

$\bar{x} = \overline{x'} \Rightarrow \exists\, m_1, m_2 \in M_1$ such that $x + m_1 = x' + m_2$. Since $m_2 \in M_1 \subseteq N$, $x' + m_2 \in N$; therefore $x + m_1 \in N$ and so $x \in N$ (because N is subtractive and $m_1 \in M_1 \subseteq N$).

Hence $x_1 + x_2 = x \in N$, therefore $x_2 \in N$ (because N is subtractive and $x \in M_1 \subseteq N$) whence $f(N/M_1) \subseteq N$. Since $f(N/M_1) \subseteq M_2$, we conclude that $f(N/M_1) \subseteq N \cap M_2 \quad (*)$

Let $x_2 \in N \cap M_2$. Then $x_2 \in M_2$; therefore there is a unique $\bar{x} \in M/M_1$ such that $x = x_1 + x_2$, where $x_1 \in M_1$, and $f(\bar{x}) = x_2$ (because f is an isomorphism).

$x_2 \in N,\ x_1 \in M_1 \subseteq N \Rightarrow x = x_1 + x_2 \in N$, therefore $\bar{x} \in N/M_1$, and hence $x_2 \in f(N/M_1)$, so $N \cap M_2 \subseteq f(N/M_1)$ $(**)$
$(*)$ and $(**)$ imply that $f(N/M_1) = N \cap M_2$. Since $N/M_1 \ll M/M_1$ and f is an isomorphism, $f(N/M_1) \ll M_2 \subseteq M$ (from Lemma 10.2), therefore $f(N/M_1) \ll M$, and hence $N \cap M_2 \ll M$.

In sum, we have $M = M_1 \oplus M_2$ such that $M_1 \subseteq N$ and $N \cap M_2 \ll M$.

Conversely, if N is a subtractive subsemimodule of M, then there is a decomposition $M = M_1 \oplus M_2$ such that $M_1 \subseteq N$, $N \cap M_2 \ll M$, and in considering the reciprocal bijection f^{-1} of f, we have $f^{-1}(N \cap M_2) = N/M_1$. Since $N \cap M_2 \ll M$ and f^{-1} is a bijection, then $N/M_1 \ll M/M_1$ (by Lemma 10.2). Thus M is s-lifting. □

Proposition 10.3 *Let M be an s-lifting R-semimodule. Then every subtractive subsemimodule N of M can be written as $N = N_1 \oplus N_2$ with N_1 a strong direct summand of M and $N_2 \ll M$.*

Proof Let N be a subtractive subsemimodule of M. Since M is s-lifting, by Theorem 10.1, there is a decomposition $M = M_1 \oplus M_2$ such that $M_1 \subseteq N$ and $N \cap M_2 \ll M$. We consider $N_1 = M_1,\ N_2 = N \cap M_2$. It is clear that N_1 is a strong direct summand of M and $N_2 \ll M$. In addition, $N = M_1 + N \cap M_2$ (because N is subtractive and $M_1 \subseteq N$); therefore $N = N_1 + N_2$.

It is very trivial to see that $N = N_1 \oplus N_2$. Indeed let $x, y \in N_1$ such that $x \equiv_{N_2} y$. Then there exist $n_2, n_2' \in N_2$ such that $x + n_2 = y + n_2'$. Since $n_2, n_2' \in M_2$ (because $N_2 \subseteq M_2$), $x \equiv_{M_2} y$; therefore $x = y$ (because $x, y \in N_1 = M_1$ and $M = M_1 \oplus M_2$), so "$(\equiv_{N_2})_{|N_1}$" is trivial.

Similarly, we prove "$(\equiv_{N_1})_{|N_2}$" is trivial, and hence $N = N_1 \oplus N_2$ with N_1 a strong direct summand of M and $N_2 \ll M$. □

Theorem 10.2 *Let M be a subtractive R-semimodule. Then the following assertions are equivalent:*

(1) *M is s-lifting.*

(2) *For every subsemimodule N of M, there is a decomposition $M = M_1 \oplus M_2$ such that $M_1 \subseteq N$ and $N \cap M_2 \ll M$.*
(3) *Every subsemimodule N of M can be written as $N = N_1 \oplus N_2$ with N_1 a strong direct summand of M and $N_2 \ll M$.*

Proof By Theorem 10.1, we have (1) $\Leftrightarrow$ (2), and by proposition 10.3 we have (1) $\Rightarrow$ (3). It stands to prove that (3) $\Rightarrow$ (1).

Let N be a subsemimodule of M. By assumption, $N = N_1 \oplus N_2$ with N_1 is a strong direct summand of M and $N_2 \ll M$. We show now $N/N_1 \ll M/N_1$.

Let H be a subsemimodule of M such that $N_1 \subseteq H$ and $N/N_1 + H/N_1 = M/N_1$.

$$\begin{aligned}
\text{Hence } N/N_1 + H/N_1 = M/N_1 &\Rightarrow (N+H)/N_1 = M/N_1 \\
&\Rightarrow N + H = M(since\ M\ is\ subtractive) \\
&\Rightarrow N_1 + N_2 + H = M \ \ (\text{because } \ N = N_1 \oplus N_2) \\
&\Rightarrow N_2 + H = M \ \ (\text{because } \ N_1 \subseteq H) \\
&\Rightarrow H = M \ \ (\text{because } \ N_2 \ll M) \\
&\Rightarrow H/N_1 = M/N_1 \\
&\Rightarrow N/N_1 \ll M/N_1.
\end{aligned}$$

And since N_1 is a strong direct summand of M, we conclude that M is lifting.

Thus, since M is subtractive, it is s-lifting (by definition 10.2). □

Remark 10.2 The condition of subtractivity is fundamental in this theorem (see the following example).

Example 10.3 Let $R = \{0; 1\}$ be the Boole semiring and the set $M = \{0; 1; a; b\}$.

Define on M the operations as follows: $0_R = 0_M$, $1_R = 1_M = 1$, $1 + 1 = 1 + a = 1 + b = a + b = 0$; $a + a = a + 0 = 0 + a = a$; $b + b = b + 0 = 0 + b = b$; $a.0 = 0.a = b.0 = 0.b = 0$; $1.a = a$; $1.b = b$.

Then $(M, +, .)$ is R-semimodule S-lifting (see Example 10.2). Clearly, M is not subtractive because $\{0; 1\}$ is not subtractive (because $1 + a = 0 \in \{0; 1\}$, but $a \notin \{0; 1\}$).

But, since the only strong direct summand of M contained in $\{0; 1\}$ is $\{0\}$ and M$\neq \{0; 1\}$, then there is no decomposition $M = M_1 \oplus M_2$ such that $M_1 \subseteq \{0; 1\}$ and $\{0; 1\} \cap M_2 \ll M$.

Corollary 10.1 *Every hollow subtractive semimodule is s-lifting.*

Proof If N is a subsemimodule of a hollow subtractive semimodule M, we have $N \ll M$ and $N = \{0\} \oplus N$. By Theorem 10.2, M is s-lifting. □

Theorem 10.3 *Any coclosed subsemimodule (and hence every strong direct summand) of a subtractive s-lifting semimodule is s-lifting.*

Proof Let N be a coclosed subsemimodule of an s-lifting subtractive R-semimodule M. Clearly, N is subtractive. Let K be a subsemimodule of N. Hence K is a subsemimodule of M, and since M is s-lifting, by Theorem 10.2, K can be written as $K = K_1 \oplus K_2$ with K_1 a direct summand of M and $K_2 \ll M$. Since $K_1 \leq N$ is a strong direct summand of M, K_1 is a strong direct summand of N (because $M = K_1 \oplus K_1' \Rightarrow\prime N = N \cap M = K_1 + N \cap K_1'$). By Proposition 10.2, M is amply supplemented (because M is s-lifting), and hence M is supplemented, so it is weakly supplemented; therefore, by Lemma 10.1, $K_2 \ll N$ (because $N \leq^{cc} M$ by proposition10.1). Hence $K = K_1 \oplus K_2$ with K_1 is a strong direct summand of N and $K_2 \ll N$, and hence, by Theorem 10.2, N is s-lifting. □

Proposition 10.4 *Let M be an s-lifting R-semimodule such that $Rad(N) \ll N$ for every subtractive subsemimodule $N \subseteq M$.*

Then every subtractive local strong direct summand of M is a strong direct summand of M.

Proof Let $\bigoplus_K M_k$ be a subtractive local strong direct summand of M, and set

$$N = \bigoplus_K M_k.$$

Since M is s-lifting and N is a subtractive subsemimodule of M, by Proposition 2.4, N can be written as $N = T \oplus S$, where T is a strong direct summand of M and $S \ll M$.

Let $x \in S$. We have $xR \subseteq\prime \bigoplus_{f \in F} M_f$, where F is a finite subset of K. Then, since $xR \ll M$ (because $xR \subseteq S \ll M$) and $\bigoplus_{k \in F} M_k$ is a strong direct summand, let us prove that $xR \ll \bigoplus_{k \in F} M_k$. Assume that there exists $L \leq M$ such that $xR + L =\prime \bigoplus_{k \in F} M_k$. Since $M =\prime \bigoplus_{k \in F} M_k \oplus M'$, we have $xR + L + M' = M$. This implies $xR + (M' + L) =\prime M$. Since $xR \ll M$, we have $M' + L = M$ and so $L =\prime \bigoplus_{k \in F} M_k$; therefore,

$$\sum_{y \in S} yR \ll Rad(N)$$

and so $S \subseteq Rad(N)$. However, by hypothesis, $Rad(N) \ll N$ and hence $S \ll N$; therefore, $N = T$ (because $N = T \oplus S$) and so N is a strong direct summand of M. □

Direct Sums of S-Lifting Semimodules

In this section we look at the question of when s-lifting is preserved by finite (or infinite) direct sums. Moreover one of the most interesting questions concerning s-lifting semimodules is when a finite (or infinite) direct sum of s-lifting semimodules is also s-lifting.

Finite Direct Sums of S-Lifting Semimodules

Theorem 10.4 *Let $K \subset V$ be two subtractive subsemimodules of semimodule M.*

(1) *If K is a supplement in M, then K is supplement in V.*
(2) *If V is a supplement, the following are equivalent:*

   (a) *K is supplement in V.*
   (b) *K is supplement in M.*

Proof

(1) If K is a supplement in M, there is some subsemimodule $P \subset M$ with $K + P = M$ and $K^{'} + P \neq M$ for a proper $K^{'} \subset K$.

By modularity, $K + (P \cap V) = V$ (because V is subtractive and $K \subset V$).

Assume $L + P \cap V = V$ for some subsemimodule $L \subset K$. Then the equality $V + P = K + P = M$ yields $M = L + (P \cap V) + P = L + P$ and so, by the minimality of K, we get $L = K$. Thus K is a supplement of P in V.

(2) By assumption there is a subsemimodule $N \subset M$ with $V + N = M$, and for any proper subsemimodule $V^{'} \subset V$, $V^{'} + N \neq M$.

$(a) \Rightarrow (b)$: There exists $L \subset V$ with $K + L = V$ and $K^{''} + L \neq V$ for any proper subsemimodules $K^{''} \subset K$. Thus $K + L + N = V + N = M$. Assume that for some $K^{'} \subseteq K$, $K^{'} + L + N = M$. Put $V^{''} = K^{'} + L$, then $V^{''} + N = M$. Since $K^{'} \subseteq K$ and $K + L = V$, then $V^{''} \subseteq V$.

We have $V^{''} \subseteq V$ and $V^{''} + N = M$; therefore, $V^{''} = V$ by the minimality of V (see above)

Hence $V = V^{''} = K^{'} + L$ and $V = K + L$; thus $K^{'} = K$ by minimality. Therefore K is a supplement of $L + N$ in M.

$(b) \Rightarrow (a)$ follows from (1). □

Theorem 10.5 *Let $M = M_1 \oplus \ldots \oplus M_l$ be a finite strong direct sum of semimodules M_i. Assume that M is subtractive, and every supplement subsemimodule of M is fully invariant. Then M is s-lifting if and only if it is amply supplemented and M_i is s-lifting for all $1 \leq i \leq l$.*

Proof We suppose that M is s-lifting. By Proposition 10.2, M is amply supplemented, and by Theorem 10.2, M_i is s-lifting, $\forall i \in \{1; \ldots; l\}$.

Conversely, we suppose that M is amply supplemented and M_i is s-lifting for all $1 \leq i \leq l$.

Let N be a supplement subsemimodule of M. Let us prove that $N = \bigoplus_{i=1}^{l} (N \cap M_i)$.

We know that $\bigoplus_{i=1}^{l} (N \cap M_i) \subseteq N$. Let us prove that $N \subseteq \bigoplus_{i=1}^{l} (N \cap M_i)$.

Since $M = \bigoplus_{i=1}^{l} M_i$, we have for $n \in N \leq M$, $n = \sum_{i=1}^{l} n_i$ with $n_i \in M_i$. It remains to prove that $n_t \in N$ for all $1 \leq t \leq l$. Consider the canonical projection $P_{M_t} : M \longrightarrow M_t$, and then $P_{M_t}(n) = \sum_{i=1}^{l} P_{M_t}(n_i)$. We know that $P_{M_t}(n_t) = n_t$ and $P_{M_t}(n_i) = 0$ for all $i \neq t$. Then $P_{M_t}(n) = n_t$. Since N is fully invariant, then $n_t = P_{M_t}(n) \in N$. Thus $N = \bigoplus_{i=1}^{l} (N \cap M_i)$.

It is clear that $N \cap M_i$ is a subsemimodule of M_i and $N \cap M_i$ is a supplement subsemimodule in N because it is a strong direct summand of N, and since N is a supplement subsemimodule in M, $N \cap M_i$ is a supplement subsemimodule in M (from Theorem 10.1); therefore, by Theorem 10.1, $N \cap M_i$ is a supplement subsemimodule in M_i, and hence $N \cap M_i$ is a strong direct summand of M_i (because M_i is s-lifting). Hence there is $M_i^{'} \leq M_i$ such that $M_i^{'} \oplus\prime (N \cap M_i) = M_i$. We have now

$$\begin{aligned} M = M_1 \oplus \ldots \oplus M_l &= M_1^{'} \oplus (N \cap M_1) \oplus M_2^{'} \oplus (N \cap M_2) \oplus \ldots \\ &\quad \oplus M_l^{'} \oplus (N \cap M_l) \\ &= (N \cap M_1) \oplus \ldots \oplus (N \cap M_l) \oplus M_1^{'} \oplus \ldots \oplus M_l^{'} \\ &= N \oplus M_1^{'} \oplus \ldots \oplus M_l^{'}. \end{aligned}$$

Hence N is a strong direct summand of M.

So M is amply supplemented, and every supplement subsemimodule of M is a strong direct summand of M; therefore, by Proposition 10.2, M is s-lifting. □

Example 10.4 Consider $(\mathbb{N};\ gcd;\ lcm)$. Set $M = \mathbb{N}$, *and we suppose* $gcd(0, 0) = 0$.

Then M is a subtractive $\mathbb{N}$-semimodule, and every supplement subsemimodule of M is fully invariant.

Clearly, M is an $\mathbb{N}$-semimodule; then, we prove that it is subtractive in showing every subsemimodule of M is a k-subsemimodule.

It is clear that every k-subsemimodule of M is of the form $n\mathbb{N}$, $n \in \mathbb{N}$, and $\{0\}$, M are trivial k-subsemimodules of M (because $\{0\} = 0\mathbb{N}$ and $M = 1\mathbb{N}$).

Let $N \neq \{0\}$ be a subsemimodule of $M = \mathbb{N}$ and $x \in\prime N$.

$N \neq \{0\}$ is a subsemimodule of $M = \mathbb{N}$, and then it has a nonzero minimal element, say m.

$x \in N$ and $m \in N \Rightarrow x +\prime m =\prime gcd(x, m) \in N$.

$gcd(x, m)|m \Rightarrow 0 \neq gcd(x, m) \leq m$, and hence $gcd(x, m) = m$ (because m is a nonzero minimal element of N).

Hence $m|x$, and therefore $x \in m\mathbb{N}$, so $N \subseteq m\mathbb{N}$ (1).

Let $y \in m\mathbb{N}$. Then there exists $\alpha \in \mathbb{N}$ such that $y = m\alpha =\prime lcm(m, \alpha)$.

Since $m \in N$, $\alpha \in \mathbb{N}$, and N is an $\mathbb{N}$-subsemimodule of M, $lcm(m, \alpha) \in N$; therefore $y \in\prime N$.

Hence $m\mathbb{N} \subseteq N$ (2).

(1) and (2) $\Rightarrow N = m\mathbb{N}$; therefore N is a k-subsemimodule of M, and M is subtractive. In addition, since the only supplement subsemimodules of M are M and $\{0\}\prime$which are fully invariant, then every supplement subsemimodule of M is fully invariant.

Arbitrary Direct Sums of S-Lifting Semimodules

Lemma 10.3 *Let K, L, and N be subtractive subsemimodules of an R-semimodule M. Assume that $K + L = M$ and $(K \cap L) + N = M$. Then $K + (L \cap N) = L + (K \cap N) = M$.*

Proof First observe that

$$K + (L \cap N) = K + (L \cap K) + (L \cap N)$$
$$= K + (L \cap ((L \cap K) + N))(\text{because } L \text{ is subtractive and } K \cap L \subset L)$$
$$= K + (L \cap M) = K + L = M.$$

Applying the same arguments to $L + (K \cap N)$ yields $L + (K \cap N) = M$. □

By subtractivity assumption and Lemma 10.3, we have the following proposition:

Proposition 10.5 *Let $K, L,$ and N be subsemimodules of subtractive R-semimodule M.*

If $M = K+L,\ L \subseteq N$, and $N/L \ll M/L$, then $(K\cap N)/(K\cap L) \ll M/(K\cap L)$.

Proof Consider a subsemimodule X such that $K \cap L \subseteq X \subseteq M$ and $M/(K \cap L) = (K \cap N)/(K \cap L) + X/(K \cap L)$. Then $M = X + (K \cap N)$.

By Lemma 10.3, $M = N + (K \cap X)$. Since $N/L \ll M/L$, this implies $M = L + (K \cap X)$. Again from Lemma 10.3, $M = X + (K \cap L)$, and hence $M = X$ (because $K \cap L \subseteq X$).

Thus $(K \cap N)/(K \cap L) \ll M/(K \cap L)$. □

Lemma 10.4 *Let M be a subtractive semimodule and $K \subseteq L \subset M$ be subsemimodules.*

If K is a supplement in M and L/K is a supplement in M/K, then L is a supplement in M.

By subtractivity assumption and Lemma 10.3, the proof is similar to those in module theory. But for the sake of completeness, we give a complete proof on the following.

Proof Let L/K be a supplement of $L^{'}/K$ in M/K, and let K be a supplement of $K^{'}$ in M. Then $M/K = L/K + L^{'}/K$ and $(L/K) \cap (L^{'}/K) \ll L/K$. Moreover $M = K + K^{'}$ and $K \cap K^{'} \ll K \subseteq L$. Hence $M = (L \cap L^{'}) + K^{'}$ and $M = L + L^{'}$ and so $M = L + (K^{'} \cap L^{'})$ (by Lemma 10.3). Now $L = L \cap (K + K^{'}) = K + (L \cap K^{'})$ (because L is subtractive and $K \subseteq L$) and $(L \cap L^{'})/K \ll L/K$ (because $(L \cap L^{'})/K \ll L/K \subset (L/K) \cap (L^{'}/K)$ and $(L/K) \cap (L^{'}/K) \ll L/K$) and so, by Proposition 10.5, $(L \cap L^{'} \cap K^{'})/(K \cap K^{'}) \ll L/(K \cap K^{'})$ and so $L \cap L^{'} \cap K^{'} \ll L$. Thus L is a supplement of $K^{'} \cap L^{'}$ in M. □

Theorem 10.6 *Let $M = \bigoplus_{i \in I} M_i$ be an arbitrary direct sum of R-semimodules M_i $(i \in I,\ |I| \geq 2)$, for some index set, such that M is subtractive and amply supplemented.*

If for every coclosed subsemimodule K of M, with $M = K + M_i$ or $M = K +' \bigoplus_{i \neq j \in I} M_j$, is a strong direct summand of M, then for every supplement K' of M_i or $\bigoplus_{i \neq j \in I} M_j$ in M, M/K' is s-lifting and K' is a strong direct summand of M.

Proof Clearly, lifting and s-lifting are equivalent because M is subtractive.

Assume that for every coclosed subsemimodule K of M, with $M = K + M_i$ or $M =' K +' \bigoplus_{i \neq j \in I} M_j$ is a strong direct summand of M.

Let K' be a supplement of M_i in M. Then, by Lemma 10.1, K' is coclosed in M, and hence, by assumption, K' is a strong direct summand of M. Since M is subtractive and amply supplemented, M/K' is amply supplemented.

Let T/K' be a supplement in M/K' with $K' \subseteq T$. Since K' is a supplement in M, T is a supplement in M (by Lemma 10.4), and hence T is coclosed in M (by Lemma 10.1); therefore, by assumption, T is a strong direct summand of M, and hence T/K' is a strong direct summand of M/K'. Hence M/K' is amply supplemented and every supplement is a strong direct summand; thus, by Proposition 10.2, M/K' is lifting, so it is s-lifting.

With the same arguments, we prove that every supplement K' of $\bigoplus_{i \neq j \in I} M_j$ in M is a strong direct summand of M and M/K' is s-lifting. □

Example 10.5 Consider $(\mathbb{N};\ gcd;\ lcm)$. Set $M = \mathbb{N}$, *and we suppose* $gcd(0, 0) = 0$. Then M is a subtractive and amply supplemented $\mathbb{N}$-semimodule.

Clearly, M is subtractive (see Example 10.4), and since the only supplement subsemimodules of M are M and {0}, M is indecomposable, and hence it is amply supplemented.

References

1. J.Clark,C.Lomp,N.Vanaja,R.Wisbauer, *Lifting modules, frontiers in Mathematics*, Birkhauser Basel 2006.
2. Oshiro, K., Lifting modules, extending modules and their applications to generalized uniserial rings, Hokkaido Math. J. 13, 339–346 (1984).

Chapter 11
A Contribution to the Study of a Class of Noncommutative Ideals Admitting Finite Gröbner Bases

Laila Mesmoudi and Yatma Diop

Abstract Considering a field $\mathbb{K}$ of characteristic 0, the n-variate commutative polynomial ring $\mathbb{K}[x_1, \ldots, x_n]$ over $\mathbb{K}$, the n-variate noncommutative polynomial ring $\mathbb{K}\langle X_1, \ldots, X_n\rangle$ over $\mathbb{K}$, and $\gamma : \mathbb{K}\langle X_1, \ldots, X_n\rangle \longrightarrow \mathbb{K}[x_1, \ldots, x_n]$ the application sending X_i to x_i, Eisenbud et al. proved that for any ideal $\mathcal{I}$ of $\mathbb{K}[x_1, \ldots, x_n]$, the ideal $\mathcal{J} = \gamma^{-1}(\mathcal{I})$ has a finite Gröbner basis.

Y. Diop and D. Sow dealt with the opposite problem and proved that any noncommutative ideal which contains all commutators and has a finite Gröbner basis is a preimage of a commutative ideal by γ.

In this work, we prove that this application γ can be replaced by any surjective homomorphism. Thus we generalize the two results previously cited.

Keywords Finite Gröbner bases · Noetherian **A**-algebra · Surjective homomorphism · Universal property

Introduction

The Gröbner bases theory originated in Buchberger's PhD thesis [3]. Recall that in the context of commutative multivariate polynomial rings $\mathbb{K}[x_1, \ldots, x_n]$ over a field $\mathbb{K}$, Hilbert's basis theorem guarantees the existence of a finite generator set (i.e., basis) for any ideal. But a generator set does not always solve the Ideal Membership Problem (IMP).

In [3], Buchberger proved the existence of a class of finite bases with some particular and interesting properties which overcome the IMP. He called them Gröbner bases in homage to his advisor Wolfgang Gröbner. The history is largely

L. Mesmoudi · Y. Diop (✉)
Department of mathematics and computer sciences, Cheikh Anta Diop University of Dakar, Dakar, Senegal
e-mail: laila.mesmoudi@ucad.edu.sn; yatma.diop@ucad.edu.sn

C. T. Gueye et al. (eds.), *Mathematics of Computer Science, Cybersecurity and Artificial Intelligence*, Springer Proceedings in Mathematics & Statistics 468,
https://doi.org/10.1007/978-3-031-66222-5_11

documented and can be found in many papers and books. Refer to [4] for more details.

Recall that a monomial order on $\mathbb{K}[x_1, \ldots, x_n]$ is a total and well order on the set of monomials of $\mathbb{K}[x_1, \ldots, x_n]$ which is compatible with the multiplication. When a monomial order $\prec$ is fixed on $\mathbb{K}[x_1, \ldots, x_n]$, then the greatest monomial of a non-nil polynomial f with respect to ($w.r.t.$ in short) $\prec$ is called the leading monomial of f and denoted $LM_\prec(f)$. This notion is extended to a set $\mathcal{G}$ of non-nil polynomials. $LM_\prec(\mathcal{G}) = \{LM_\prec(f),\ f \in \mathcal{G}\}$. Then $\mathcal{G}$ is a Gröbner basis of an ideal $\mathcal{I} \subset \mathbb{K}[x]$ $w.r.t.$ a monomial order $\prec$ if it satisfies $\langle LM_\prec(\mathcal{G})\rangle = \langle LM_\prec(\mathcal{I})\rangle$.

Gröbner bases have several other applications. The Polynomial System Solving (PoSSo) is the most known among them. Namely, in the context of polynomial systems, Gröbner bases are an analog of the Gaussian elimination. Their efficiency in the resolution of such systems combined to the natural apparition of these systems in many modelizations (cryptology, robotic, statistics, etc.) is one of the causes of the success of this new mathematical domain.

Gröbner bases were then extended in several ways and several algebras. In [2], Bergman generalized them in nonassociative algebras.

Their extension in the noncommutative polynomial rings $\mathbb{K}\langle X_1, \ldots, X_n\rangle$ over a field $\mathbb{K}$ [2, 8, 9] has as an immediate consequence the lost finiteness, i.e., there exist noncommutative polynomial ideals that do not have a finite Gröbner basis whatever the monomial order one considers. The classical example is the principal ideal $\mathcal{J} = \langle X_1X_2X_1 - X_2X_1\rangle \subset \mathbb{K}\langle X_1, X_2\rangle$. Whatever the field $\mathbb{K}$ and whatever the monomial order, a Gröbner basis of $\mathcal{J}$ must contain the sequence $f_n = X_1X_2^nX_1 - X_2^nX_1,\ n \in \mathbb{N}^*$. It means that $\mathcal{J}$ does not have a finite Gröbner basis.

Hence, one must ask to know how to characterize noncommutative ideals that have a finite Gröbner basis $w.r.t.$ some monomial order. There is no general response to this question. But it is partially solved in [6] and [5].

Previous Works In [6], the authors consider the homomorphism $\gamma : \mathbb{K}\langle X_1, \ldots, X_n\rangle \longrightarrow \mathbb{K}[x_1, \ldots, x_n]$ which replaces X_i by x_i and proved in Theorems 11.1 and 11.2 that if $\mathbb{K}$ is of characteristic 0, then any ideal $\mathcal{J}$ of the type $\mathcal{J} = \gamma^{-1}(\mathcal{I})$ has a finite Gröbner basis. Furthermore, they proposed a nice technique to compute a finite Gröbner basis of a such type of noncommutative ideal $\mathcal{J}$ from one of the corresponding commutative ideals $\mathcal{I}$.

The paper [5] is about the opposite problem. Its authors keep the same application $\gamma : \mathbb{K}\langle X_1, \ldots, X_n\rangle \longrightarrow \mathbb{K}[x_1, \ldots, x_n]$ and proved in Theorem 3.5 that any noncommutative ideal $\mathcal{J}$ which contains all commutators $X_jX_i - X_iX_j,\ i < j$ and has a finite Gröbner basis can be expressed $\mathcal{J} = \gamma^{-1}(\mathcal{I})$ for some ideal $\mathcal{I}$ of $\mathbb{K}[x_1, \ldots, x_n]$.

Our Contribution In this work, we prove that γ can be replaced by any surjective homomorphism. Our reflexion is first of all a mathematical curiosity, but the result can be useful if one finds some surjective homomorphism with some properties. Our proof is based on a construction, and we will take advantage from the well-known universal property.

Organization In section "Preliminaries", we will recall some notions and results that we need to present our contribution in section "Our Contribution".

In all the rest of the chapter, $\mathbf{A}$ is a commutative ring with unity, and $(\mathbf{B}, +_{\mathbf{B}}, \times_{\mathbf{B}})$ and $(\mathbf{C}, +_{\mathbf{C}}, \times_{\mathbf{C}})$ are two $\mathbf{A}$-algebras.

Preliminaries

In this section, we recall some well-known properties. They are essential in some steps of our contribution.

Definition 11.1 Let $f : \mathbf{B} \longrightarrow \mathbf{C}$ be an application.

1. f is called an $\mathbf{A}$-morphism if for any $a, b \in \mathbf{B},\ \alpha \in \mathbf{A}$, we have

   $f(a +_{\mathbf{B}} b) = f(a) +_{\mathbf{C}} f(b).$
   $f(a \times_{\mathbf{B}} b) = f(a) \times_{\mathbf{C}} f(b).$
   $f(1_{\mathbf{A}}) = 1_{\mathbf{B}}.$
   $f(\alpha a) = \alpha f(a).$

2. A morphism $f : \mathbf{B} \longrightarrow \mathbf{C}$ is called an endomorphism if $\mathbf{B} = \mathbf{C}$.
3. A morphism $f : \mathbf{B} \longrightarrow \mathbf{C}$ is called an isomorphism if it is bijective.

Theorem 11.1 (Universal Property of Polynomial Ring) *Let $(b_1, \ldots, x_n) \in \mathbf{B}^n$. Then there exists a unique morphism of $\mathbf{A}$-algebras $\varphi : \mathbf{A}[x_1, \ldots, x_n] \longrightarrow \mathbf{B}$ which satisfies $\varphi(x_i) = b_i$ for any $1 \leq i \leq n$.*

There is a general version of this property in [7, p.124]

Proposition 11.1 *Let f be a surjective endomorphism over a Noetherian ring. Then f is an isomorphism.*

Proof Let $f : \mathbf{A} \longrightarrow \mathbf{A}$ be morphism. Then $(\ker(f^n))_{n\in\mathbb{N}^*}$ is an increasing ideal sequence. Since $\mathbf{A}$ is Noetherian, then there exists $n_0 \in \mathbb{N}$ such that $\ker(f^{n_0}) = \ker(f^n)\ \forall n \geq n_0$.

Let us now prove that f is injective.

Let $x \in \mathbf{A}$ such that $f(x) = 0$. Since f is surjective, then f^n is surjective for any $n \in \mathbb{N}^*$.

Since f^{n_0} is surjective, then there exists $a \in A$ such that $f^{n_0}(a) = x$. Thus, $f^{n_0+1}(a) = f(f^{n_0}(a)) = f(x) = 0$.

Then $a \in \ker(f^{n_0+1}) = \ker(f^{n_0})$. So $x = f^{n_0}(a) = 0$. It follows that f is injective and so it is an isomorphism.

Definition 11.2 Elements $b_1, \ldots, b_n$ of $\mathbf{B}$ are said to generate $\mathbf{B}$ if every element of $\mathbf{B}$ can be expressed as a polynomial in the b_i with coefficients in $\mathbf{A}$. We then write $\mathbf{B} = \mathbf{A}[b_1, \ldots, b_n]$, and we say that $\mathbf{B}$ is an $\mathbf{A}$-algebra of finite type or finitely generated.

Proposition 11.2 ([1, p.81]) *If* $\mathbf{A}$ *is Noetherian, then any* $\mathbf{A}$*-algebra of finite type is Noetherian.*

Our Contribution

Recall that our goal is to generalize the two main results of the papers [6] and [5]. For that, we have first the following result.

In the two following lemmas, the ring $\mathbf{A}$ is supposed to be Noetherian.

Lemma 11.1 *Let* $\{\alpha_1, \ldots, \alpha_n\}$ *be a generator set of* $\mathbf{A}[x_1, \ldots, x_n]$.

Then there exists an automorphism $g : \mathbf{A}[x_1, \ldots, x_n] \longrightarrow \mathbf{A}[x_1, \ldots, x_n]$ *such that* $g(\alpha_i) = x_i \ \forall i$.

Proof Let $(\alpha_1, \ldots, \alpha_n) \in (\mathbf{A}[x_1, \ldots, x_n])^n$. The universal property implies that there exists a unique homomorphism $\alpha : \mathbf{A}[x_1, \ldots, x_n] \longrightarrow \mathbf{A}[x_1, \ldots, x_n]$ such that $\alpha(x_i) = \alpha_i \ \forall i$. By the definition of α, we have $\mathbf{Im}(\alpha) = \mathbf{A}[\alpha_1, \ldots, \alpha_n]$. Also $\mathbf{A}[\alpha_1, \ldots, \alpha_n] = \mathbf{A}[x_1, \ldots, x_n]$ since $\{\alpha_1, \ldots, \alpha_n\}$ is a generator set of $\mathbf{A}[x_1, \ldots, x_n]$. It follows that α is a surjective endomorphism over a Noetherian ring. So α is an isomorphism. Then we take $g = \alpha^{-1}$

Lemma 11.2 *Let* $f : \mathbf{A}\langle X_1, \ldots, X_n\rangle \longrightarrow \mathbf{A}[x_1, \ldots, x_n]$ *be a surjective morphism and* $\gamma : \mathbf{A}\langle X_1, \ldots, X_n\rangle \longrightarrow \mathbf{A}[x_1, \ldots, x_n]$ *which associates* X_i *with* x_i. *Then there exists an isomorphism* $g : \mathbf{A}[x_1, \ldots, x_n] \longrightarrow \mathbf{A}[x_1, \ldots, x_n]$ *which satisfies* $\gamma = g \circ f$.

Proof Let $\alpha_1, \ldots, \alpha_n \in \mathbf{A}[x_1, \ldots, x_n]$ such that $f(X_i) = \alpha_i \ \forall i$. Since f is surjective, then the family $\{\alpha_1, \ldots, \alpha_n\}$ generates the $\mathbf{A}$-algebra $\mathbf{A}[x_1, \ldots, x_n]$. By Lemma 11.1, there exits an isomorphism $g : \mathbf{A}[x_1, \ldots, x_n] \longrightarrow \mathbf{A}[x_1, \ldots, x_n]$ such that $g(\alpha_i) = x_i \ \forall i$.

Thus $\gamma(X_i) = x_i = g(\alpha_i) = g(f(X_i)) = g \circ f(X_i) \ \forall i$. So $\gamma = g \circ f$.

Now we can present our result. Note that the proof is based on the previous lemma and the main results in [6] and [5].

Theorem 11.2 *Let* $\mathbb{K}$ *be a field of characteristic 0 and* $f : \mathbb{K}\langle X_1, \ldots, X_n\rangle \longrightarrow \mathbb{K}[x_1, \ldots, x_n]$ *a surjective homomorphism. Let* $\mathcal{J} \subset \mathbb{K}\langle X_1, \ldots, X_n\rangle$ *be an ideal containing all commutators. Then the two following statements are equivalent.*

1. $\mathcal{J}$ *has a finite Gröbner basis.*
2. *There exists an ideal* $\mathcal{I} \subset \mathbb{K}[x_1, \ldots, x_n]$ *such that* $\mathcal{J} = f^{-1}(\mathcal{I})$.

Proof Let $\mathcal{J} \subset \mathbb{K}\langle X_1, \ldots, X_n\rangle$ be an ideal containing all commutators.

1) $\Longrightarrow$ 2)

Suppose that $\mathcal{J}$ has a finite Gröbner basis. Then Theorem 3.5 of [5] implies that there exists an ideal $\mathcal{I}_0$ of $\mathbb{K}[x_1, \ldots, x_n]$ such that $\mathcal{J} = \gamma^{-1}(\mathcal{I}_0)$.

By Lemma 11.2, we can write $\gamma = g \circ f$, where g is an automorphism of $\mathbb{K}[x_1, \ldots, x_n]$.

$$\begin{aligned}\mathcal{J} = \gamma^{-1}(\mathcal{I}_0) &\Longrightarrow \mathcal{J} = (g \circ f)^{-1}(\mathcal{I}_0)\\ &\Longrightarrow \mathcal{J} = f^{-1}(g^{-1}(\mathcal{I}_0))\\ &\Longrightarrow \mathcal{J} = f^{-1}(\mathcal{I}),\ with\, \mathcal{I} = g^{-1}(\mathcal{I}_0).\end{aligned}$$

2) $\Longrightarrow$ 1).
Suppose that $\mathcal{J} = f^{-1}(\mathcal{I})$ for some ideal $\mathcal{I} \subset \mathbb{K}[x_1, \ldots, x_n]$.
Let $\mathcal{I}_1 = g(\mathcal{I})$. Then $\mathcal{I}_1$ is an ideal since g is an automorphism.

$$\begin{aligned}\mathcal{J} = f^{-1}(\mathcal{I}) &\Longrightarrow \mathcal{J} = f^{-1}(g^{-1}(g(\mathcal{I})))\\ &\Longrightarrow \mathcal{J} = f^{-1}(g^{-1}(\mathcal{I}_1))\\ &\Longrightarrow \mathcal{J} = (g \circ f)^{-1}(\mathcal{I}_1)\\ &\Longrightarrow \mathcal{J} = \gamma^{-1}(\mathcal{I}_1)\end{aligned}$$

$\Longrightarrow$ **Theorem 2** of [6] implies that $\mathcal{J}$ has a finite Gröbner basis.

Conclusion

Now we have seen how to move from γ to a surjective homomorphism f. As we underlined it in the **Introduction** section, this is first of all a mathematical curiosity, but if $\mathcal{J} = \gamma^{-1}(\mathcal{I}) = f^{-1}(\mathcal{I}_0)$, it can be more useful to work with the second expression if the ideal $\mathcal{I}_0$ has some particularities. In other words, our result allows having a choice into how to express the noncommutative ideal which contains all commutators and which has a finite Gröbner basis.

References

1. M. F. ATIYAH, I. G. MACDONALD *Introduction to commutative algebra, University of Oxford*
2. G. M. BERGMAN *The diamond lemma for ring theory, Adv. Math. 29(1978), 178–218.*
3. B. BUCHBERGER *Bruno Buchberger's PhD thesis 1965 : An algorithm for finding the basis elements of the residue class of a zero dimensional ideal, Journal of Symbolic Computation 41(2006) 475–511.*
4. D. COX, J. LITTLE, D. O'SHEA *Ideals, Varieties and Algorithms An Introduction to Computational Algebraic Geometry and Commutative Algebra. Second edition Springer, Undergraduate Texts in Mathematics 1997*
5. Y. DIOP, D. SOW *On finite noncommutative Gröbner bases, Algebra Colloquium 27:3 (2020) 381–388.*
6. D. EISENBUD, I. PEEVA, B. STURMFELS *Noncommutative Gröbner bases for commutative algebras, Proc. Amer. Math. Soc. 126(1998) 687–691*
7. N. JACOBSON *Basic Algebra I Second Edition, Library of Congress Cataloging in Publication Data*

8. T. MORA *An introduction to commutative and non-commutative Gröbner Bases, Journal of Theoretical Computer Science, 13(1994) 131–173*
9. R. B. MURRAY *Free associative algebras, non-commutative Gröbner bases, and universal associative envelopes for non-associative structures, CMUC: Comment. Math. Univ. Carolin. 55, 3(2014) 341–379*

Chapter 12
Construction of Numbers with the Same "Normality" Properties as a Given Number

Khabane Ngom and Ismaila Diouf

Abstract Let x be a positive real number and b an integer greater than or equal to 2. In this chapter, we will construct from the expansion of x in base b an another real number y such that:

– If x is normal in base b, then y is normal in base b.
– If x is simply normal in base b, then y is simply normal in base b.
– If x is abnormal in base b, then y is abnormal in base b.

Keywords Normal numbers · Expansion of x in base b · Permutations · Lebesgue's measure

Introduction

Any positive integer A is written uniquely in the form $A = \sum_{k=0}^{N} a_k b^k$ with $N \in \mathbb{N}$ and, for all $i \in \{0, \ldots, N\}$, $a_i \in \{0, 1, \ldots, b-1\}$. The representation of A in base b is $a_N a_{N-1} \ldots a_1 a_0$. Throughout the rest, for any integer $b \geqslant 2$ such that $b \neq 10$, $\lfloor x \rfloor_b$ represents the integer part of x in base b.

Let x be a positive real number and b an integer, $b \geqslant 2$. There always exists a unique sequence of integers $(x_n)_{n \geqslant 1}$ between 0 and $b-1$ that verifies

$$x = \lfloor x \rfloor + \sum_{k=1}^{+\infty} \frac{x_k}{b^k}.$$

K. Ngom (✉) · I. Diouf
Département de Mathématiques et Informatique, Faculté des Sciences et Techniques, Université Cheikh Anta Diop, Dakar, Senegal
e-mail: khabane.ngom@ucad.edu.sn; ismaila.diouf@ucad.edu.sn

C. T. Gueye et al. (eds.), *Mathematics of Computer Science, Cybersecurity and Artificial Intelligence*, Springer Proceedings in Mathematics & Statistics 468,
https://doi.org/10.1007/978-3-031-66222-5_12

The expression $\lfloor x \rfloor_b, x_1x_2x_3\ldots$ is called the expansion of x in base b. Let $A = \{0, 1, \ldots, b-1\}$ and S be a finite sequence of A. We denote $N(S, n)$ as the number of times that S appears among the first n digits after the comma of the expansion of x in base b. The real x is said:

- **Simply normal** In base b, if $\forall a \in A \quad \lim\limits_{n\to+\infty} \dfrac{N((a), n)}{n} = \dfrac{1}{b}$.
- **Normal** in base b, if it is simply normal in base b^k for any integer $k \geqslant 1$, which means

$$\forall k \geqslant 1, \forall S \in A^k, \lim_{n\to+\infty} \frac{N(S, n)}{n} = \frac{1}{b^k}.$$

- **Abnormal** in base b, if it is not normal in base b.
- **Absolutely normal**, if it is normal in any base.
- **Absolutely abnormal**, if it is abnormal in any base.

In 1909, Borel [6] introduced the concept of normal numbers and proved that almost all numbers are absolutely normal with respect to Lebesgue's measure. Since rational numbers constitute an infinite class of absolutely abnormal numbers, therefore by this theorem irrational numbers are serious candidates for normal and absolutely normal numbers. A large number of irrationals have been proved to be normal or absolutely normal.

In 1933, Champernowne [2] proved that the number $C_{10} = 0.123456789101112 1314\ldots$, formed from the concatenation of consecutive positive integers, is normal in base 10. He even extends this result to any base b. He observes that the number obtained by concatenating the sequence of consecutive positive integers written in any base b represents a b-normal number.

Copeland and Erdös [3] proved in 1946 that the number $0.23571113171923293137\ldots$, obtained by the concatenation of consecutive prime numbers, is normal in base 10. In general, they show that if $(a_n)_{n\geqslant 1}$ is an increasing sequence of positive integers (written in base b) such that for any positive real number $\theta < 1$, $\#\{a_i \leqslant x\} > x^\theta$, for $x \geqslant x_0(\theta)$, then the number $0.a_1a_2a_3\ldots$ is normal in base b.

Nakai and Shiokawa [5] proved that if $f \in \mathbb{R}[X]$ such that $f(x) > 0$ for $x > 0$, then the number $0.\lfloor f(1)\rfloor\lfloor f(2)\rfloor\lfloor f(3)\rfloor\ldots$, where $\lfloor f(n)\rfloor$ represents the integer part of $f(n)$ in base b, is normal in base b. Since we have no shortage of normal numbers, it would be nice to see some abnormal numbers other than the rational numbers.

Bailey and Crandall [1] show that for any $x \in [0, 1]$, $f(x) = \sum\limits_{n=1}^{+\infty} \dfrac{\lfloor nx \rfloor}{2^n}$ is an abnormal number in base 2.

In May 2000, at a survey conference organized by Glynm Harman, Andrew Granville asked about an absolutely abnormal number. In response, Carl Pomerance suggested considering the Liouville number $l = \sum\limits_{n=1}^{+\infty} (n!)^{-n!}$. Recall that a number

β is said to be a Liouville number if, given any integer m, there exists a rational $\frac{p}{q}$ such that $0 < \left|\beta - \frac{p}{q}\right| < \frac{1}{q^m}$. Note that it is known that any Liouville number is transcendent. So far, no one has proven that l is absolutely abnormal.

Intrigued by Granville's question, Martin [4] considered the very fastly growing sequence:

$$d_2 = 2^2, d_3 = 3^2, d_4 = 4^3, d_5 = 5^{16}, d_5 = 5^{16}, d_6 = 6^{30517578125} \dots$$

with the recursive rule

$$d_j = j^{d_{j-1}/(j-1)} \quad (j \geq 3).$$

Then he proved that the number

$$\alpha = \prod_{j=2}^{+\infty}\left(1 - \frac{1}{d_j}\right) = 0.6562499999956991 \quad \underbrace{999\dots999}_{23,747,291,5599s} \quad 85284042016\dots$$

is a Liouville number and in fact an absolutely abnormal number.

More generally, given any sequence of positive integers $n_2, n_3, \dots$, and

$$d_j = j^{n_j d_{j-1}/(j-1)} \qquad (j \geqslant 3)$$

and considering the number

$$\alpha = \prod_{j=2}^{+\infty}\left(1 - \frac{1}{d_j}\right).$$

Martin proved that α is an absolutely abnormal number, thus providing an uncountable family of absolutely abnormal numbers.

Our objective here is, from a positive real number x, to construct other real numbers y having the same normality properties as x, using the expansion of x in base b and the permutation on all the digits after the comma in this expansion.

In section "Introduction" we will provide the necessary background for the demonstration of the main result, and in section "Preliminaries of the Main Result" we will present the main result and its demonstration.

Preliminaries of the Main Result

Remark 1.1 Denote 0_b, the representation of 0 in any base b.

We write $\{x\}$ to designate the fractional part of x, and we thus have $\{x\} = x - \lfloor x \rfloor$. It is easy to establish the following properties:

Proposition 1.1

1. $\lfloor x \rfloor \leqslant x < \lfloor x \rfloor + 1$, $0 \leqslant x - \lfloor x \rfloor < 1$, $x = \lfloor x \rfloor + \theta$ *avec* $0 \leqslant \theta < 1$.
2. $\lfloor x \rfloor + \lfloor y \rfloor \leqslant \lfloor x + y \rfloor \leqslant \lfloor x \rfloor + \lfloor y \rfloor + 1$.

Expansion of a Positive Real Number in Any Base

Let $x \in \mathbb{R}_+$ and b an integer greater than or equal to 2. We will construct two real sequences which will converge to x. Let $P_n = \lfloor b^n x \rfloor$. We construct the sequences α_n and γ_n in the following way:
$\alpha_n = b^{-n} P_n$ and $\gamma_n = b^{-n}(P_n + 1)$.
$P_n = \lfloor b^n x \rfloor \Longrightarrow P_n \leqslant b^n x < P_n + 1$
$\Longrightarrow b^{-n} P_n \leqslant x < b^{-n}(P_n + 1)$
so $\alpha_n \leqslant x < \gamma_n$.

Let us show that the sequence $(\alpha_n)_{n \in \mathbb{N}}$ is increasing.
$\dfrac{\alpha_{n+1}}{\alpha_n} = \dfrac{b^{-n-1} P_{n+1}}{b^{-n} P_n} = \dfrac{P_{n+1}}{b P_n}$ or $P_n \leqslant b^n x$ by multiplying by b, and we have $b P_n \leqslant b^{n+1} x \Longrightarrow \lfloor b P_n \rfloor \leqslant \lfloor b^{n+1} x \rfloor$ because the integer part function is increasing.

Therefore $b P_n \leqslant P_{n+1}$, thus $\dfrac{P_{n+1}}{b P_n} \geqslant 1$ and $\dfrac{\alpha_{n+1}}{\alpha_n} \geqslant 1$, and hence the sequence $(\alpha_n)_{n \in \mathbb{N}}$ is increasing.

Let us show that the sequence $(\gamma_n)_{n \in \mathbb{N}}$ is decreasing.
$\dfrac{\gamma_{n+1}}{\gamma_n} = \dfrac{b^{-n-1}(P_{n+1} + 1)}{b^{-n}(P_n + 1)} = \dfrac{P_{n+1} + 1}{b(P_n + 1)}$.

We know that $P_n = \lfloor b^n x \rfloor$.

Therefore $b^n x < P_n + 1 \Longrightarrow b^{n+1} x < b(P_n + 1)$

$$\begin{aligned}
&\Longrightarrow \lfloor b^{n+1} x \rfloor < b(P_n + 1) \\
&\Longrightarrow P_{n+1} < b(P_n + 1) \\
&\Longrightarrow P_{n+1} + 1 \leqslant b(P_n + 1) \\
&\Longrightarrow \frac{P_{n+1} + 1}{b(P_n + 1)} \leqslant 1 \\
&\Longrightarrow \frac{\gamma_{n+1}}{\gamma_n} \leqslant 1;
\end{aligned}$$

hence the sequence $(\gamma_n)_{n \in \mathbb{N}}$ is decreasing.

In addition, $\gamma_n - \alpha_n = b^{-n}(P_n + 1) - b^{-n} P_n = b^{-n}$.

We have $\lim\limits_{n \to +\infty} (\gamma_n - \alpha_n) = \lim\limits_{n \to +\infty} b^{-n} = 0$. Therefore the sequences $(\alpha_n)_{n \geqslant 0}$ and $(\gamma_n)_{n \geqslant 0}$ are adjacent to the common limit x.

Proposition 1.2

1. *For all* $n \in \mathbb{N}^*$, *the sequence* $c_n = P_n - bP_{n-1}$ *is between* 0 *and* $b-1$.
2. $\forall n \in \mathbb{N}^*, \alpha_n = \alpha_0 + \sum_{k=1}^{n} c_k b^{-k}$.

Proof

1. We have for all $n \in \mathbb{N}^*$, $\gamma_n = b^{-n}(P_n + 1)$. As the sequence $(\gamma_n)_{n \geqslant 0}$ is decreasing, therefore, $\gamma_n \leqslant \gamma_{n-1}$. $b^{-n}(P_n+1) \leqslant b^{-n+1}(P_{n-1}+1)$. This implies that
   $P_n + 1 \leqslant b(P_{n-1} + 1)$; therefore $P_n - bP_{n-1} \leqslant b - 1$. Therefore $c_n = P_n - bP_{n-1} \leqslant b - 1$. Thus $c_n \leqslant b - 1, \forall\, n \geqslant 1$.

   On the other hand, for the second inequality, we use a property of the integer part, namely $\lfloor x + y \rfloor \geqslant \lfloor x \rfloor + \lfloor y \rfloor$ for all real numbers x and y.
   $P_n = \lfloor b^n x \rfloor = \lfloor b(b^{n-1}x) \rfloor$
   $$= \left(\overbrace{\lfloor b^{n-1}x + \cdots\cdots + b^{n-1}x \rfloor}^{btimes} \right)$$
   $$\geqslant \left(\overbrace{\lfloor b^{n-1}x \rfloor + \cdots\cdots + \lfloor b^{n-1}x \rfloor}^{btimes} \right)$$
   $$\geqslant b\lfloor b^{n-1}x \rfloor = bP_{n-1}.$$

   Hence, $c_n = P_n - bP_{n-1} \geqslant 0$.
   Therefore $\forall\, n \in \mathbb{N}^*, 0 \leqslant c_n \leqslant b - 1$.

2. Let us show by induction on $n \in \mathbb{N}^*$ that $\alpha_n = \alpha_0 + \sum_{k=1}^{n} c_k b^{-k}$.

   If $n = 1$, we show that $\alpha_1 = \alpha_0 + c_1 b^{-1}$.
   We have $c_n = P_n - bP_{n-1}$, so $c_1 = P_1 - bP_0$
   $\alpha_0 + c_1 b^{-1} = \alpha_0 + b^{-1}(P_1 - bP_0)$
   $$= \alpha_0 + b^{-1}P_1 - P_0$$
   $$= \alpha_0 - P_0 + b^{-1}P_1$$
   $$= b^{-1}P_1 = \alpha_1, \text{ so the equality is verified at rank 1.}$$

   Assume that the property is true up to n.
   Let us prove that the property holds for $n + 1$.
   $\alpha_n = b^{-n}P_n \Longrightarrow \alpha_{n+1} = b^{-n-1}P_{n+1}$, since $c_n = P_n - bP_{n-1}$,
   we have $P_{n+1} = c_{n+1} + bP_n$, and then $\alpha_{n+1} = b^{-n-1}(c_{n+1} + bP_n)$

$$
\begin{aligned}
&= b^{-n-1}c_{n+1} + b^{-n}P_n \\
&= b^{-n}P_n + b^{-(n+1)}c_{n+1} \\
&= \alpha_n + b^{-(n+1)}c_{n+1} \\
&= \alpha_0 + \sum_{k=1}^{n} c_k b^{-k} + b^{-(n+1)}c_{n+1} \\
&= \alpha_0 + \sum_{k=1}^{n+1} c_k b^{-k}.
\end{aligned}
$$

□

Let us show that the series $\sum_{k\geqslant 1} c_k b^{-k}$ converges.

$\forall\, k \in \mathbb{N}$, $c_k b^{-k} \leqslant (b-1)\cdot b^{-k}$, since the series $\sum_{k\geqslant 1}(b-1)\cdot b^{-k}$ has the same nature (in terms of convergence) as the geometric series $\sum_{k\geqslant 1} b^{-k}$, and the latter is convergent because its reason $\left|\frac{1}{b}\right| < 1$, so $\sum_{k\geqslant 1}(b-1)\cdot b^{-k}$ converges and the same is true for $\sum_{k\geqslant 1} c_k b^{-k}$.

We have $x = \lim_{n\to+\infty} \alpha_n = \lim_{n\to\infty}(\alpha_0 + \sum_{k=1}^{n} c_k b^{-k}) = \alpha_0 + \sum_{k=1}^{+\infty} c_k b^{-k}$.

So

$$x = \lfloor x \rfloor + \sum_{k=1}^{+\infty} c_k b^{-k}.$$

Definition 1.1 The expression $\lfloor x \rfloor_b, c_1c_2c_3c_4\dots$ is called the expansion of x in base b.

Case of Rational Numbers

Theorem 1.1 *A positive real number x is rational if and only if its decimal expansion is finite or infinite periodic.*

Proof Let $x = \frac{p}{q}$ be rational. If the quotient is in irreducible form, that is, p and q are relatively prime, we have several cases.

- If q is the product of a power of 2 and a power of 5, the quotient has a finite decimal expansion. If this is not the case, the quotient has an infinite periodic decimal expansion whose shortest period has a length less than $q-1$ depending only on q. If, in addition, q is prime with 10, this period begins immediately after the comma. When we perform the division of p by q, at each step, there are only q possible rests because the rest is always strictly less than the quotient. The division stops at the first nonzero rest, and the quotient has a finite decimal writing. If there is a nonzero rest, the quotient has a finite decimal writing. If this is not the case, we set $p = Nq + r_0$ and carry out successive divisions of $10r_k$ by q, giving the quotient x_{k+1} and the rest r_{k+1}. The quotient $\frac{p}{q}$ is the written $\lfloor x \rfloor, x_1x_2x_3 \ldots$. The rests are then always between 1 and $q-1$. We cannot perform q steps without encountering two identical rests. If we denote r_k and r_{k+l}, as the first two identical rests, the divisions of $10r_k$ and $10r_{k+l}$ by q will have the same quotient $x_{k+1} = x_{k+1+l}$ and the same rest $r_{k+1} = r_{k+1+l}$, and so on. We therefore see a period of length l.
- If q is a product of powers of 2 and 5, then the quotient is a decimal. If $q = 2^k5^m$, then $\frac{p}{q} = \frac{2^m5^kp}{2^m5^kq} = \frac{2^m5^kp}{10^{m+k}}$ is a decimal and therefore has a finite expansion.
- If q and 10 are coprime, the period begins just after the comma. The different rests r_i are the rests of the Euclidean division of 10^ir_0 by q. If $r_k = r_{k+l}$, then 10^kr_0 and $10^{k+l}r_0$ have the same rest, so $10^kr_0(10^l-1)$ is a multiple of q. As q is prime with 10, q divides $r_0(10^l-1)$, which allow us to say that 10^lr_0 and r_0 have the same rest. $r_0 = r_l$, so $x_1 = x_{l+1}$, and the period starts at x_1.

Reciprocally, we can assume $0 \leqslant x < 1$. If the expansion of x is periodic, we have $x = 0, b_1b_2\cdots b_sa_1a_2\cdots a_na_1a_2\cdots a_na_1a_2\cdots a_n\cdots$. By multiplying by 10^s, we therefore have
$10^sx = b_1b_2\cdots b_s, a_1a_2\cdots a_na_1a_2\cdots a_n\cdots = b_1b_2\cdots b_s+0, a_1a_2\cdots a_na_1a_2\cdots a_n\cdots$. We are thus led to prove that the number $y = 0, a_1a_2\cdots a_na_1a_2\cdots a_n\cdots$ is rational. We then multiply by 10^n, and we have $10^ny = a_1a_2\cdots a_n, a_1a_2a_na_1a_2\cdots a_n\cdots = a_2a_2\cdots a_n + y$; hence
$y = \frac{a_1a_2\cdots a_n}{10^n-1} \in \mathbb{Q}$. □

Remark 1.2

1. A real number x is rational if and only if its expansion in any base is finite or infinite periodic.
2. The expansion of an irrational in any base is infinite nonperiodic.

Examples 1.1

1. The decimal expansion of the rational $13/11$ is $1, 18181818181\ 818181\ldots$
2. The expansion of π in base 10 is $3, 14159265358979323846264 3\ldots$
3. The expansion of π in base 2 is $11, 001001000011111110110\ldots$

Definition 1.2 Let E be a set. The group of permutations of E denoted $S(E)$ is the set of bijective applications of E on itself. Thus $\delta \in S(E)$ if and only if δ is a bijective map of E on E.

Note that if $card(E) = n$, then $card(S(E)) = n!$.

Main Result

Theorem 2.1 *Let x be a positive real number and b an integer greater than or equal to* 2.

There exists a unique sequence of integers $(x_n)_{n\geqslant 1}$ between 0 *and* $b-1$ *such that the expansion of x in base b is $\lfloor x \rfloor_b, x_1x_2x_3x_4x_5\ldots$. Let $A = \{x_n, n \in \mathbb{N}^*\}$. In other words, A is the set of digits appearing after the comma in the development of x in base b.*

If $\delta \in S(A)$, then the real y whose expansion in base b is $\lfloor x \rfloor_b, \delta(x_1)\delta(x_2)\delta(x_3)$ $\delta(x_4)\ldots$ is such that:

- *If x is normal in base b, then y is normal in base b.*
- *If x is simply normal in base b, then y is simply normal in base b.*
- *If x is abnormal in base b, then y is abnormal in base b.*

Examples 2.1 We will give simple examples of applications of the theorem in some bases.

1. The number x whose expression in base 2 is $0.1010101010\ldots$ is simply normal in base 2 because here the frequency of appearance of the digits 0 and 1 is equal to $\frac{1}{2}$. Here $A = \{0, 1\}$. Let $\delta \in S(A)$ such that $\delta(0) = 1$, and $\delta(1) = 0$. According to Theorem 2.1, the real y whose expression in base 2 is
   $0.\ \delta(1)\delta(0)\delta(1)\delta(0)\delta(l)\delta(0)\delta(1)\delta(0)\delta(1)\delta(0)\ldots = 0.0101010101\ldots$
   is simply normal in base 2.
2. The number of Champernowne $x = 0.123456789101112131415161718192021$ $222324252627\ldots$ is normal in base 10. Here $A = \{0, 1, 2, \ldots, 9\}$. Let $\delta \in S(A)$ such that $\delta(0) = 5$, $\delta(1) = 3$, $\delta(2) = 4$, $\delta(3) = 6$, $\delta(4) = 9$, $\delta(5) = 7$, $\delta(6) = 8$, $\delta(7) = 2$, $\delta(8) = 1$ and $\delta(9) = 0$. According to Theorem 2.1, the real number
   $y = 0.\delta(1)\delta(2)\delta(3)\delta(4)\delta(5)\delta(6)\delta(7)\delta(8)\delta(9)\delta(1)\delta(0)\delta(1)\delta(1)\ldots = 0.3469782103533\ldots$
   is normal in base 10.
3. The base 2 expansion of π is $11.0010010000111111\text{0110}\ldots$. Here $A = \{0, 1\}$.
   Let $\delta \in S(A)$ such that $\delta(0) = 1$ and $\delta(1) = 0$. According to Theorem 2.1, if we manage to show that the real y whose writing in base 2 is
   $11.\delta(0)\delta(0)\delta(1)\delta(0)\delta(0)\ldots = 11.\ 11011\ldots$
   is normal in base 2, then π would be normal in base 2.

Proof Let x be a positive real number and $b \geqslant 2$ a given integer.

The expansion of x in base b is $\lfloor x\rfloor_b, x_1x_2x_3x_4\ldots$ with $(x_n)_{n\geqslant 1}$ a sequence between 0 and $b-1$. Let $A=\{x_n, n\in\mathbb{N}^*\}$. We have $A\subset\{0,1,\ldots,b-1\}$. Let $\delta\in S(A)$. Let the number y be written in base b as $\lfloor x\rfloor_b, \delta(x_1)\delta(x_2)\delta(x_3)\delta(x_4)\ldots$.

Let for $j\in\mathbb{N}$, $L=\delta(x_{j+1})\delta(x_{j+2})\ldots\delta(x_{j+k})$ be a sequence of length k in the base b expansion of y. Let $S=x_{j+1}x_{j+2}\ldots x_{j+k}$. Denote by $N'(L,n)$ the number of occurrences of L among the first n digits after the comma in the expansion of y in base b and $N(S,n)$ the number of occurrences of S among the first n digits after the comma in the expansion of x in base b.

Let us show that $N'(L,n)=N(S,n) \quad \forall n\in\mathbb{N}^*$.

Two cases are possible: Either L appears, or it does not appear in the expansion of y in base b.

- If L does not appear in the expansion of y, then neither does S in the expansion of x because δ is bijective. So $N'(L,n)=N(S,n)=0\;\forall\, n\in\mathbb{N}$.
- If L appears in the expansion of y,

we will draw the following schema to establish the idea:

$\lfloor x\rfloor_b,\quad x_1\quad x_2\quad x_3\ldots\ldots x_{j+1}\quad x_{j+2}\ldots\ldots\ldots x_{j+k}\ldots\ldots x_{j+1}\quad x_{j+2}\ldots\ldots x_{j+k}\ldots x_n\ldots$

$\lfloor x\rfloor_b,\quad \delta(x_1)\delta(x_2)\delta(x_3)\ldots\delta(x_{j+1})\delta(x_{j+2})\ldots\delta(x_{j+k})\ldots\delta(x_{j+1})\delta(x_{j+2})\ldots\delta(x_{j+k})\ldots\delta(x_n)\ldots$.

We can clearly see here that each x_i is aligned vertically with $\delta(x_i)$. Thus each time S appears in the expansion of x at a position $t\leqslant n-k$, L appears at the same position t in the expansion of y and vice versa because δ is a bijection. So we see clearly that $N'(L,n)=N(S,n) \quad \forall\, n\in\mathbb{N}^*$.

1. Assume that x is normal in base b. Let us show that in this case y is normal in base b.

   Since x is b-normal, then $\lim\limits_{n\to+\infty}\dfrac{N(S,n)}{n}=\dfrac{1}{b^k}$. It therefore follows that
$$\lim_{n\to+\infty}\frac{N'(L,n)}{n}=\lim_{n\to+\infty}\frac{N(S,n)}{n}=\frac{1}{b^k}.$$
   So $\lim\limits_{n\to+\infty}\dfrac{N'(L,n)}{n}=\dfrac{1}{b^k}$. Therefore y is normal in base b.
2. Assume that x is simply normal in base b. Let us show that in this case y is simply normal in base b.

   $\lim\limits_{n\to+\infty}\dfrac{N'(L,n)}{n}=\lim\limits_{n\to+\infty}\dfrac{N(S,n)}{n}=\dfrac{1}{b}$, because S is a sequence of length k and we know that x is simply normal in base b. So $\lim\limits_{n\to+\infty}\dfrac{N'(L,n)}{n}=\dfrac{1}{b}$. Therefore y is simply normal in base b.
3. Finally, assume that x is not normal in base b.

   So there exists $F=x_{l+1}x_{l+2}\ldots x_{l+k}$ a sequence of length k such that $\lim\limits_{n\to+\infty}\dfrac{N(F,n)}{n}\neq\dfrac{1}{b^k}$.

By setting $G = \delta(x_{l+1})\delta(x_{l+2})\ldots\delta(x_{l+k})$, $\lim_{n\to+\infty}\frac{N'(G,n)}{n}$ $= \lim_{n\to+\infty}\frac{N(F,n)}{n} \neq \frac{1}{b^k}$.

So $\lim_{n\to+\infty}\frac{N'(G,n)}{n} \neq \frac{1}{b^k}$. Therefore y is also an abnormal number in base b, which completes the demonstration. □

References

1. D. H. Bailey and R. E. Crandall *On the random character of fundamental constant expansion*, Experimental Math., **10** (2001), (2), 175–190.
2. D.G. Champernowne, *The construction of decimals normal in the scale of ten*, J. London Math. Soc. 8 (1933), 254–260.
3. A.H. Copeland and P. Erdös, *Note on normal decimals*, Bull. Amer. Math. Soc. 52(1946), 857–860.
4. G. Martin, *Absolutely abnormal numbers*, Amer. Math. Monthly 108 (2001), no.8, 746–754.
5. Y. Nakai and I. Shiokawa, *Normality of Numbers generated by the values of polynomials at primes*, Acta Arith. **81** (1997), no.4, 345–356.
6. W. Sierpinski, *Démonstration élémentaire du théorème de M. Borel sur les nombres absolument normaux et détermination effective d'un tel nombre*, Bull. Soc. Math. France 45 (1917), 127–132.

Part III
Contributed Talks: Computer Science and Telecommunications

Chapter 13
Robustness of Imputation Methods with Backpropagation Algorithm in Nonlinear Multiple Regression

Castro Gbêmêmali Hounmenou, Milognon Boris Behingan, Christophe Archille Chrysostome, Kossi Essona Gneyou, and Romain Glèlè Kakaï

Abstract Missing observations constitute one of the most important issues in data analysis in applied research studies. The magnitude and their structure impact parameters estimation in the modeling with important consequences for decision-making. This chapter aims to evaluate the efficiency of imputation methods combined with the backpropagation algorithm in a nonlinear regression context. The evaluation is conducted through a simulation study including sample sizes (50, 100, 200, 300, and 400) with different missing data rates (10, 20, 30 40, and 50%) and three missingness mechanisms (MCAR, MAR, and MNAR). Four imputation methods (Last Observation Carried Forward, Random Forest, Amelia, and MICE) were used to impute datasets before making prediction with backpropagation algorithm. 3-MLP model was used by varying the activation functions (Logistic-

C. G. Hounmenou (✉)
Institut de Mathématiques et de Sciences Physiques, University of Abomey-Calavi, Dangbo, Benin

Laboratoire de Biomathématiques et d'Estimations Forestières, University of Abomey-Calavi, Dangbo, Benin

M. B. Behingan
Laboratoire de Biomathématiques et d'Estimations Forestières, University of Abomey-Calavi, Dangbo, Benin

Laboratoire de Recherche Avicole et de Zoo-Economie, University of Abomey-Calavi, Dangbo, Benin

C. A. Chrysostome
Laboratoire de Recherche Avicole et de Zoo-Economie, University of Abomey-Calavi, Dangbo, Benin

K. E. Gneyou
Laboratoire de Modélisations Mathématiques et Applications, University of Lome, Lomé, Togo

R. G. Kakaï
Laboratoire de Biomathématiques et d'Estimations Forestières, University of Abomey-Calavi, Dangbo, Benin

C. T. Gueye et al. (eds.), *Mathematics of Computer Science, Cybersecurity and Artificial Intelligence*, Springer Proceedings in Mathematics & Statistics 468,
https://doi.org/10.1007/978-3-031-66222-5_13

Linear, Logistic-Exponential, TanH-Linear, and TanH-Exponentiel), the number of nodes in the hidden layer (3–15), and the learning rate (20–70%). Analysis of the performance criteria (R^2, r, and $RMSE$) of the network revealed good performances when it is trained with TanH-Linear functions, 11 nodes in the hidden layer, and a learning rate of 50%. MICE and Random Forest were the most appropriate for data imputation. These methods can support up to 50% of missing rate with an optimal sample size of 200.

Keywords Multilayer perceptron neural network · Regression model · Backpropagation · Missing data · Imputation method

Introduction

Let Y be a real random variable revealed mean depends on $\mathbf{x} = (x_1, \cdots, x_p) \in \mathbb{R}^p$, replications of the random vector $\mathbf{X}$, and the dependence may be nonlinear, $\mathbb{E}(Y|x_1, \cdots, x_p) = \zeta(x_1, \cdots, x_p)$. This relation is equivalent to $Y = \zeta(x_1, \cdots, x_p) + \epsilon$ with $\mathbb{E}(\epsilon) = 0$. Let a parametric nonlinear regression model be represented by $Y = \zeta(x_1, \cdots, x_p; \theta) + \epsilon$, where ζ is nonlinear with respect to θ, the set of model parameters. This means that, for at least one θ_i, the derivative of ζ with respect to θ_i depends on at least one of the parameters. For example, $\zeta(\mathbf{x}; \theta) = \frac{\theta_1 x_1}{1+\theta_2 x_2}$ is used by chemists. Differentiating ζ with respect to θ_1 and θ_2 gives $\frac{\partial \zeta}{\partial \theta_1} = \frac{x_1}{1+\theta_2 x_2}$ and $\frac{\partial \zeta}{\partial \theta_2} = \frac{-x_1 x_2 \theta_1}{(1+\theta_2 x_2)^2}$. One of the nonlinear models that has received great attention last few years is the model based on artificial neural networks (ANNs). They are used in the fields of prediction and classification, fields in which regression models and other related statistical techniques have traditionally been used [1–4]. Multilayer perceptron neural networks (MLPs) are one of the architectures of ANNs acting as a type of regression model, not necessarily parametric, which enables complex functional forms to be modeled [5, 6]. In breeding, the knowing of production is necessary for specialists who need simple and accurate techniques to predict the production of meat, eggs, milk, etc. Production is influenced by interdependent factors, and MLPs offer more flexibility in describing their relationships. But data collected in the case of production are often small due to the cost of experimentation and seldom complete. Missing data are one of the most common problems for researchers in breeding [7]. It occurs because of human error, equipment failure, death of animal during the experiment, data collected with difficulty, official statistics not available, etc. [8–10, 12–19]. Analysis of incomplete datasets results in problems such as biased parameter estimates, inflation of standard errors, loss of information, and weak generalizability of results [11, 12, 17, 18]. Apart from Kohonen network [20], most of statistical analysis methods assume the absence of missing data and are only able to include observations for which every variables are measured [21]. To overcome this situation, rows with missing values can be deleted (deletion), but it leads to a loss of precision [22, 24] with weak sample size. To avoid this situation, imputation methods can be used. Different imputation methods exist based on different approaches: single imputation, multiple imputation, etc. [25, 26]. With imputation techniques, researchers can obtain complete data for their prediction.

Despite the success of MLPs in breeding and other disciplines, there are exist some factors that can affect its performances such as activation functions, learning rate, number of hidden layers, number of neurons in each hidden layer, etc. [28]. However there are no clear guidelines on which activation function performs better [29] and also about the value of the learning rate [28, 29]. Yet, a drawback of this type of network is that it requires a full set of input data. Therefore our study aims to evaluate the empirical robustness of imputation methods in nonlinear regression with backpropagation (BP) algorithm.

The main objective of this chapter is to analyze the behavior of the imputation methods combined to the BP algorithm for the management of missing data. Specifically, we (i) analyze the effect of imputation methods on the structure of hyperparameters and (ii) determine the best imputation method according to sample size and the missing data rate with the best structure of hyperparameters for the multilayer perceptron neural network.

Framework, Specification of Model, and Generation of a Data Population

Types of Missing Data and Their Management

Let $X = [x_{ij}]$ be a data matrix of dimension (n, p) of elements $x_{ij} \in \mathbb{R}$, where n and p elements of $\mathbb{N}^*$ are, respectively, the number of observations and the number of variables, and x_{ij} is the value of the variable $j \in [\![1, p]\!]$ for the observation $i \in [\![1, n]\!]$. Let $Z = [z_{ij}]$, an indicator matrix of missing data elements z_{ij}, such that $z_{ij} = 1$ if x_{ij} is missing and $z_{ij} = 0$ otherwise, then we have $X = \{X_{obs}, X_{mis}\}$. The matrix Z describes the structure of the missing data and is useful to treat it as a stochastic matrix. The statistical model for missing data are $P(Z|X, \kappa)$, where κ is the parameter of the missing data process and $P(\cdot)$ denotes the conditional distribution of Z given X of parameters κ. The mechanism of missingness is determined by the dependency of Z on the variables in the dataset. According to [8], three categories of missing data can be distinguished: Missing Completely at Random (MCAR), Missing at Random (MAR), and Missing Not at Random (MNAR).

Definition 1 (Missing Data Are "Missing Completely at Random") Missing data are said to be MCAR when the fact of not having a value is totally independent of the variables X and we have

$$\forall X, P(Z|X, \kappa) = P(Z|\kappa). \tag{13.1}$$

When the missing data are not MCAR, we need to know if differences in the characteristics of nonrespondents and respondents can be explained by variables common to respondents and nonrespondents. We note X_{obs}, the observed part of the data X and X_{mis}, the missing part.

Definition 2 (Missing Data Are "Missing at Random") The data are said to be MAR when the distribution of Z given X depends only on the variables recorded in the database X_{obs}, and we have

$$\forall X_{mis}, P(Z|X_{obs}, X_{mis}, \kappa) = P(Z|X_{obs}, \kappa). \tag{13.2}$$

Definition 3 (Missing Data Are "Missing Non At Random") The data are said to be MNAR when the distribution of Z given X also depends on X_{mis}, and we have

$$\forall X_{obs} \; and \; X_{mis}, P(Z|X_{obs}, X_{mis}, \kappa) = P(Z|X_{obs}, X_{mis}, \kappa). \tag{13.3}$$

There are two basic methods for managing data matrices with missing values: (i) the *deletion method* and (ii) the *imputation method* [23]. The first one considers only the individuals for which all the data are available, i.e., to delete any individual having at least one missing value. The second consists in replacing the missing values in the dataset by estimated ones. Two imputation approaches are used: *simple imputation* and *multiple imputation* [30]. Single imputation is to fill in each missing value with a value. The second approach covers methods whose procedures are based on models. This is done by replacing the missing values with several simulated values to properly reflect the uncertainty that is attached to the missing data [31].

Factors Affecting the Predictive Performance of a Multilayer Perceptron Neural Network and Backpropagation Algorithm

A multilayer perceptron neural network (MLP) is a feedforward neural network, consisting of a number of units (called neurons) connected by weight links. The units are organized in several layers, the first one is an input layer, the last one is an output layer, and the intermediate one can have one or several hidden layers. The input layer receives an external activation vector and transmits it via weighted connections to the units of the first hidden layer. These compute their activations and transmit them to the neurons in succeeding layers, see Fig. 13.1.

Although multilayer perceptron neural networks have shown good predictive performance compared to classical methods, they are often affected by factors such as the number of neurons and layers, the choice of transfer functions, and the sample size. For more details, see [27]. The estimation of the network weights is done by minimizing a quadratic cost function. It can be done, among other things, by the BP algorithm, whose procedure is summarized as follows:

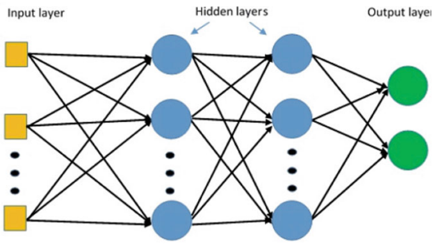


Fig. 13.1 Example of a multilayer perceptron network with two hidden layers

1. Initialize all weights to small random values in the interval [−0.9, 0.9].
2. Normalize the training data.
3. Randomly permute the training data.
4. For each training data k:

   (a) Compute the observed outputs by forward propagating the inputs.
   (b) Adjust the weights by backpropagating the observed error from the output layer toward the input layer:

$$\begin{aligned} w_{ij}(k+1) &= w_{ij}(k) + \Delta w_{ij}(k+1) \\ &= w_{ij}(k) + \eta \delta_j(k+1) y_i(k+1) \end{aligned} \tag{13.4}$$

   with $w_{ij}(k+1)$, the adjusted weight for the neuron j; $w_{ij}(k)$, the previously computed weight for the neuron j; $0 \leq \eta \leq 1$ representing the learning rate; $\delta_j(k+1)$ the local gradient computed for the neuron j, and $y_i(k+1)$ representing either the output of neuron i on the previous layer, if it exists, or the input i otherwise.
5. Repeat steps 3 and 4 up to a maximum number of iterations or until the root mean square error is less than a certain threshold.

Specification of Model and Generation of a Data Population

The nonlinear regression model considered is a multilayer perceptron neural network with a hidden layer, and its expression is

$$Y = \zeta_\theta(\mathbf{x}) + \epsilon \tag{13.5}$$

with $\mathbb{E}(\epsilon) = 0$; $\mathbf{x} \in \mathbb{R}^p$ is a vector of p inputs, and $Y \in \mathbb{R}$, $\zeta_\theta(\mathbf{x}) \in \mathbb{R}$ are, respectively, the observed output and the predicted output.

$$\zeta_\theta(\mathbf{x}) = f_2\left(\mathbf{w}^{(2)}\mathbf{f_1}\left(\mathbf{w}^{(1)}\mathbf{x} + \mathbf{b}^{(1)}\right) + b^{(2)}\right), \tag{13.6}$$

where $\theta = \left(\mathbf{w}^{(1)}, \mathbf{b}^{(1)}; \mathbf{w}^{(2)}, b^{(2)}\right)$ and
$\theta = \left(w_{11}^{(1)}, \cdots, w_{1p}^{(1)}, \cdots, w_{m1}^{(1)} \cdots, w_{mp}^{(1)}; b_{10}^{(1)} \cdots, b_{m0}^{(1)}; w_{11}^{(2)}, \cdots, w_{1m}^{(2)}, b^{(2)}\right)$ are the model's parameters, and the total number of parameters is

$$n_\theta = m(p+2) + 1, \tag{13.7}$$

with m the number of neurons in the hidden layer, f_2 is a transfer function applied to only neuron of the output layer, and $\mathbf{f_1}$ is a vector composed of the same transfer function applied to each neuron in the hidden layer.

In order to have data with multicollinearity, a nonlinear relationship between variables, and for predictive purposes, we used the results of Insect as Feed for West Africa project [51] which evaluated the effect of maggot meal on the growth and economic performance of guinea fowl. The dependent variable is $y = food\ economic\ efficiency$, and independent variables are $\mathbf{x} = (x_1 = dose\ with\ three\ modality\ (0,\ 50,\ and\ 100), x_2 = age, x_3 = food\ consumption, x_4 = weight)$. The predictive model obtained is

$$\mathbb{E}(y_t) = \sum_{i=1}^{11} \mathbf{w}_i^{(2)} \frac{\exp\left(\langle \mathbf{x}_t, \mathbf{w}_i^{(1)} \rangle + b_i^{(1)}\right) - \exp\left[-\left(\langle \mathbf{x}_t, \mathbf{w}_i^{(1)} \rangle + b_i^{(1)}\right)\right]}{\exp\left(\langle \mathbf{x}_t, \mathbf{w}_i^{(1)} \rangle + b_i^{(1)}\right) + \exp\left[-\left(\langle \mathbf{x}_t, \mathbf{w}_i^{(1)} \rangle + b_i^{(1)}\right)\right]} + b^{(2)}, \tag{13.8}$$

where y_t represents the tth observation ($t \in [\![1, n]\!], n \in \mathbb{N}^*$), $\mathbf{w}_i$ is the weight vector associated with the ith neuron in the hidden layer ($i \in [\![1, 11]\!]$), and b_i and b are, respectively, the bias of the ith neuron in the hidden layer and the bias applied to output neuron of 3-MLP model. The optimal parameters are $\theta = (\mathbf{w}^{(1)}, \mathbf{b}^{(1)}, \mathbf{w}^{(2)}, b^{(2)})$ with

$$\mathbf{w}^{(1)} = \begin{bmatrix} -0.06 & -0.87 & 0.33 & -0.10 & -0.15 & 0.08 & -0.13 & 0.60 & -0.07 & 0.04 & -0.17 \\ -0.06 & 0.41 & -0.38 & 0.29 & -0.18 & -0.31 & 0.22 & -0.44 & -0.12 & 0.16 & 0.28 \\ -0.13 & 0.47 & -0.16 & -0.27 & 0.22 & 0.20 & -0.36 & 1.42 & -0.13 & 0.27 & -0.48 \\ 0.22 & 0.35 & 0.21 & 0.03 & 0.15 & 0.01 & 0.27 & 0.55 & 0.32 & -0.43 & 0.37 \end{bmatrix}$$

$$\mathbf{b}^{(1)} = \left[0.31\ 0.65\ 0.45\ -0.82 - 0.39\ -0.38\ -0.86\ -0.01\ -0.79\ 0.79\ -0.43\right];$$

$$\mathbf{w}^{(2)} = \left[0.01\ -0.38\ 0.15\ -0.03\ -0.05\ 0.04\ -0.04\ 0.61\ -0.01\ -0.02\ -0.05\right];$$

$$b^{(2)} = -0.82.$$

The total number of parameters is $n_\theta = 67$.

A population of size $N = 10000$ was obtained from Eq. (13.8) to which we added the error ϵ of the Eq. (13.5) to compute Y. The error was generated according to $\mathcal{N}(\mu = 0, \sigma^2 = 1)$. The input variables X_1–X_4 related to Y were defined using their respective distributions, X_1 by resampling techniques, $X_2 \sim \mathcal{N}(\mu = 4.5, \sigma^2 = 2.30)$, $X_3 \sim \mathcal{N}(\mu = 29.95, \sigma^2 = 13.04)$, and $X_4 \sim \mathcal{N}(\mu = 239.76, \sigma^2 = 117.11)$.

Simulation Study

Seven factors were considered in this study. The study considered various factors, including the sample size (5 different sizes), the missingness mechanism (3 mechanisms), the missing data rate (5 rates), and the imputation methods (4 methods). Additionally, the factors influencing the predictive and explanatory performance of the MLP model were analyzed: the activation function (4 functions), the number of hidden neurons (13 sizes), and the learning rate (6 rates). For each sample size, we have a combination of 936,000 items, which is replicated 100 times.

Sampling Size, Simulating Missingness, and Missing Data Imputation

Five samples of different sizes n_i ($n_i = 50$, 100, 300, and 400) were extracted from the population using the bootstrap technique [32]. Three missingness mechanisms were considered, MAR, MCAR, and MNAR (see section "Framework, Specification of Model, and Generation of a Data Population") with five missing data rates (MRs) (10, 20, 30, 40, and 50%) to generate incomplete datasets. Missingness simulation is conducted on each of the five complete data using *MICE* package [33] from software R 3.3.6 [34]. Each of these previously obtained missing data are imputed with Last Observation Carried Forward (LOCF), Random Forest (RF), Amelia (AMELIA), and Multivariate Imputation by Chained Equation (MICE) methods in R using, respectively, *zoo* [35], *missForest* [36], *Amelia* [21], and *MICE* package [33].

Prediction with 3-MLP in R Software

Before performing the prediction, 75% of each imputed dataset is used to train the neural network and 25% to test trained network concerning its generalization capacity. Before performing the training and testing, the imputed datasets were normalized using min-max normalization technique [37]:

$$new_v = \frac{v - \min_z}{\max_z - \min_z}(new \max_z - new \min_z) + new \min_z, \tag{13.9}$$

where v is an observation of vector z and new_v is a normalized observation.

The function "*mlp*" of *RSNNS* package [38] was used for the prediction. A 3-MLP model (see Eq. (13.5)) was used by varying hyperparameters for each sample size of imputed dataset. Four combinations of activation functions (AFs) (f_1 and f_2, see Eq. 13.5) were used: (i) Logistic-Linear (LL), (ii) Logistic-

Exponential (LE), (iii) TanH-Linear (TL), and (iv) TanH-Exponentiel (TE). The expression of activation functions considered is Linear, $h(x) = x$, Logistic, $h(x) = \frac{1}{1+e^{-x}}$, Exponential, $h(x) = e^x$, and Tangent hyperbolic, $h(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$. In additional, 13 numbers of nodes (Node) in the hidden layer were considered: 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, and 15. In addition, six learning rates (LRs) were considered: 20%, 30%, 40%, 50%, 60%, and 70%, as well as five sample sizes of imputed dataset (Size). The considered learning algorithm is standard backpropagation (see section "Framework, Specification of Model, and Generation of a Data Population").

A total of 100 replications was performed on each size of imputed dataset to analyze the performance of the method. Initial weights were generated randomly according to the uniform law in the range of −3 and 3. The stopping criteria used are the combination of a fixed number of epochs, NE = 1000, and a sufficiently small training error less than or equal to 10^{-6}.

Performance Criteria and Statistical Method Comparison

The performance criteria used are (i) coefficient of correlation, r, (ii) coefficient of determination, R^2, and (iii) Root Mean Squared Error, $RMSE$ [39, 40]. In the formula below, y and ζ_θ, respectively, denote observed outputs and predicted outputs, $\bar{y}$ and $\bar{\zeta}_\theta$ their mean, and n the test data size.

$$r = \frac{\sum(y_t - \bar{y}_t)(\zeta_\theta(x_t) - \bar{\zeta}_\theta(x_t))}{\sqrt{\sum(y_t - \bar{y}_t)^2} \times \sqrt{\sum(\zeta_\theta(x_t) - \bar{\zeta}_\theta(x_t))^2}} \tag{13.10}$$

$$R^2 = \frac{\sum(y_t - \zeta_\theta(x_t)) \times (\sum y_t \times \sum \zeta_\theta(x_t))}{\sqrt{(\sum y_t^2 - (\sum y_t)^2)(\sum \zeta_\theta(x_t)^2 - (\sum \zeta_\theta(x_t))^2)}} \tag{13.11}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(\zeta_\theta(x_t) - y_t\right)^2}. \tag{13.12}$$

The appropriate imputation method for a missing data mechanism giving the best configuration of model characteristics (13.5) with the BP algorithm and for an optimal sample size is the model for which we observe a high correlation between predicted and observed data ($|r| \geq 0.8$) [39], with R^2 close to "1" [41] and with a low value of $RMSE$ [39].

To assess effects of factors (Size, MR, AF, Node, and LR) which affect performance of the 3-MLP model, the generalized linear models based on the beta distribution were run on R^2, $|r|$, and the linear fixed effects models on $RMSE$ for each missing data mechanism and by imputation method.

Interaction plot was considered for significant interactions between the MLP hyperparameters by missing data mechanism.

$Mean$, $minimum$, $maximum$, and $coefficient\ of\ variation$ of the criteria considered (R^2, $|r|$, and $RMSE$) were used to compare imputation method performances.

Results

Effect of Imputation Methods by Missing Data Mechanism on the Performance of the Hyperparameter Structure of the 3-MLP Model

Table 13.1 shows the results of the effect of the imputation methods (Amelia, LOCF, RF, and MICE) by missing data mechanism (MAR, MACR, and MNAR) on the performance of the hyperparameter structure (AF, LR, and Node) of the 3-MLP model. The analysis shows that AF, LR, and Node significantly affect the performances of the imputation methods whatever the missing data mechanism ($p < 0.05$). However, the second-order interaction of these factors (AF:LR:Node) did not impact ($p > 0.05$) the performances of imputation methods for across the three missing data mechanisms. The predictive performances (R^2 and r) of the imputation methods used were not affected by the interaction between learning rate and the number of neurons in the hidden layer (LR:Node) but had a significant impact on the root mean square error (RMSE) for each missing data mechanism. By considering the interaction between the activation function and the number of neurons in the hidden layer (AF:Node), we observed that from a missing data mechanism to another, the predictive performances of AMELIA and LOCF were not affected by this interaction. However, those of RF and MICE were significantly affected. About the RMSE, apart the one of LOCF under MAR assumption, others were significantly affected by this interaction. Results also revealed a significant effect on the performances of imputation methods concerning the interaction between the activation function and the learning rate (AF:LR) for all missing data mechanism.

Effect of Imputation Methods by Missing Data Mechanism on the Performance of Activation Function and Learning Rate

The interaction plots revealed that under MAR assumption, the performances of imputation methods increase with the learning rate when we use activation functions

Table 13.1 Effect of imputation methods by missing data mechanism on the structure of hyperparameters: results of GLM and linear models

	Amelia			LOCF			RF			MICE		
Factors	R^2	RMSE	r	R^2	RMSE	r	R^2	RMSE	r	R^2	RMSE	r
MAR												
AF	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
LR	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Node	0.002	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
AF:LR	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
AF:Node	0.439	0.001	0.434	0.526	0.092	0.500	0.001	0.001	0.001	0.001	0.001	0.001
LR:Node	0.999	0.001	0.999	0.999	0.098	0.999	0.999	0.001	0.999	0.999	0.001	0.999
AF:LR:Node	0.999	0.001	0.999	0.999	0.001	0.999	0.999	0.001	0.999	0.999	0.001	0.999
MCAR												
AF	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
LR	0.001	0.001	0.001	0.997	0.001	0.994	0.001	0.001	0.001	0.001	0.001	0.001
Node	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
AF:LR	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
AF:Node	0.999	0.001	0.999	0.999	0.001	0.999	0.001	0.001	0.001	0.001	0.001	0.001
LR:Node	0.999	0.001	0.086	0.999	0.098	0.999	0.999	0.203	0.999	0.999	0.001	0.999
AF:LR:Node	0.999	0.001	0.999	0.999	0.001	0.999	0.999	0.001	0.999	0.999	0.001	0.999
MNAR												
AF	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
LR	0.001	0.001	0.001	0.997	0.001	0.994	0.001	0.001	0.001	0.001	0.001	0.001
Node	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
AF:LR	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
AF:Node	0.783	0.001	0.776	0.872	0.001	0.870	0.001	0.001	0.001	0.001	0.001	0.001
LR:Node	0.999	0.001	0.999	0.999	0.036	0.999	0.999	0.001	0.999	0.999	0.002	0.999
AF:LR:Node	0.999	0.001	0.999	0.999	0.001	0.999	0.999	0.001	0.999	0.999	0.001	0.999

Cells contain p-value; AF, activation function; LR, learning rate

such as TanH-Linear (TL), Logistic-Linear (LL), and Logistic-Exponential (LE), see Fig. 13.2. Contrary to those activation functions, TanH-Exponential (TE) starts to decrease after 30% of learning rate. The best values of R^2 and r was obtained with the TanH-Linear activation function followed by Logistic-Exponential and Logistic-Linear. About the RMSE, the Logistic-Exponential yields the highest values indicating that the network commits more error with this activation function. TanH-Exponential activation function gave the best RMSE. With this activation function, the error varies slightly the learning rate increases contrary to the other function. The latter increased when the learning rate increases. For TanH-Linear and Logistic-Linear, the RMSE was closed but after 40%. Logistic-Linear yields an RMSE greater than the other one.

Similar trends have been observed when the missingness mechanism is either MCAR or MNAR. Thus, the highest values of R^2 and r have been observed with TanH-Linear, while the lowest values have been observed with TanH-Exponential.

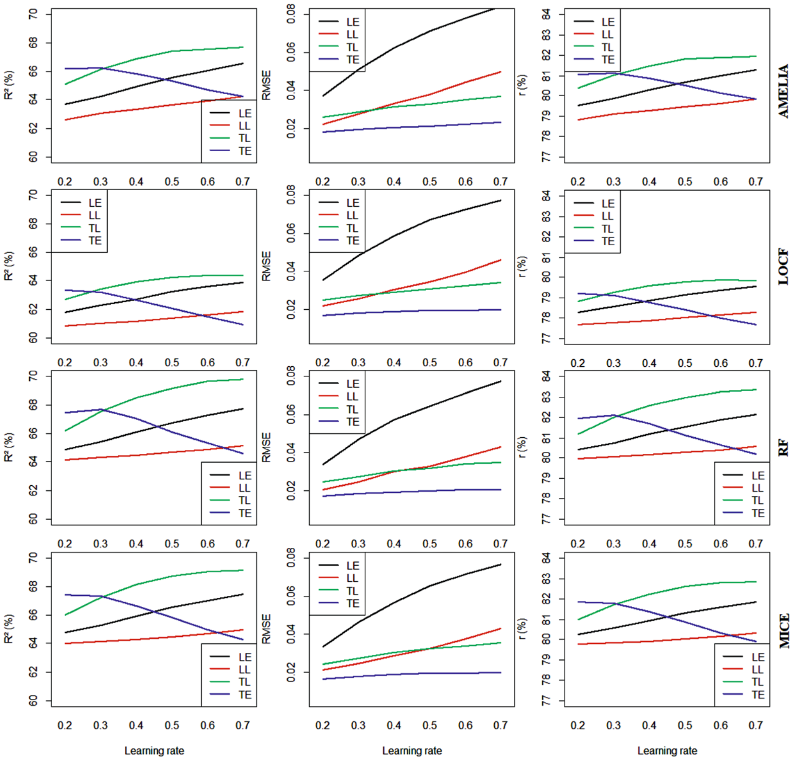


Fig. 13.2 Interaction plot of AF:LR for R^2, RMSE, and r under MAR assumption

As observed under MAR assumption, the TanH-Exponential function commits little error when the data are MCAR or MNAR. For the three missingness mechanisms, the predictive performances of the imputation methods vary slightly after 50% of learning rate indicating that the neural network can be trained with 50% of learning rate for each activation function. More a little variation of the error has been observed from 50% of learning rate for Tanh-Exponential and TanH-Linear contrary to Logistic-Exponential and Logistic-Linear which continues to increase.

Effect of Imputation Methods by Missing Data Mechanism on the Performance of Activation Function and Node

The performances of imputation methods according to the activation function and the number of node in the hidden layer for the three missingness mechanisms revealed almost the same performance, see only Fig. 13.3. The predictive performances of imputation methods improve with the increase of the number of nodes for all missing data mechanism. When data are MAR, R^2, and r values for the

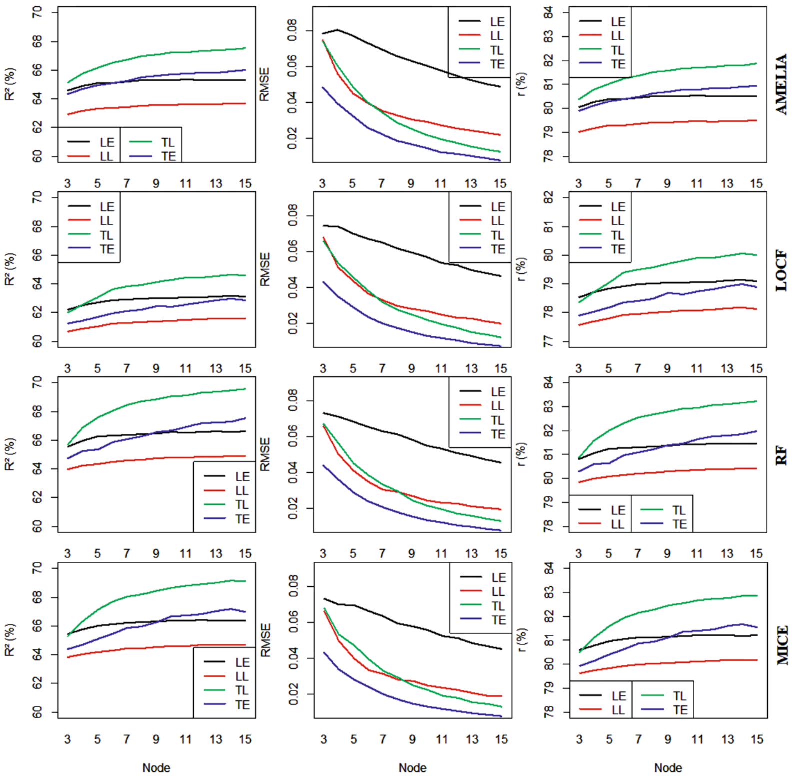


Fig. 13.3 Interaction plot of AF:Node for R^2, RMSE, and r under MAR assumption

TanH-Linear activation function were greater than the values recorded with the other activation functions. The Logistic-Linear yields the lowest values of R^2 and r. The same trend has been observed under MNAR assumption. However, under MCAR assumption, it is TanH-Exponential activation function which had the lowest values of R^2 and r for LOCF method. Concerning the errors commit by the model, it became more and more lower when the number of hidden neurons increased and this for all the imputation methods for the three missing data mechanisms. The model commits more errors with the Logistic-Exponential activation function when the TanH-Exponential functions commit fewer errors. The errors when the activation functions are Logistic-Linear and TanH-Linear were lower than the one with Logistic-Exponential. For these two activation functions, the RMSE was similar from 3 to 7 neurons. After seven nodes, the model with TanH-Linear was better than the one with Logistic-Exponential. The trend of RMSE observed under MAR assumption was similar to the one observed under MCAR and MNAR assumptions.

We also noticed that for the three missing data mechanisms a low variation of the performances (R^2, r, and RMSE) was observed whatever the imputation method used after 11 nodes in the hidden layer.

Effect of Imputation Methods on Size and the Missing Data Rate

Table 13.2 shows how sample size, missing rate, and their interaction affect the performances of imputation method according to the missing data mechanism. The interaction between size and missing rate highly significantly affected the performances of imputation methods whatever the missing data mechanism ($p < 0.01$). The interaction plot under MAR, MCAR, and MNAR is showed similar trends and only for MAR presented in Fig. 13.4.

When the data are MAR, an improvement of R^2 and r had been noticed for LOCF method from 50 to 200 sample size whatever the missing data rate considered. But after 200, the predictive performances start to decrease. The RMSE for this method under the same missingness assumptions followed the same trend. The values were close between the missing data rates and varied slightly for sample sizes between 50 and 200 but increased for sample sizes above 200. For RF and MICE the predictive performances increased when the sample size is between 50 and 200 whatever the missing data rate. However after 200, predictive performances vary slightly. Under 200, RMSE did not vary greatly but increased from 200. With Amelia, a large difference has been noticed among missing rate under 200 sample size for R^2 and r. However, for low missing rate (10 and 20), the performances were better. After 200 sample size, no major difference has been noticed. The trend of RMSE shows that the error is less with low missing data rate and increases after 100 sample size.

Under MCAR assumption, the error was closed for all missing data rates for LOCF, Random Forest, and MICE. It varies slightly under 200 sample size. The predictive performance of LOCF was best with 10% and 40% of missing rate at 200 and 300 sample sizes, respectively. However, the RMSE was greater with 40% of

Table 13.2 Effect of imputation methods on size and missing data rate: results of GLM and linear models

	Amelia			LOCF			RF			MICE		
Factors	R^2	RMSE	r	R^2	RMSE	r	R^2	RMSE	r	R^2	RMSE	r
MAR												
Size	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
MR	0.001	0.001	0.001	0.001	0.101	0.001	0.001	0.045	0.001	0.001	0.074	0.001
Size:MR	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
MCAR												
Size	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
MR	0.001	0.001	0.001	0.001	0.108	0.001	0.001	0.016	0.001	0.001	0.165	0.001
Size:MR	0.001	0.001	0.001	0.001	0.129	0.001	0.001	0.060	0.001	0.001	0.001	0.001
MNAR												
Size	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
MR	0.001	0.001	0.001	0.001	0.038	0.001	0.001	0.182	0.001	0.001	0.005	0.001
Size:MR	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001

Cells contain p-value

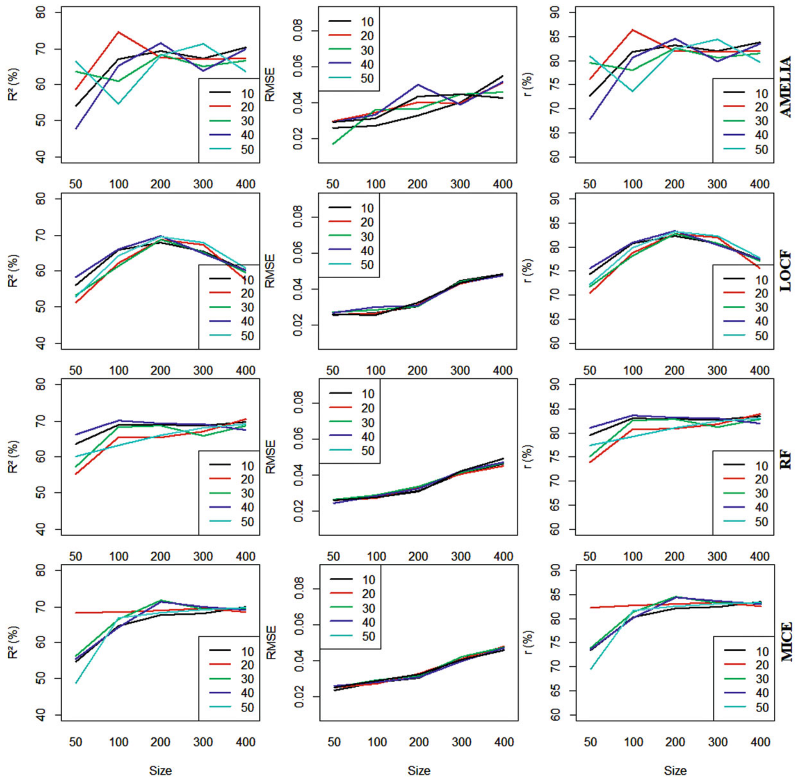


Fig. 13.4 Interaction plot of Size:MR for R^2, RMSE, and r under MAR assumption

missing rate. About Random Forest and MICE, R^2 and r values were better with 10% and 20% of missing rate, respectively, at 200 sample size.

When data are MNAR, the performances obtained for LOCF were similar to what obtained under MAR assumptions. R^2 and r increased between 50 and 200 sample sizes whatever the missing data rate but decrease after 200. Error was closed between missing rates and was best under 200 sample size. The performance of Random Forest method is best with 40% of missing rate at 200 sample size. However, with 20, 30, and 50% of missing rate, values of R^2 and r were closed to the one obtained with 40%. The error did not vary among missing data rate. With MICE method, the error did not vary among missing data rate as observed with Random Forest. Values of R^2 and r were better with 10, 30, and 50% of missing data rate than the values with 20 and 40%. However, the differences were not important. Concerning Amelia method, large variation had been observed among missing data rate for all sample sizes. The error varied slightly for sample sizes below 100 and increased for sample sizes above 100. R^2 and r performed better for 10% and 20% missing data rates at a sample size of 100.

Table 13.3 Mean and coefficient of variation of performances criterion according to imputation method and missing data mechanism

		MAR	MCAR	MNAR
AMELIA	Rsquare	65.22(15.88)	66.68(14.37)	64.49(16.79)
	RMSE	0.038(88,53)	0.037(87.74)	0.036(88.32)
	r	80.45(8.74)	81.43(7.55)	79.96(9.33)
LOCF	Rsquare	62.59(17.58)	59.28(17.82)	63(16.76)
	RMSE	0.035(89.15)	0.034(91.40)	0.035(89.41)
	r	78.76(9.57)	76.64(9.66)	79.04(9.14)
RF	Rsquare	66.46(12.62)	65.77(15.56)	67.35(13.79)
	RMSE	0.035(88.26)	0.035(87.71)	0.035(88.94)
	r	81.34(6.58)	80.78(8.90)	81.84(7.39)
MICE	Rsquare	66.19(15.46)	66.34(14.85)	66.82(13.03)
	RMSE	0.035(89.17)	0.035(88.59)	0.034(88.86)
	r	81.07(8.35)	81.19(7.98)	81.55(6.81)

Comparison of Imputation Methods

The mean and coefficient of variation the performance criteria according to imputation method and missing data mechanism are presented in Table 13.3. There is not a great variation among imputation method under the assumption that data are missing at random. However, LOCF method has the lowest value of R^2 (62.59%). For the other imputation methods used in this study, the values obtained were closed (65.22%, 66.46%, and 66.19%, respectively, for AMELIA, RF, and MICE). The error committed by the model (RMSE) was similar for all imputation methods. The coefficient of correlation was also low with LOCF (78.76%) compared to the others methods (80.45%, 81.44%, and 81.07%, respectively, for AMELIA, RF, and MICE).

When data are MCAR, a similar trend is observed like under MAR assumption. The lowest R^2 and r were obtained with LOCF method (59.28%) when those of AMELIA, RF, and MICE were, respectively, 66.68%, 65.77%, and 66.34% for R^2. Regarding the coefficient of correlation, it is also lower (76.64%) when using LOCF for imputation compared to AMELIA (81.43%), RF (80.78%), and MICE (81.19%). The RMSE was similar between methods (0.035 for RF and MICE; 0.037 and 0.034, respectively, for AMELIA and LOCF).

Under the assumption that data are MNAR, the trend for R^2 and r was different contrary to the values observed when data are MAR and MCAR. R^2 was lower with AMELIA (64.49%) and LOCF (63%) than those of RF (67.35%) and MICE (66.82%). As observed with the other missing data mechanism, RMSE was similar under MNAR assumption. The error was 0.036 for AMELIA, 0.035 for LOCF and RF, and 0.034 for MICE. About the coefficient of correlation, it does not vary greatly from a method to another. Thus we recorded 79.96%, 79.03%, 81.84%, and 81.55% for AMELIA, LOCF, RF, and MICE, respectively.

Discussion

Effect of Imputation Method by Missing Data Mechanism on the Structure of Hyperparameters of 3-MLP Models

For each imputation method, the interaction between AF:LR and AF:Node significantly impacts the performances of the network for any missing data mechanism. The performance of the 3-MLP models is best when the network is trained with the TanH-Linear activation function, 11 nodes in the hidden layer, and a learning rate of 50%. The accuracy of TanH-Linear activation function is much better than the other functions. For the number of nodes, even if the error continuous to decrease after 11 nodes, the gain of the model in terms of prediction has not increased considerably. More, [40, 41, 49] suggest to set the number of nodes in the hidden layer to a minimum as possible because a network with a large number of nodes increases the computational time needed for training. Our findings about the number of nodes in the hidden layer are in agreement with those of [50] which states that the best approach to set the number of nodes in the hidden layer is to start with a small number of nodes and increase until no major improvement in the performances is obtained. As regards the learning rate, after 50%, the predictive ability of the network is still increasing. But this increase in the predictive ability is not important, and a larger learning rate causes network to be more unstable as the error increases. Our results for the optimum learning rate are in agreement with those of the author in [50] who said that if the value of the learning rate is large, the network may show oscillatory response because of the larger changes in the synaptic weight which may cause network to be unstable. However our optimum value of learning rate (50%) is less than 60% suggested by Rajasekaran and Pai [52]. Another study conducted by Nagori [29] set the optimum learning rate as 35% which is less than the one of [52] and the one of our study. This difference can be due to the domain of application which is different.

On the other hand, it should be underlined that the BP algorithm used to train models is intrinsically sensitive to the quality of the data with which it is fed. When data are incomplete due to missing values, this can disrupt the learning process and influence model performance. What is more, every imputation method, whether based on simple techniques such as Last Observation Carried Forward (LOCF) or more complex ones such as Random Forest methods, introduces a certain level of bias into the imputed data. This bias can have a significant impact on the hyperparameter structure of 3-MLP models, as it influences the distribution of the data used for training. Similarly, different missing data mechanisms, such as completely random missing data (MCAR), random missing data (MAR), and nonrandom missing data (MNAR), can have different effects on the way data are imputed and, consequently, on the structure of model hyperparameters. Finally, the interaction between model hyperparameters, such as the activation function, the number of nodes in the hidden layer, and the learning rate, can also play a crucial role in overall model performance. By understanding and taking account of these different factors, it is possible to improve the robustness and generalizability of machine learning models in contexts where missing data are commonplace.

Effect of Imputation Methods on Size and Missing Data Rate

Both size and missing data rate affect imputation methods. No matter the mechanism and the method used, the error increased when the sample size increased for all missing rates. Apart from Amelia, the optimal size is 200 for the others methods under the three missingness mechanisms. For Amelia the optimal sample size is 200 under MAR and MCAR assumption but 100 when the missingness mechanism is MNAR. LOCF can support missing data rate up to 50% with an optimal sample size of 200 under MAR and MNAR assumptions. This method performs better with 10% under MCAR assumption. Since differences between missing data rate are not important, Random Forest and MICE can support up to 50% of missing rate at an optimal sample size of 200. Results are not in agreement with those reported by the authors in [42] who found that error decreased when the sample size increased no matter the missing rate. The difference might be explained by the fact that in our study imputed data pass through the network before evaluating the performances.

The complexity of imputation methods plays a crucial role. Some methods, such as Amelia, rely on more complex multiple imputation models, which can make them more sensitive to different sample sizes and missing data mechanisms than simpler methods such as last observation carried forward (LOCF). Secondly, sensitivity to missing data mechanisms can be considered. Each method may react differently depending on whether the missing data follows a completely random, random, or nonrandom pattern. Finally, the interaction between the imputation process and the learning process, as observed in our study where imputed data are used in the neural network, may modify the optimal performance and sample size required. These different factors could underline the importance of considering the diversity of imputation methods and missing data mechanisms in analytical decision-making.

Comparison of Imputation Methods

Four imputation methods have been used in this study, and results show that there is not a great variation among imputation method under the assumption that data are MAR and MNAR. However, Random Forest method and MICE seem to perform well than AMELIA and LOCF since they have less error. Our findings are in agreement with the results of [43–45]. These authors compare nine imputation methods by considering the three missingness mechanisms (MAR, MCAR, and MNAR). They found that MICE multiple imputations are overall the best approach. Another research of [36] compared the random forest method to kNN imputation [46], MissPALasso (a method based on EM algorithm, proposed by Städler and

Bühlmann [47] and MICE [33]). For these authors, random forest could outperform other imputation methods. Results of these authors are similar to our findings. Indeed in this study random forest and MICE yield similar performances.

Under MCAR assumptions, LOCF is not indicated to handle missing data since it gives low R^2. This confirms the conclusion of [48] which states that single imputation and LOCF are not optimal approaches for missing value imputation, as they can cause bias and lead to invalid conclusions. More, Ref. [13] states that single imputation is not solidly grounded in mathematical foundations, and they exist merely for their ease of implementation. Most of the imputation methods assume that data are missing at random. Our results show that even if this assumption is violated, they perform well since the performances recorded for AMELIA, RF, and MICE do not vary greatly from a missing data mechanism to another. These findings are in agreement with [14] which states that MICE is especially suitable in MAR settings. But the authors in [15] and [16] point out that MICE is also capable to deal with MNAR schemes.

It is important to stress out that, when performing the presented imputation methods, the default settings were used, and tweaking parameters may improve the performance of these methods. This was also mentioned by Stavseth et al. [54] in his study on a comparison of six different imputation methods for categorical questionnaire data. Since missing data can impact significantly the quality of the analyses, it may impact decision-making processes. For [54], regardless of the quality of the statistical method and the robustness of the results, no imputation method can really compensate for the fact that data are effectively missing. Some considerations such as the nature of missing data, the percentage of missing data, the relationships among variables, the types of the data, and the domain of application should guide the choice of method [53, 55].

Conclusion

The possibility to combine imputation methods to multilayer neural network has been accessed in this study through four methods (Amelia, LOCF, Random Forest, and MICE) for any missing data mechanism by controlling the hyperparameters (activation function, number of hidden neurons, and learning rate). From our findings, single imputation is not an optimal approach to deal with missing data. However MICE multiple imputation and RF are more appropriate. Even if these methods outperform the two others (Amelia and LOCF), the best solution is to employ maximal efforts to avoid missing data during data collection. With regard to hyperparameters, to learn the model with the BP algorithm, the performance criteria showed that the combination of TanH-Linear activation functions is best suited to implement the network with 11 nodes in the hidden layer with a learning rate of 50%. However, for further studies, most adapted and developed methods have to be compared with the best method found in this study using other learning methods.

Disclosure of Interests No potential conflict of interest was reported by the author(s).

References

1. Ul-Saufie, A.Z. and Yahya, A.S. and Ramli, N.A. and Hamid, H. A.: Comparison between multiple linear regression and feed forward back propagation neural network models for predicting PM10 concentration level based on gaseous and meteorological parameters. International Journal of Applied **1**(4), 42–49 (2011)
2. Karthikeyan, K. and Kumar, N. and Yousuf, A. and Bhople, B.S. and Naitam, R.K.: Land Evaluation: A General Perspective. Watershed Hydrology, Management and Modeling **175** (2019)
3. Khaled, A.Y. and Abdel Aziz, S. and Bejo, S.K. and Nawi, N.M. and Jamaludin, D. and Ibrahim, N.U.: A comparative study on dimensionality reduction of dielectric spectral data for the classification of basal stem rot (BSR) disease in oil palm. Computers and Electronics in Agriculture **170**, 105288 (2020)
4. Zhang, C. and Pan, X. and Li, H. and Gardiner, A. and Sargent, I. and Hare, J. and Atkinson, P.M.: A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification. ISPRS Journal of Photogrammetry and Remote Sensing **140**, 133–144 (2018)
5. Caselli, M. and Trizio, L. and De Gennaro, G. and Ielpo, P.: A simple feedforward neural network for the PM 10 forecasting: Comparison with a radial basis function network and a multivariate linear regression model. Water, Air, and Soil Pollution **201**(1), 365–377 (2009)
6. Ghazanfari-Hashemi, S. and Etemad-Shahidi, A. and Kazeminezhad, M.H. and Mansoori, A.R.: Prediction of pile group scour in waves using support vector machines and ANN. Journal of Hydroinformatics **13**(4), 609–620 (2011)
7. Torres Munguía, J.A.: Comparison of imputation methods for handling missing categorical data with univariate pattern. Revista de Métodos Cuantitativos para la Economía y la Empresa **17**, 101–120 (2014)
8. Little, R.J., Rubin, D.B.: Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. Annual review of public health **21**(1), 121–145 (2000)
9. Sijtsma, K. and Van der Ark, L.A.: Investigation and treatment of missing item scores in test and questionnaire data. Multivariate Behavioral Research **38**(4), 505–528 (2003)
10. Graham, J.W. and Taylor, B.J. and Olchowski, A.E. and Cumsille, P.E.: Planned missing data designs in psychological research. Psychological methods **11**(4), 323 (2006)
11. Schafer, J.L.: Analysis of incomplete multivariate data. CRC Press (1997)
12. Finch, W.K.: Imputation methods for missing categorical questionnaire data: A comparison of approaches. Journal of Data Science **8**(3), 361–378 (2010)
13. Schafer, J.L. and Graham, J.W.: Missing data: our view of the state of the art. Psychological methods **7**(2), 147 (2002)
14. Janssen, K.J.M. and Donders, A.R.T. and Harrell Jr, F.E. and Vergouwe, Y. and Chen, Q. and Grobbee, D.E. and Moons, K.G.M.: Missing covariate data in medical research: to impute is better than to ignore. Journal of clinical epidemiology **63**(7), 721–727 (2010)
15. He, Y. and Zaslavsky, A.M. and Landrum, M.B. and Harrington, D.P. and Catalano, P.: Multiple imputation in a large-scale complex survey: a practical guide. Statistical methods in medical research **19**(6), 653–670 (2010)
16. White, I.R. and Royston, P. and Wood, A.M.: Multiple imputation using chained equations: issues and guidance for practice. Statistics in medicine **30**(4), 377–399 (2011)
17. De Leeuw, E.D. and Hox, J.J. and Huisman, M.: Prevention and treatment of item nonresponse. Journal of Official Statistics **19**, 153–176 (2003)

18. Dong, Y. and Peng, C.Y.J.: Principled missing data methods for researchers. SpringerPlus **2**(1), 1–17 (2013)
19. Baraldi, A.N. and Enders, C.K.: An introduction to modern missing data analyses. Journal of school psychology **48**(1), 5–37 (2010)
20. Cottrell, M. and Ibbou, S. and Letrémy, P.: Traitement des donnees manquantes au moyen de l'algorithme de kohonen. arXiv preprint arXiv:0704.1709 (2007)
21. Honaker, J. and King, G. and Blackwell, M.: Amelia II: A program for missing data. Journal of statistical software **45**(7), 1–47 (2011)
22. Hox, J.: A review of current software for handling missing data. Kwantitatieve methoden **20**, 123–138 (1999)
23. Gregori, P. and Felip-Bardoll, J.V. and Gras, R.: Imputation multiple de donnees manquantes par l'Analyse Statistique Implicative. Quaderni di Ricerca in Didattica (Mathematics), 20–1 (2010)
24. Preda, C. and Duhamel, A. and Picavet, M. and Kechadi, T.: Tools for statistical analysis with missing data: application to a large medical database. Studies in health technology and informatics **116**, 181–186 (2005)
25. Alison P.D.: Multiple imputation for missing data: A cautionary tale. Sociological methods & research **28**(3), 301–309 (2000)
26. Donzé, L.: L'imputation des données manquantes, la technique de l'imputation multiple, les conséquences sur l'analyse des données: l'enquête 1999 KOF/ETHZ sur l'innovation. Ecole polytechnique fédérale de Zurich, Centre de recherches conjoncturelles, (2001)
27. Hounmenou, C.G. and Tohoun, R. and Gneyou, K.E. and Glèlè Kakaï, R.: Empirical determination of optimal configuration for characteristics of a multilayer perceptron neural network in nonlinear regression. Afrika Statistika **15**(3), 2413–2429 (2020)
28. Chatuverdi, D.K.: Factors affecting the performance of artificial neural network models. Soft Computing: Techniques and its Applications in Electrical Engineering, 51–85 (2008)
29. Nagori V.: Fine tuning the parameters of back propagation algorithm for optimum learning performance. In: 2nd International Conference on Contemporary Computing and Informatics (IC3I), pp. 7–12. IEEE, (2016)
30. Huisman, M.: Imputation of missing item responses: Some simple techniques. Quality and Quantity **34**(4), 331–351 (2000)
31. Zainuri, N.A. and Jemain, A.A. and Muda, N.: A comparison of various imputation methods for missing values in air quality data. Sains Malaysiana **44**(3), 449–456 (2015)
32. Tibshirani, R.? and Efron, B.: An introduction to the bootstrap. Monographs on statistics and applied probability **51**, 1–436 (1993)
33. Groothuis-Oudshoorn, K. and Van Buuren, S.: Mice: multivariate imputation by chained equations in R. Journal of Statistical Software **45**(3), 1–67 (2011)
34. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Version 3.3.6 (2019)
35. Zeileis, A. and Grothendieck, G. and Ryan, J. and Andrews, F.: S3 Infrastructure for Regular and Irregular Time Series (Z's Ordered Observations) R Package Version. R Foundation for Statistical Computing, (2015)
36. Stekhoven, D.J. and Bühlmann, P.: missForest: Nonparametric Missing Value Imputation using Random Forest R package version 1.3. R package, (2012)
37. Priddy, K. L and Keller, P.E.: Artificial neural networks: an introduction. SPIE Press, (2005)
38. Bergmeir, C.N. and Benítez Sánchez, J.M.: Neural networks in R using the Stuttgart neural network simulator: RSNNS. American Statistical Association, (2012)
39. Kazem, H.A. and Jabar, H.Y.: Comparison of prediction methods of photovoltaic power system production using a measured dataset. Energy Conversion and Management **148**, 1070–1081 (2017)
40. Elarabi, H. and Taha, N.F.: Effect of different factors of neural network on soil profile of Khartoum state. American Journal of Earth Sciences **3**(1), 62–66 (2014)
41. Shahin, M.A. and Jaksa, M.B. and Maier, H.R.: State of the art of artificial neural networks in geotechnical engineering. Electronic Journal of Geotechnical Engineering **8**(1), 1–26 (2008)

42. Kalkan, O.K. and Yusuf, K.A.R.A. and Kelecoioğlu, H.: Evaluating performance of missing data imputation methods in IRT analyses. International Journal of Assessment Tools in Education **5**(3), 403–416 (2018)
43. Glasson-Cicognani, M. and Berchtold, A.: Imputation des données manquantes: Comparaison de différentes approches. In: 42èmes Journées de Statistique (2010)
44. Durant G.B.: Imputation methods for handling item-nonresponse in the social sciences: a methodological review. ESRC National Centre for Research Methods and Southampton Statistical Sciences Research Institute, (2005)
45. Stekhoven, D.J. and Bühlmann, P.: MissForest—non-parametric missing value imputation for mixed-type data. Bioinformatics **28**(1), 112–118 (2012)
46. Troyanskaya, O. and Cantor, M. and Sherlock, G. and Brown, P. and Hastie, T. and Tibshirani, R. and Botstein, D. and Altman, R.B.: Missing value estimation methods for DNA microarrays. Bioinformatics **17**(6), 520–525 (2001)
47. Städler, N. and Bühlmann, P.: Mice: multivariate imputation by chained equations in R. Journal of Statistical Software **45**(3), 1–67 (2011)
48. Kang, H.: The prevention and handling of the missing data. Korean journal of anesthesiology **64**(5), 402 (2013)
49. Macukow, B.: Neural networks state of art, brief history, basic models and architecture. In: International conference on computer information systems and industrial management. Springer, Cham, (2016)
50. Uma Rao, K.: Artificial Intelligence and Neural Networks. Pearson Education, (2011)
51. IFWA.: Assess the effect of maggot meal flour on the growth and economic performances of guinea fowl. Laboratoire de Recherche Avicole et de Zoo-Economie, (2017)
52. Rajasekaran, S. and Pai, G.A.V.: Neural networks, fuzzy logic and genetic algorithm: synthesis and applications (with cd). PHI Learning Pvt. Ltd., (2003)
53. Jadhav, A. and Pramod, D. and Ramanathan, K.: Comparison of Performance of Data Imputation Methods for Numeric Dataset. Applied Artificial Intelligence **33**(10), 913–933 (2019)
54. Stavseth, M.R. and Clausen, T. and Roislien, J.: How handling missing data may impact conclusions: A comparison of six different imputation methods for categorical questionnaire data. SAGE Open Medicine **7**, 1–12 (2019)
55. Bennett, D.A.: How can I deal with missing data in my study. Australian and New Zealand Journal of Public Health **25**(5), 464–469 (2001)

Chapter 14
A Better Random Forest Classifier: Labels Guided Mondrian Forest

Ismaël Koné, Adama Samaké, Behou Gérard N'Guessan, and Lahsen Boulmane

Abstract A novel class of Random Forests (RFs), namely Mondrian Forests (MFs), which are an ensemble of Mondrian Trees, achieves competitive performance relatively to classical Breiman RFs. They have attractive properties like performing Bayesian inference at the tree level and being trainable online. However, they perform poorly in the presence of less or low predictive power features. Thus, we propose to extend MF by using label information during splits in order to make them more accurate and robust. We showed an increase in performance when using labels during splits on four datasets where we notice a big improvement on a dataset containing many non-predictive features which is very important as feature relevancy is unknown at first. Additionally, this extension yields equal or superior performance relatively to classical RFs.

Keywords Mondrian forests · Random forests · Entropy · Information gain · Bayesian inference

Introduction

Mondrian forests [1] (MFs) are a recently introduced class of classification random forest (RF) algorithms trainable online that perform Bayesian inference at the tree level. Consequently, they provide better calibrated probabilities compared to classical random forests [2]. However, no label information is involved during splitting operations when building each Mondrian tree. Thus, when data contains many low predictive power features, Mondrian forests perform poorly [1], while this

I. Koné (✉) · B. G. N'Guessan
Université Virtuelle de Côte d'Ivoire, Abidjan, Côte d'Ivoire
e-mail: ismael21.kone@uvci.edu.ci; behou.nguessan@uvci.edu.ci

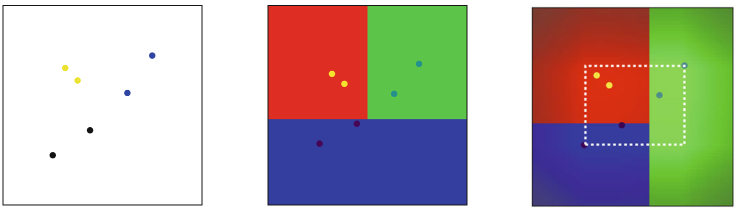
A. Samaké · L. Boulmane
Faculté des sciences, Université Moulay Ismail, Meknes, Morocco
e-mail: a.samake@edu.umi.ac.ma; l.boulmane@umi.ac.ma

C. T. Gueye et al. (eds.), *Mathematics of Computer Science, Cybersecurity and Artificial Intelligence*, Springer Proceedings in Mathematics & Statistics 468,
https://doi.org/10.1007/978-3-031-66222-5_14

is almost always the case in real-life scenarios, where low predictive power features are unknown at first. To deal with these shortcomings, we propose to introduce a random node optimization operation using labels to guide splits as in random forests (section "Method"). Then we will study the effect of such change on the performance of the algorithm (section "Experiment 1") comparatively to the original Mondrian forests. We will also study the impact on the online training property (section "Experiment 2"). Finally, we will compare it to classical random forests (section "Experiment 3").

Background

RFs are a class of machine learning algorithms that have successfully been deployed in many systems. This is due in part to their underlying simplicity, robustness, accuracy, and scalability [3]. Also they have this nice property of feature importance ranking which allows assessing predictive power of features. Nevertheless, they lack well-calibrated probabilities [4]. For example, we train a random forest on a training data. If we apply the trained random forest on a new data different from those of the training set, it will give a result with high confidence. Thus, they are not suitable for applications requiring good probability assessment such as in medicine. Conversely, MFs provide well-calibrated probability assessment at the tree level [5]. This is the case as the feature space is split only inside the bounding box or extent of feature values of the training data. The bounding box is obtained by taking the minimum and maximum values on each feature of the training data. Therefore data outside that bounding box will receive a lower probability proportional to how far it is from the training data. To illustrate, let us visualize an example of decision tree (DT) versus Mondrian tree (MT) with a toy dataset (Fig. 14.1). We notice that DT affects the same high probability (bright color) inside each class of decision boundaries, while MT gives lower probability (dark color) to data outside the extent or bounding box of the toy dataset.



(a) A dataset of 3 classes (one per color)

(b) Decision boundaries from a decision tree

(c) Decision boundaries from a mondrian tree

Fig. 14.1 Difference between MT (**c**) and DT (**b**) on a toy dataset (**a**). Brighter color means higher probability of the associated class, and darker color means lower probability. In (**c**) the dashed rectangle is the bounding box of the data

Another interesting property of MT is that it can be trained in an online fashion. Moreover, the distribution of the online version matches that of the batch version on the whole dataset which is a very important property notably when models need to be updated with new incoming data. Thus, we do not need to retrain models with the whole updated datasets which save computational resources and time.

Despite these properties, splits occurring during MT construction do not use label information. This results in a poor performance when a dataset contains many low predictive power features. Consequently, we cannot use them to rank feature importance which is a very important property in building models and making decisions. Consequently, the main question is can we get Mondrian forests to use label information to fully benefit from random forests' properties additionally to its well-calibrated probabilities?

Method

Our approach consists of applying a random node optimization as in random forests to split the data $\mathcal{D}(\subset \mathbb{R}^D, D \in \mathbb{N})$ based on label information. We choose the information gain (*IG*) derived from the Shannon entropy (*H*) [6]:

$$IG(\delta, \xi, \mathcal{D}) = H(\mathcal{D}) - \frac{1}{|\mathcal{D}|} \sum_{i \in \{R,L\}} |\mathcal{D}_i| H(\mathcal{D}_i) \tag{14.1}$$

$$H(\mathcal{D}) = -\sum_{c \in \mathcal{C}} p(c|\mathcal{D}) log_2(p(c|\mathcal{D})) \tag{14.2}$$

with

$$\mathcal{D}_R = \{x \in \mathcal{D} | x_\delta \geq \xi\}, D_L = \{x \in \mathcal{D} | x_\delta < \xi\}$$

$\delta \in \{1, ..., D\}$ is a feature or dimension of the data $\mathcal{D}$, $\xi \in \mathbb{R}$ is a location or threshold on δ, and $|.|$ means the cardinality of a set. $\mathcal{C}$ is the set of classes or labels associated with each data in $x \in \mathcal{D}$. $\mathcal{D}_R$ and $\mathcal{D}_L$ are subsets of $\mathcal{D}$ after the split based on feature δ at threshold ξ. The probability $p(c|\mathcal{D})$ is the proportion of data with label c in $\mathcal{D}$.

The optimization aims at selecting the couple (δ^*, ξ^*) out of $T \times Q$ randomly sampled split candidates $(\delta_t, \xi_q) \in \{1, ..., D\} \times \mathbb{R}$ that maximize the information gain (*IG*):

$$\delta^*, \xi^* = \underset{\{(\delta_t, \xi_q), (t,q) \in \{1,...,T\} \times \{1,...,Q\}\}}{\text{argmax}} IG(\mathcal{D}, \delta_t, \xi_q) \tag{14.3}$$

Algorithm 1 implements this selection.

Algorithm 1 MondrianTreeNodeOptimization($\mathcal{D}$, T, Q)

1: **Input**: node data $\mathcal{D} := \{(X_i, y_i)\}_{i=1}^{n}$, number of feature and threshold candidates (T, Q)
2: **Output**: split feature and location (δ^*, ξ^*)
3: **Algorithm:**
4: Initialize best Information Gain (IG): $g^* = -1$
5: **for** $t = 1, ..., T$ **do**
6: Choose δ_t with probability proportional to $(X_{\delta_t})_{max} - (X_{\delta_t})_{min}$ ▷ *dimension-wise* min *and* max
7: **for** $q = 1, ..., Q$ **do**
8: Choose ξ_{tq} uniformly from interval $[(X_{\delta_t})_{min}, (X_{\delta_t})_{max}]$
9: $g = IG(\delta_t, \xi_{tq}, \mathcal{D})$
10: **if** $g^* < g$ **then**
11: $(g^*, \delta^*, \xi^*) = (g, \delta_t, \xi_{tq})$
12: **end if**
13: **end for**
14: **end for**

Note that if we set T and Q to 1, we fall back to the original MF split procedure where no label is involved. We can notice that the selected feature has a probability proportional to the length of its range of values (line 6). This is due to the Mondrian process [7] which is the probabilistic generative process of the Mondrian tree construction. It has the effect of selecting features with a wider range of values more often than smaller ones.

Given we plug in this algorithm in the Mondrian tree construction, what is the result on the performance? Does that really improve the performance when low predictive power features are present? How does it compare to the classical random forests?

The next section attempts to answer these questions empirically.

Experiments

The main purpose of our experiments is to study the impact of using labels during splits in Mondrian tree and forests. Below are specific experiments we are going to conduct with their associated goals:

- Experiment 1: effect of label guided splits relative to the original Mondrian forest on batch mode
- Experiment 2: effect of labels on the online mode performance and its relation to batch mode
- Experiment 3: comparison with classical random forests on batch mode

We have used four datasets: usps [8–10], letter [11], and dna [12]. Table 14.1 presents an overview of these datasets. Appendix provides more information about them and technical specifications of the experiments.

Table 14.1 Brief structural description of the four datasets used for experiments

	Datasets			
Description	usps	satimage	letter	dna
No. of features	256	36	16	180
No. of classes	10	6	26	3
Training set size	7291	3104	15, 000	1400
Testing set size	2007	2000	5000	1186

Also, we used the software developed by Balaji, the main author of the original Mondrian forest paper [1], ported to Python 3 and adapted it to implement our method accessible on GitHub.[1] For comparisons with random forests, we used the sklearn implementation [14].

Results and Analysis

All results in this section are reported performance on test partitions of each dataset (Table 14.1).

Experiment 1

Figure 14.2 shows the results of using labels to guide splits during Mondrian forest training. In the case of one tree, the improvement is significant with a jump between 5% and 15% in accuracy. However, when ensembling trees, results are quite similar with a slight improvement (1.3%) in favor of our approach. One result is very singular, the dna dataset. We observe a big improvement of 25.3% in accuracy which is quite remarkable. This answers positively our question as the dna dataset contains many low predictive power features. Therefore using labels improved drastically the performance of Mondrian forests in the presence of low predictive power features and thus reinforces its robustness.

Experiment 2

Case of One Tree

Looking at Fig. 14.3a, we see an important variance in results for online mode as we run the experiments five times (blue and green plots). This is a well-known

[1] https://github.com/iskode/mondrianforest-labelsplit/

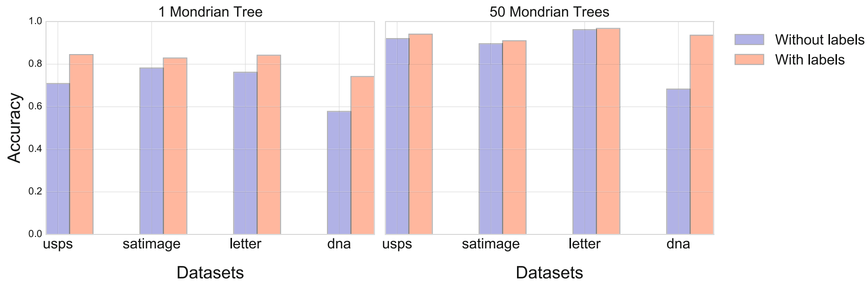
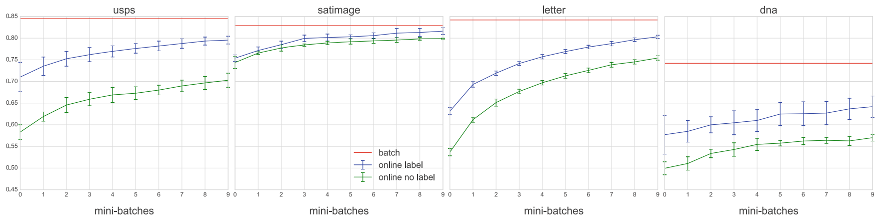
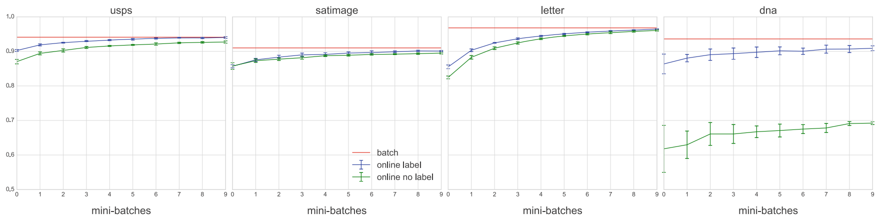


Fig. 14.2 Experiment 1 results. Comparison between the original MF and our modified version in batch mode on the four datasets



(a) 1 Mondrian Tree



(b) 50 Mondrian Trees

Fig. 14.3 (**a**) Comparison between performance of a Mondrian tree (**a**) trained with (blue) and without (green) labels in online mode and the reference batch mode without label (red). (**b**) The same for 50 trees. Vertical segments quantify each mini-batch's variance

property of decision tree: They have moderate to high variance [16, 17]. This variance is bigger for our approach than the original Mondrian tree across all four datasets, despite, being less accurate, it seems that the original Mondrian tree is more stable. Additionally, we notice that a Mondrian tree trained on batch mode gives consistently significant better results than being trained online on the same dataset.

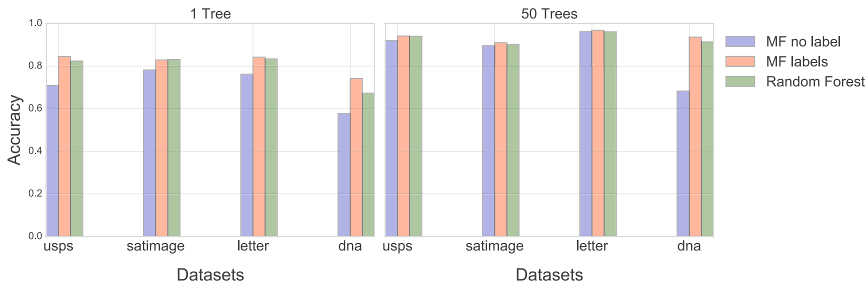


Fig. 14.4 Comparison between classical random forest and our extended Mondrian forest

Case of 50 Trees

Now (Fig. 14.3b), we notice a significant reduction of variance due to ensembling [15–17]. Conversely, on the dna dataset, the variance has increased for the original Mondrian forest. It seems that ensembling trees make it stable. For our approach, the variance decreases but is still significant. That is a sign that we can increase the number of trees to still improve the performance. Now, the online version of our method matches approximately its batch counterpart except for the dna dataset where the former is closer but still below the later.

Experiment 3

Looking at Fig. 14.4, we notice that our method is equal to or slightly better than random forest from sklearn library for both one tree and a forest of 50 trees. The difference is more nuanced on the dna dataset. Also the gap is more important for a single tree than the ensemble of 50 trees. We can notice that the original Mondrian forest has comparable performance to both our extension and random forests [1] except when the dataset contains many low predictive power features as of the dna.

Related Work

There is a rich and huge literature on Random forest algorithms. One of the main best reviews on the subject is the technical report [18]. There are many specific random forest algorithms [19] sharing similarities with the original Mondrian forests as pointed by the authors [1]. Recall that MF has two main desirable features [1]:

1. It performs Bayesian inference by a hierarchy of normalized stable process (HNSP).

2. It can be trained on batch, and online and theoretical results show their distributions are both equivalent.

There are many variants of random forest algorithms that only focus on one of those features and never accumulate them.

We have many random forest algorithms that perform Bayesian inference [20, 21]. Most of them impose a prior over the decision trees that depends on the data and approximate the posterior by Markov Chain Monte Carlo (MCMC) techniques [22]. One of the problems of these approaches is that they are computational heavy. It has been also shown that Bayesian model average performs worse than ensemble model combination [23]. Conversely, MF is on the one hand an ensemble model combination which is better than Bayesian model average. On the other hand, it performs Bayesian inference approximation at the tree level with efficiency comparable to top-down traversal in classical random forest inference.

On the online tree algorithms, we have ORF-Saffari [24] and ORF-Denil [25]. However, they are memory inefficient [1]. Other algorithms focus only on the online growth of trees, not their ensemble, which are better than individual trees.

It is clear that Mondrian forests uniquely gather these two interesting properties that are difficult to gather in a random forest-based algorithm. But they lack the involvement of label for finding optimal splits. Therefore, our contribution to use label during splits fills this gap and shows on par or better performance than classical random forests.

One important remark is that using labels makes splits dependent on the order of data arrival. It means different orders of batch of data during online training will produce different splits and then different tree distributions. Therefore, the theoretical guarantee that batch and online MF distributions match does not hold anymore. Nonetheless, our experiments show that they are nearly equal except on the dna dataset where the gap is 3.5% which is still close.

Conclusion

In this chapter, we have extended Mondrian forests, a class of random forests, which uniquely gather many sought properties such as efficiency, well-calibrated probability assessment, and online training. Our extension incorporates label information during splits using information gain maximization based on the entropy measure. Our experiments show an overall improvement in batch and online settings. Additionally, the resulting Mondrian forests yield equal or better performance than classical random forests based on our experiments. Over and above conducting more experiments on a large number of datasets, a good research venue would be to add the same extension to Mondrian forests for regression [26].

Acknowledgments We would like to thank Balaji Lakshminarayanan, the main author of the original Mondrian forests for his availability to discuss and answer our questions about his fantastic work.

Appendix

Datasets

We will use the same four datasets as in the original Mondrian forests paper [1]:

- usps [8, 9] represents normalized handwritten digits, automatically scanned from envelopes by the US Postal Service.
- satimage [10] consists of the multi-spectral values of pixels in 3×3 neighbourhoods in a satellite image and the classification associated with the central pixel in each neighbourhood.
- letter [11] represents English alphabet recognition based on black and white rectangular pixels.
- dna [12] represents 180 indicator binary variables where each 3 stands for nucleotides among only A.G,T,C. In particular, it has many low predictive power features [13]; precisely, only features from A61 to A120 (60 out of 180) provide good performance, which is suitable to answer our hypothesis.

usps, satimages, letter, and dna datasets [8, 9]. The datasets (Table 14.1) have a different level of difficulties with a varying number of features and training versus testing size. In particular, the dna dataset has many low predictive power features [13], precisely only features from A61 to A120 (60) provide good performance, which is suitable to answer our hypothesis.

Technical Specifications

We used the software developed by Balaji, the main author of the original Mondrian forest paper[2] and adapted it to implement our method. For comparisons with random forests, we used the sklearn implementation [14].

We will run our experiments on a single tree (DT or MT) and on a forest of 50 trees (RF or MF).

For training MT or MF, we set split parameters (number of split candidates) $T = \sqrt{D}$ as is the default value of sklearn random forests and $Q = 6$ as it gives a good result across all datasets. But these parameters can be further optimized via

[2] https://github.com/balajiln/mondrianforest

cross validation as part of the model selection. Additionally, we do not use bootstrap which consists of using a different subset of the data to train each tree.

In order to compare fairly RF from sklearn and our MF, we use the same settings as MF. In other words, we used the following settings (from sklearn API) for the RF:

- `n_estimators = 1` or `50` (number of trees)
- `criterion = 'entropy'` (information gain IG)
- `bootstrap = False`
- `max_features = 'auto'` equivalent to $\sqrt{D}$

References

1. Lakshminarayanan B., Roy D. M., and Teh Y. W., Mondrian Forests: Efficient Online Random Forests, Advances in Neural Information Processing Systems, 27, 3140–3148 (2014)
2. Breiman, L. Random Forests, Machine Learning, 45, 5–32 (2001)
3. Siddharth Misra, Yaokun Wu, Machine Learning for Subsurface Characterization, 289–314. Gulf Professional Publishing, United States (2020)
4. Niculescu-Mizil A. and Caruana R., Predicting Good Probabilities with Supervised Learning, Proceedings of the International Conference on Machine Learning - ICML, 22, (2005)
5. Lakshminarayanan B., Thesis: Decision Trees and Forests: A Probabilistic Perspective, UCL, United Kingdom (2016)
6. Shannon C. E., A Mathematical Theory of Communication, The Bell System Technical Journal, 27, 379–423 (1948)
7. Daniel M. Roy, Yee Teh, The Mondrian Process, Advances in Neural Information Processing Systems, 21, 1377–138 (2008)
8. Hull J. J., A database for handwritten text recognition research, IEEE Transactions on Pattern Analysis and Machine Intelligence, 16, 550–554 (1994)
9. Feng C., Sutherland A., King S., Muggleton S. & Henery R., Comparison of Machine Learning Classifiers to Statistics and Neural Networks, AI & Stats Conference, 93(1993).
10. Srinivasan A., Statlog (Landsat Satellite). UCI Machine Learning Repository (1993)
11. Slate D., Letter Recognition. UCI Machine Learning Repository (1991).
12. Michalski,R.S. and Chilausky,R.L.., Soybean (Large). UCI Machine Learning Repository (1988).
13. Vanschoren J., Rijn J. N. V., Bischl B. and Torgo L., OpenML: networked science in machine learning. SIGKDD Explorations 15(2), pp 49–60, 2013
14. Pedregosa et al., Journal of Machine Learning Research, 12, 2825–2830 (2011)
15. Breiman L., “Bagging predictors”, Machine Learning, 24(2), 123–140, 1996.
16. Dietterich T. G, Kong E. B., Machine Learning Bias, Statistical Bias, and Statistical Variance of Decision Tree Algorithms, 1–14. Technical Report. Department of Computer Science, Oregon State University (1995)
17. Breiman L., Bias, variance, and arcing classifiers, Technical Report, 460, Statistics Department, University of California, Berkeley, CA, USA (1996)
18. Criminisi A., Shotton J., Konukoglu E., Decision Forests: A Unified Framework for Classification, Regression, Density Estimation, Manifold Learning and Semi-Supervised Learning, Foundations and Trends Computer Graphics Vision, 7, 81–227 (2012)
19. Geurts P., Ernst. D., and Wehenkel L., Extremely randomized trees, Machine Learning, 63, 3–42 (2006)

20. Chipman H. A., George E. I., and McCulloch R. E., Bayesian CART model search, 93, Journal of the American Statistical Association, 935–948 (1998)
21. Denison D. G. T., Mallick B. K., and Smith A. F. M., A Bayesian CART algorithm, Biometrika, 85, 363–377 (1998)
22. Chipman H. A., George E. I., and McCulloch R. E., BART: Bayesian Additive Regression Trees, The Annals of Applied Statistics, 4, 266–298 (2010)
23. Domingos P., Bayesian Averaging of Classifiers and the Overfitting Problem, International Conference of Machine Learning, 747, 223–230 (2000)
24. Saffari A., Leistner C., Santner J., Godec M., Bischof H., On-line random forests. Computer Vision Workshops - ICCV Workshops, 12, 1393–1400 (2009)
25. Denil M., Matheson D., Freitas N., Consistency of online random forests, International Conference of Machine Learning, 28, 1256–1264 (2013)
26. Lakshminarayanan B., Roy D. M., and Teh Y. W., Mondrian Forests for Large-Scale Regression when Uncertainty Matters, Proceedings of the 19th International Conference on Artificial Intelligence and Statistics, JMLR: W&CP, 51, 1478–1487 (2016)

Chapter 15
Remote Sensing of Artisanal Mines Buried in the Ground by Infrared Thermography Using UAV

Adama Coulibaly, Ibrahima Ngom, Jean Marie Dembele, Ibrahima Diagne, Ousmane Sadio, Marc Momar Tall, Moustapha Ndiaye, and Abdou Diop

Abstract The antipersonnel and anti-tank landmines create a lot of human and material damage in the Sahel countries affected by terrorism. Explosive mine detection methods are based on tools handled by human operators and target industrial metal mines. These methods are risky and limited because the types of mines most commonly used in the Sahelian context are mainly homemade and are encased in various local materials such as metal, plastic, glass, ceramic, or wood. This chapter presents a solution for remote sensing of artisanal mines buried in the ground using infrared thermography. A DJI Phantom 4 Quadcopter equipped with a FLIR thermal camera and a GNSS sensor performs an automatic low-level flyover of the potentially mined road. Thermal images of the road are collected with an overlap rate of 80% and referenced with the GNSS sensor. Photogrammetry algorithms are used to process the thermal images to detect and locate anomalies related to the presence of buried mines. Despite the limitations due to environmental influences, the model showed a detection rate of 75% during flights at an altitude of 6 m and a speed of 3 m/s. The experimental results show a good correlation between the

Authors Ibrahima Ngom, Jean Marie Dembele, Ibrahima Diagne, Ousmane Sadio, Marc Momar Tall, Moustapha Ndiaye, and Abdou Diop have contributed equally to this work

A. Coulibaly (✉)
Applied Sciences and Technologies, University Gaston Berger, Nationale 2, Saint Louis, Senegal

Ecole Supérieure Polytechnique, University Cheikh Anta Diop, Dakar, Senegal
e-mail: adama.coulibaly@esp.sn

I. Ngom · I. Diagne · O. Sadio · M. M. Tall · M. Ndiaye · A. Diop
Ecole Supérieure Polytechnique, University Cheikh Anta Diop, Dakar, Senegal
e-mail: ibrahima.ngom@esp.sn; ibrahima.diagne@esp.sn; ousmane.sadio@esp.sn; marcmomar.tall@esp.sn; moustapha.diagne@esp.sn; abdou.diop@esp.sn

J. M. Dembele
Applied Sciences and Technologies, University Gaston Berger, Nationale 2, Saint Louis, Senegal
e-mail: jean-marie.dembele@ugb.edu.sn

C. T. Gueye et al. (eds.), *Mathematics of Computer Science, Cybersecurity and Artificial Intelligence*, Springer Proceedings in Mathematics & Statistics 468,
https://doi.org/10.1007/978-3-031-66222-5_15

thermal contrast of the mathematical model and the cooler areas containing a mine-related chemical substance.

Keywords Artisanal explosive mines · Buried in the ground · Infrared thermography · Remote sensing · UAV

Introduction

Burkina Faso is facing a serious security crisis manifested by terrorist acts of all kinds in several regions of the country. The army, in the process of recapturing areas under terrorist control, is carrying out offensive actions against armed groups and also humanitarian actions for the benefit of the populations of regions cut off from the capital. However, the pursuit of these actions by road is strongly disrupted by improvised explosive devices consisting of antipersonnel and anti-tank mines. These types of homemade explosive landmines are encased in wood, ceramic, or plastic and are difficult to detect because they are completely buried in the ground. Existing landmine detection methods are based on metal sensors manipulated by human operators or mounted on demining vehicles. These methods are ineffective in searching for plastic, ceramic, or wooden cased mines. They also present risks related to the proximity of the operators to the explosive devices. With the technological advances marked by the advent of artificial intelligence, drones are increasingly used in several fields, including landmine detection. They have the advantage of integrating different types of mine sensors and being autonomous or remotely piloted, which improves safety because the operator is distant from the dangerous area. The objective of this chapter is to study a solution for remote sensing of homemade mines by infrared thermography using UAV. The remainder of this chapter is organized as follows. Section "Related Works" provides an overview of related works on the topic. Section "Materials and Methodology" describes the working methodology and materials used. Section "Results" presents the experimental results of the studied model. And section "Discussions" is devoted to the discussion of the results.

Related Works

A great deal of research has been conducted in the field of explosive landmine detection. These researches have focused on both detection techniques and communication of information. Gooneratne et al. [1] and Siegel [2] reviewed landmine detection technologies. They identified two main families of mine detection methods: biological (guard dogs, rats, plant leaves, bees, and bacteria) and technological (metal sensors, nuclear quadruple resonance (NQR), ground-penetrating radar (GPR), laser, X-ray, ultrasound, and microwave). Their study revealed that most technologies could be installed on a robot or drone to avoid loss of life. Yoo et al.

[3] worked on a drone equipped with a magnetometer for metal mine detection. Their prototype includes a ground control system, a flight control system, and a magnetometer. The metallic AP (M16) and AT (M15) mines and low metal content AT (M19) mines were detected successfully in field tests under the proposed conditions. The low metal content AP (M14) mine was not detected. Castiblanco et al. [3] proposed a low-cost ARdrone 2.0 drone as a complementary tool for visual landmine detection in rural scenarios. Their system consists of a UAV and a base station that manages both the flight controls and the landmine detection algorithms. Experimental results show an effective percentage of the detection over 80% at an altitude of 1 m with a flight speed of 2.2 m/s. But their prototype does not take into account the precise location of the detected mine. Ganesh et al. [4] adopted a two-step detection methodology: detection of the metal mine and its photography by an infrared camera. The system consists of a metal detector, an infrared camera, a GPS sensor, a GSM modem, and an Arduino UNO board. The operation of their system is independent of the drone. Their system was able to detect mines and send a message containing the longitude and latitude of the detected mine by the GPS using a GSM module. Colorado et al. [5] integrated computer vision algorithms into a UAV to perform terrain mapping and visual detection of landmine-like objects in real time. Despite hardware limitations, their system was able to detect partially buried objects in different types of terrain with a detection percentage of over 80%. The authors [6, 7] have worked on mine detection using infrared thermography. Their system is not embarked on a drone but allows detecting in the invisible. Their experience has shown that mine detection using infrared thermography is valid under specific conditions such as on fine sand. Gracias et al. [8] worked on a dataset of thermographic images for the detection of buried landmines. Their work consisted of capturing aerial infrared images of a terrain where elements with characteristics similar to antipersonnel mines type leg-breaker were buried. The dataset had 2700 thermographic images acquired at different heights, using a Zenmuse XT infrared camera, embedded in the DJI Matrice 100 drone. An automatic detection methodology for leg-breaker Antipersonnel Landmines (APLs) was developed by Forero-Ramirez et al. [9] based on digital image processing techniques and pattern recognition, applied to thermal images acquired by means of an Unmanned Aerial Vehicle (UAV) equipped with a thermal camera. They obtained remarkable results using a Multilayer Perceptron (MLP) classifier, reaching average percentages of success in detecting suspicious areas with the presence of these artifacts about 97.1% for images acquired at 1 m from the ground and 88.8% at higher altitudes. Baur et al. [10] focused their work on developing and testing an automated technique of remote landmine detection and identification of scatterable antipersonnel landmines in wide-area surveys. Their findings are based on the analysis of multispectral and thermal datasets collected by an automated UAV-survey system, which used scattered PFM-1-type landmines as test objects. The study also presents results from efforts to automate landmine detection using supervised learning algorithms, specifically a Faster Regional-Convolutional Neural Network (Faster R-CNN). The RGB visible light Faster R-CNN demo yielded a 99.3% testing accuracy for a partially withheld testing set and 71.5% testing accuracy for a completely withheld testing set.

These related works have explored several robots and UAV-based solutions for detecting landmines. But most of these solutions are only applicable to industrial metal mine partially buried. The scope of these researches does not include artisanal mines, which are encased in a variety of local materials with different properties.

Materials and Methodology

Method

The method used in this chapter includes three phases.

1. **Data collection**: Dji Phantom 4 UAV equipped with FLIR thermal camera and GNSS sensor performs an automatic mapping mission of the road to be inspected at an altitude of 6 m and a speed of 3 m/s. The geo-referenced thermal images are captured with a coverage rate of 80% and sent in real time to the supervision tablet via the radio control.
2. **Image alignment and reconstruction of the road to be inspected**: The images are imported into the FLIR photogrammetry tool, which allows them to be grouped and matched using GNSS coordinates in order to obtain a complete thermal image of the inspected road.
3. **Thermal processing**: The photogrammetry algorithm processes the thermal contrast of the image to identify temperature differences and generate reports. The target areas are then located using their GNSS coordinates.

Presentation of the Explosive Mine

A mine is a device designed to be placed under or on the ground and to be exploded by the presence, proximity, or contact of a person or a vehicle [11]. If the mine is intended to disable, injure, or kill one or more persons, it is termed antipersonnel (AP). When the target is a vehicle, the mine is called an anti-tank (AT) mine. In general, the mine consists of an ignition device (igniter, detonator), a warhead, a safety device, and the whole contained in an envelope with projectiles (refer Fig. 15.1a) [12]. The homemade variant of the mines (refer Fig. 15.1b) contains an excessive quantity of composite products including explosives, recovered ammunition (bombs, shells, etc.), and various casings (metal, plastic, wood, and ceramic).

Infrared Thermography

In the electromagnetic spectrum, infrared is located between the visible and the microwaves. The main source of infrared radiation is heat or thermal radiation. Any object with a temperature above absolute zero (−273.15C or 0K) emits radiation in

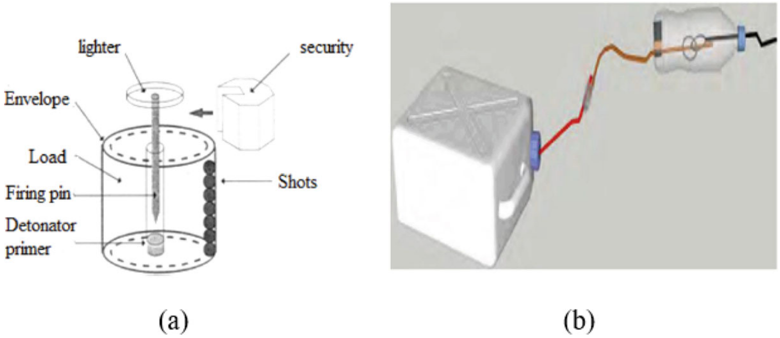


Fig. 15.1 Mine overview [12, 13]. (**a**) Structure of mine. (**b**) Plastic coated mine

the infrared range. The infrared thermography is the process consisting in carrying out thermal images by infrared by highlighting by sets of color the differences of temperature [14]. The thermal imaging or thermogram consists in transforming measurements of the infrared radiation into a radiometric image corresponding to the values of temperature. Thus, each pixel of the radiometric image is a temperature measurement. Thermal cameras are designed to transform infrared signatures into something visible to humans.

Mathematical Modeling

The detection of a mine buried in the ground is based on the fact that the ground does not return infrared in the same way depending on whether or not there is a buried mine [15]. Thus, thermal infrared detection is based on the differential heating or cooling of the ground. This difference is small and is only noticeable under favorable thermal conditions. The night is best suited for the detection of the mine by the thermal method. The principle is as follows:

1. During the day, the mine (plastic, wood, and ceramic) does not conduct the Sun's heat well, so that the thermal energy supplied by the Sun remains at the surface: The ground above the mine is warmer during the day [15–18].
2. At night, since this surface has not collected heat, it does not release any and remains relatively cold [15–18].

But several parameters must be taken into account, some of which are shown in Fig. 15.2:

– The emissivity of the buried mine $e = l - p$ (p reflection factor of the body and l the wavelength)
– The albedo $A = \frac{R_r}{R_s}$ (with R_r the reflected radiation and R_s the received solar radiation)

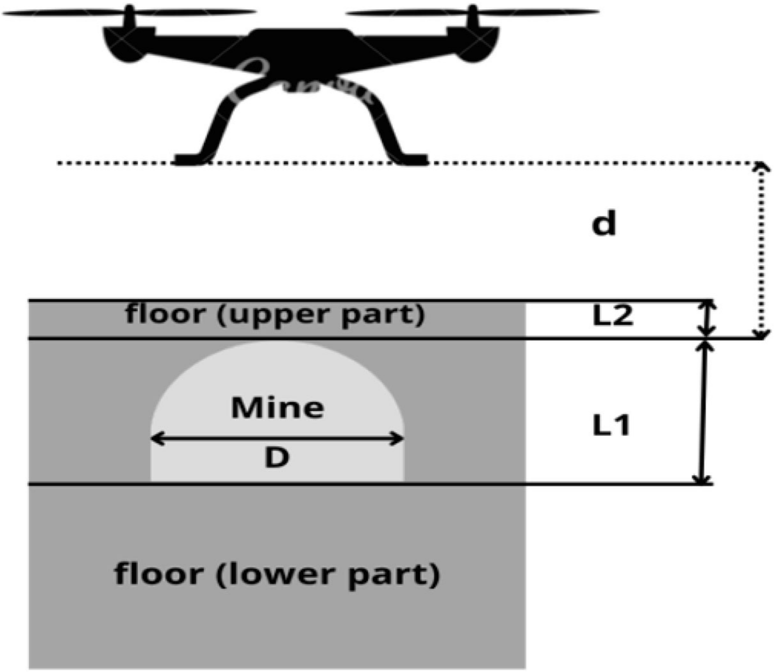


Fig. 15.2 Mathematical model [9]

- The temperature of the environment T_{env}
- The atmospheric temperature T_{atm}
- The altitude of the drone d
- The depth of the mine L_2
- The received solar radiation R_s
- The type of mine

For a black body, the spectral radiance is given by Planck's law:

$$L_{0(l,T)} = \frac{2 \times h \times c^2}{(e^{\frac{h \times c}{l \times k \times T}} - 1) \times l^5} \text{ [9]} \tag{15.1}$$

with $L_{0(l,T)}$: spectral radiance $(W \times m^{-3} \times sr^{-1})$
h: Planck's constant $(6.62617 \times 10^{-34} \times J \times s)$
c: celerity $(299792458 m \times s^{-1})$
k: Boltzmann's constant $(1.38 \times 10^{-23} \times J \times K^{-1})$
T: absolute temperature of black body in Kelvin
Radiation absorbed by the surface $R_{abs} = (1 - -A) \times R_s$ with $A = 36\%$ [17] for the Earth which gives $R_{abs} = 64\% \times R_s$. The soil has a more or less great capacity to

store thermal energy by conduction in depth during the day and to restore it during the night. This capacity is called thermal inertia noted I:

$$I = \sqrt{k_t \times p_c} \tag{15.2}$$

with k_t thermal conductivity, p_c volumetric heat capacity, and $I(J \times m^{-2} \times K^{-1} \times s^{-1/2})$.

Stefan Boltzmann's law gives the power of thermal radiation captured by the camera:

$$P = e \times j \times T^4 \tag{15.3}$$

with j Stefan's constant ($5.671 \times 10^{-8} \times W \times m^{-2} \times K^{-4}$ and $e = 1$ case of black body ($e < 1$ for real materials) [7].

Let

$$y = \frac{h \times c}{l \times k} \tag{15.4}$$

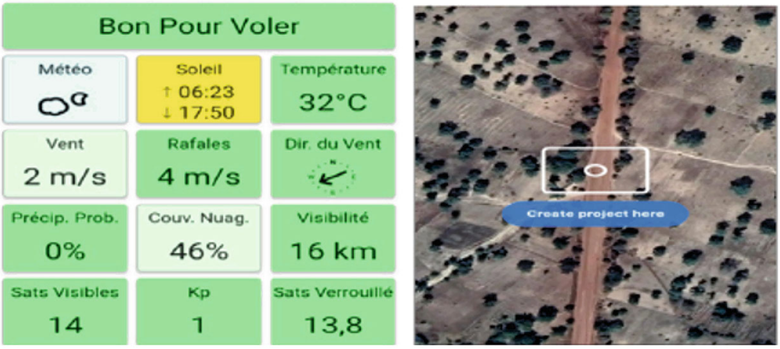
The radiometric temperature is written as

$$T_l = \frac{y}{ln \frac{2 \times y \times c \times k}{l^4 \times L_{0(l,T)} + 1}} \; [9] \tag{15.5}$$

Materials and Tools Used

The experiment required a DJI Phantom 4 drone (Fig. 15.3) equipped with a FLIR View 336 thermal camera and GNSS geotagger to perform the low-level automatic mapping mission of the potentially mined road. The high-resolution long-wave thermal camera (7.5−13.5 um) is mounted on a dial that stabilizes the image and

Fig. 15.3 UAV, radio control, and FLIR control module



(a) (b)

Fig. 15.4 Mission planning. (**a**) Atmospheric conditions. (**b**) Site of the experiment

allows the lens to rotate 360°. The drone has a 15 minutes flight time and 128 GB internal storage capacity. A tablet combined with the radio control and the FLIR thermal control module is used to supervise the flight. Processing is done on a PC HP EliteBook Intel(R) Core(TM) I7-7600U, CPU @ 2.8 GHz 2.9 GHz with 16 GB of RAM. Twelve inert landmines are used for the experiment. The software tools used are:

– UAV Forecast for pre-flight weather and safety forecast (Fig. 15.4a)
– PiX4Dcapture for the planning of the mapping mission (Fig. 15.4b)
– DJI GO 4 for flight supervision
– FLIR UAS for thermal camera setup
– FLIR Tools Studio for thermal image processing (Fig. 15.6)

Results

The experimental study is carried out on an unpaved road with a maximum width of 6 m (Fig. 15.4b). For this study, we used inert mines which are safe in handling and are used to study the operation of the real mine. The model used is materialized in Fig. 15.2 with the parameters of Table 15.1. The values of emissivity e and albedo A are chosen according to the soil type of the road portion. The albedo is 36% for soil, and the emissivity is 0.39 for clay soil [17]. The parameters L_1 and D are the actual dimensions of the mine envelope used. As for the value of L_2, it represents the maximum depth from which thermal detection is possible according to the sunshine data of the area. The section of road used for the experiment is 1.5 km long. Twelve inert handmade mines distributed in Table 15.2 are buried at random 100 m apart along the road for the experiment.

Table 15.1 Parameters table

Environment parameters		UAV-mine parameters			
Emissivity	Albedo	Altitude	Height	Width	Depth
e	A	d	L_1	D	L_2
0.39	0.36	6 m	15 cm	40 cm	5 cm

Table 15.2 Repartition of mines

Plastic	Metal	Wood	Ceramic
3	3	3	3

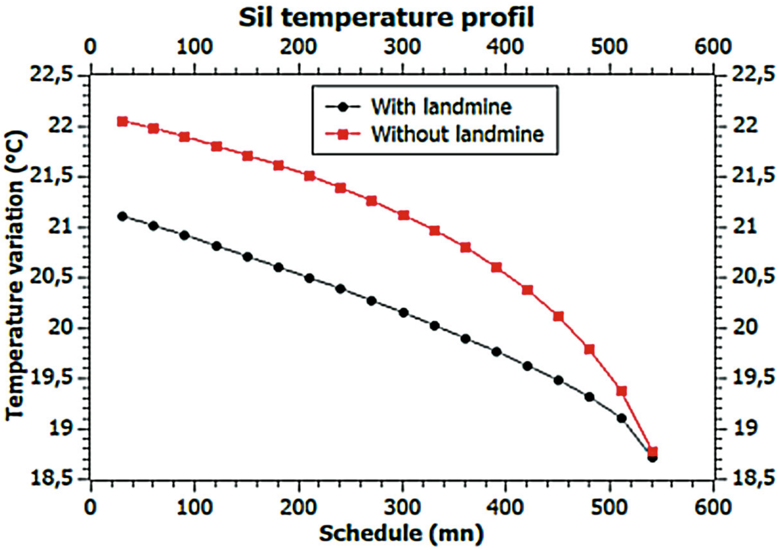
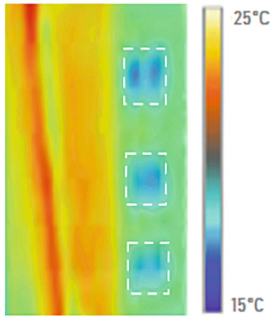


Fig. 15.5 Temperature variation of ground with and without buried mine

The experiment is carried out in two phases:

1. **Phase 1**: To highlight the difference in temperature between the soil with mine and the soil without mine, we extract the temperature profiles of the two surfaces using an infrared thermometer. This phase takes place in the time interval from 8:00 pm to 5:00 am. The temperature of the two surfaces close to the road, one over a plastic-coated mine and the other without a mine, is measured every 5 minutes. The temperature profile readings from this experiment are plotted in Fig. 15.5. The observation is that the temperature of the soil with mine is lower than that of the soil without mine, and the average difference is 0.818C. This reflects the fact that the area above the buried mine is relatively cooler than the area without mine.
2. **Phase 2**: The UAV then conducts mapping flights of the area of interest according to the scheduled missions to collect thermal images. Three mapping missions were scheduled at night time slots: mission 1: 6:00 pm–7:00 pm, mission 2:

Parameters	
Emissivity	0.39
Reflected temperature	72.7°F
Distance	6.00m
Atmospheric temperature	57.7°F
Relative humidity	28%
External optics temperature	-
External optics transmission	-



(a) (b)

Fig. 15.6 Thermography results. (**a**) Infrared image. (**b**) Thermal treatment result

Table 15.3 Detection results

Mission	Plastic	Metal	Wood	Ceramic	Total	
Mission 1	0	0	0	0	0	0%
Mission 2	3	0	3	3	9	75%
Mission 3	1	0	0	2	3	25%

8:00 pm–9:00 pm, and mission 3: 11:00 pm–0:00 am. The choice of these three time intervals is motivated by the variations observed in the previous experiment. According to the parameters set on the drone: speed (3 m/s), altitude (6 m), and coverage rate (80%), each mission collected 750 thermal images. These images from each mission are processed by the FLIR Report Studio photogrammetry tool. The results of the image alignment and the thermal contrast analysis report are given in Fig. 15.6a, b, respectively. The color palette and the target temperature range are automatically set by the photogrammetry tool but can be modified if needed. The colors of the palette have the following meanings: Yellow indicates a hot zone, green symbolizes a medium temperature, and blue corresponds to the cold zone.

The results of mine detection during the three mapping missions are shown in Table 15.3. Mission 1 detected no mine, i.e., a rate of 0%. Mission 2 detected 9 mines out of 12, including plastic-cased, wooden, and ceramic mines, i.e., a rate of 75%. Mission 3 detected one plastic and two ceramic mines, i.e., a rate of 25%. Each detected mine is located through its associated coordinates.

Discussions

The experimental results of the system show a high detection rate for plastic, ceramic, and wooden cased mines during mission 2. The detection rate is low for the two other missions and also for the metal clad mines. Indeed, metal has a high thermal conductivity compared to wood, plastic, and ceramic. Therefore the results of the phase 1 experiment are not verified with the metal mine. The low detection rates during missions 1 and 3 are explained by the heat restitution conditions during these periods. In Burkina, the time interval from 6:00 pm to 7:00 pm corresponds to nightfall. Thus, the process of heat restitution is not yet effective, which does not allow us to observe the differences in temperature between zones. As for the time interval of mission 3, it corresponds to a period when the ground is quite cooled, which also does not allow the differentiation of temperature between different areas. Our solution focuses on the detection of artisanal mines completely buried on the roads. The difference between an industrially manufactured mine and a handcrafted mine lies essentially in the envelope. Industrial lead is generally metal-jacketed, while artisanal lead can be plastic, wooden, or ceramic-jacketed. These local materials do not have the same physical characteristics as metal. Previous work has focused on the detection of metal mines partially buried in the ground. In our case, the mine is completely buried in the ground, making detection uncertain. Our results are therefore better for plastic, wood, and ceramic mines, confirming the results of our first experiment. But our experiment failed to detect metal mines, whereas earlier work showed better results for metal mines and limited results for plastic, wood, and ceramic mines. Our results would be negatively impacted by the effect of vegetation and also by the presence of other types of objects buried on the road. Thus extraneous objects may constitute false positives in the results. Taking environmental factors into account and choosing the right period for data collection will ensure reliable results.

Conclusion

The remote detection of explosive mines buried in the ground remains a major concern for Sahelian countries facing the terrorist phenomenon. In this chapter, we have studied a solution for remote sensing of artisanal mines buried in the ground by infrared thermography using a drone. This first experiment that concerned roads gave satisfactory results for wooden, plastic, and ceramic mines, despite the influence of many external factors. Each detected target is located through its GNSS coordinates provided. The accuracy in localization can be improved by using an RTK drone with centimeter accuracy. As for the accuracy of the detection by infrared thermography, it is strongly dependent on environmental factors, the judicious choice of the period of collection of the thermal images as well as the setting up of the photogrammetry tool.

The performance of the model used in this chapter can be improved by adding the thermal and magnetic characteristics of metal mines for a complete solution.

Acknowledgments This work received funding from the Partnership for Skills in Applied Sciences, Engineering, and Technology/Regional Scholarship and Innovation Fund (PASET/RSIF).

References

1. Gooneratne, C. P., Mukhopahyay, S.C., Sen Gupta, G.: A Review of Sensing Technologies for Landmine Detection: Unmanned Vehicle Based Approach. In: 2nd International Conference on Autonomous Robots and Agents, pp. 401–407, IEEE, New Zealand, (2004).
2. Siegel, R.: Land Mine Detection, IEEE Instrumentation and Measurement Magazine, vol. 2, 22–28 (2002)
3. Yoo, L.S., Lee, J.H., Ko, S.H., Jung, S.K., Lee, S.H., Lee, Y.K.: A Drone Fitted with a Magnetometer Detects Landmines, In: IEEE geoscience and remote sensing letters, vol. 49111, pp. 1–5, South Korea (2019)
4. Castiblanco, C., Rodriguez, J., Mondragon, I., Parra, C., Colorad, J.: Air Drones for Explosive Landmines Detection. In: ROBOT2013: First Iberian Robotics Conference, Advances in Intelligent Systems and Computing, pp. 107–114, Springer International Publishing Switzerland (2014).
5. Ganesh, Y., Raju, R., Hegde, R.: Surveillance Drone for Landmine Detection. In: International Conference on Advanced Computing and Communications, pp. 33–38, IEEE, (2015)
6. Colorado, J., Mondragon, I., Rodriguez, J., Castiblanco, C.: Geo-Mapping and Visual Stitching to Support Landmine Detection Using a Low-cost UAV, International Journal of Advanced Robotic Systems, pp. 1–12, (2015)
7. Abdelmoumni, N., Mansiaux, F., Xu, C.: Détection Des Mines Antipersonnel Par Thermographie Infrarouge, Olympiades de la physique, France pp. 1–23, (2009)
8. Tenorio-Tamayo, H. A., Forero-Ramírez, J. C., García, B., Loaiza-Correa, H., Restrepo-Girón, A.D., Nope-Rodríguez, S.E., Barandica-López, A., Buitrago-Molina, J.: Dataset of thermographic images for the detection of buried landmines, Elsevier Inc., vol. 49(109443), 1–8 (2023) https://doi.org/10.1016/j.infrared.2022.104307
9. Forero-Ramírez, J. C., García, B., Tenorio-Tamayo, H.A., Restrepo-Giron, A.D., Loaiza-Correa, H., Nope-Rodríguez, S.E., Barandica-Lopez, A., Buitrago-Molina, J.T.: Detection of "legbreaker" antipersonnel landmines by analysis of aerial thermographic images of the soil, Elsevier Inc., vol. 125(104307), 1–8 (2022) https://doi.org/10.1016/j.infrared.2022.104307
10. Baur, J., Steinberg, G., Nikulin, A., Chiu, K., De Smet, T.S.: Applying Deep Learning to Automate UAV-Based Detection of Scatterable Landmines, Remote Sens., vol. 12(859), 1–16, (2020). https://doi.org/10.3390/rs12050859
11. Antoine, R., Lopez, T.: Télédétection Dans l'infrarouge Thermique, Hal open science, vol. 02865102, pp. 156–203, (2020)
12. SLAM: Glossaire des termes et abréviations concernant l'action contreles mines, Nilam 04.10, 2nd édition, New York, (2003)
13. Franco-mines: Les Mines, Forum francophones des actions contre les mines.
14. GICHD: Guide des bonnes pratiques pour la neutralisation des engins explosifs improvisés, (2020)
15. http://www.flyingeye.fr/les-bases-des-drones-thermiques. Last accessed 30 Dec 2022
16. https://bibliomines.org/wp-content/uploads/les-techniques-dedeminage1.pdf. Last accessed 10 Jan 2023
17. Jolivot, A., Gomez-Candon, D., Labbé, S., Virlet, N., Regnard, J.L.: Acquisition d'images thermiques par drone : corrections radiométriques à partir de données ter-rain, pp-156–203 (2014)
18. Le Barny, P., Obert, E., Simic, V., Leray, I.: La Détection d'explosifs Vers de Nouvelles Solutions Technologiques, L'actualité chimique 104–107 (2007)

Chapter 16
Implementation of EdDSA in the Ethereum Blockchain

Mamadou Cherif Kasse and El Hadj Modou Mboup

Abstract Blockchain technology is widely used across various domains for its security and distributed ledger capabilities. To secure transactions, most blockchain platforms such as Ethereum employ the Elliptic Curve Digital Signature Algorithm (ECDSA).

However, the use of ECDSA can pose risks, such as the inadvertent exposure of the private key in case of errors, thus facilitating obtaining corresponding signatures for various documents. To address this issue, a solution emerges: the integration of the Edwards-curve Digital Signature Algorithm (EdDSA). By opting for EdDSA to generate transaction signatures, several advantages emerge, such as increased speed, optimal performance, and enhanced independence in random number generation. Indeed, this innovative proposition significantly bolsters security compared to the conventional use of ECDSA, marking a substantial advancement within the Ethereum ecosystem.

Furthermore, we have implemented both algorithms to sign and verify Ethereum transactions to make a performance comparison. The implementation is carried out in Python on an Intel Core i3 processor with 8 GB of RAM and a 64-bit operating system.

Keywords EdDSA · Ed25519 · ECDSA · Ethereum · Transaction · Blockchain

Introduction

The blockchain proves to be an exceptionally promising and revolutionary technology because it contributes to reducing security risks, eliminating fraud, and

---

M. C. Kasse (✉)
Université Cheikh Anta Diop de Dakar, FST, DMI, LACGAA, Dakar, Senegal

E. H. M. Mboup
Université Iba Der Thiam de Thiés, Dakar, Senegal
e-mail: elmodou.mboup@univ-thies.sn

C. T. Gueye et al. (eds.), *Mathematics of Computer Science, Cybersecurity and Artificial Intelligence*, Springer Proceedings in Mathematics & Statistics 468,
https://doi.org/10.1007/978-3-031-66222-5_16

establishing unparalleled transparency. Thanks to this technology, it becomes possible to ensure transparency in food supply chains, secure medical data, and reshape our approach to data processing and ownership. Blockchain technology also enables cryptocurrencies like Ethereum and other digital information to flow freely between individuals without the intervention of intermediaries.

By integrating this advancement, signature algorithms play a crucial role in consolidating this freedom of exchange. Indeed, through sophisticated digital signature mechanisms, the blockchain secures transactions while allowing for smooth and autonomous circulation of cryptocurrencies and digital data among concerned parties. This combination of technologies ensures increased integrity and trust in a constantly evolving digital environment. In the context of blockchain, a signature scheme generally works as follows:

- Each network participant generates a key pair: a private key and a public key.
- When a transaction is initiated, it is signed with the sender's private key, creating a unique digital signature attached to the transaction. This signature serves as proof of authentication and guarantees the transaction's integrity. Network nodes, using the corresponding public key, can verify the signature's validity and, consequently, the legitimacy of the transaction.

All these operations are based on public key cryptography, specifically elliptic curve cryptography, which relies on the properties of elliptic curves in finite fields.

The Ethereum platform uses the ECDSA, specifically with the secp256k1 curve, which is widely accepted and recommended by various standardization and normalization organizations, including the Federal Information Processing Standards (FIPS) [1], to provide transactional security. However, ECDSA uses a randomly generated value during the signature called a nonce, and it has been established that even a slight nonce leakage can potentially lead to full key recovery, as demonstrated by Diego F. et al. in [2]. It was mentioned in [3] that this vulnerability manifested in Sony's ECDSA implementation when signing code for the PlayStation 3, thereby exposing Sony's long-term secret key. ECDSA is also vulnerable to side-channel attacks. To illustrate, Nguyen and Shparlinski described in [4] an algorithm that exploits lattice techniques to derive the long-term ECDSA key using only 3 bits of the nonce value from several hundred signatures. These 3 bits can come from side-channel attacks.

However, EdDSA [5], developed by Daniel J. Bernstein [6], in addition to its security level and runtime performance, uses deterministic nonces, making it less vulnerable to the aforementioned attacks. The idea of producing signatures deterministically was proposed by Barwood in [7]. EdDSA encompasses two variants, Ed25519 and Ed448. Ed25519 is preferred over Ed448 because it requires fewer resources while ensuring entirely adequate security [5]. EdDSA is based on the Schnorr signature algorithm, which provides advantages such as the use of key and signature aggregation for multiple parties. These features are particularly beneficial for blockchain platforms like Ethereum because they allow for efficient grouping of keys and signatures, which can enhance system performance and overall efficiency. In other words, EdDSA offers advanced aggregation mechanisms that make it particularly well suited for blockchain scenarios where speed and efficiency are crucial. It has been proven that EdDSA is much faster in terms of signature generation and signature verification than ECDSA, as demonstrated by Guruprakash J et al. in [8]. With all these elements, EdDSA is a promising candidate for enhanced security with better performance in the Ethereum blockchain.

Related Work

The comparative evaluation of digital signature algorithms, particularly ECDSA and EdDSA, in the context of blockchains, is an active research topic. Numerous studies have been conducted to understand the advantages and disadvantages of each algorithm in terms of security, performance, and efficiency.

Guruprakash J et al. [8] conducted a comprehensive performance comparison between ECDSA and EdDSA, focusing on their usage in blockchain and IoT. Their results showed that, despite the higher complexity of EdDSA, it offered significant advantages in terms of signature computation time and energy consumption.

Research by Bernstein et al. [6] demonstrates that EdDSA is resistant to side-channel attacks and that its performance potentially surpasses that of ECDSA. Basha et al. [9] conducted an in-depth analysis of how the EdDSA can enhance the security of digital signatures in blockchain transactions. They highlight specific advantages that EdDSA brings in terms of resistance to side-channel attacks and fault attacks, making it an attractive choice for ensuring transaction integrity in blockchain environments.

Barenghi et al. [10] examined the vulnerability of ECDSA and EdDSA to fault attacks, concluding that the robustness of EdDSA was superior due to its deterministic construction.

Our Contributions

In this chapter, we propose the integration of the EdDSA digital signature algorithm (particularly, the Ed25519 variant) into Ethereum transactions, replacing the ECDSA. To do this, we first evaluate the performance of the ECDSA and EdDSA signature algorithm by implementing both algorithms in Python. We measured execution time, memory consumption, and other key parameters to assess their efficiency in a simulated Ethereum blockchain environment. The obtained data was thoroughly analyzed to determine which algorithm provides better performance in terms of speed and efficiency. Following our comparative analysis, we then embarked on a crucial step: the integration of EdDSA using the Ed25519 signature scheme within the Ethereum environment for transaction signing. We examined the specific requirements of Ethereum's transaction format, including how signatures are handled in transactions to enable the use of EdDSA instead of ECDSA. Our implementation of EdDSA for Ethereum transaction signing provides a compelling proof of concept for the viability of this alternative in a real-world context.

Organization of the Paper

This document is organized as follows: In section "Preliminaries", we discuss the transition to Elliptic Curves and EdDSA, along with a brief comparison between EdDSA and ECDSA. In section "Ethereum Transaction", we present Ethereum transactions and how they utilize signature algorithms. Section "Ed25519 Signature" covers the Ed25519 signature algorithm, detailing its key generation, signing, and signature verification operations. In section "Integration", we propose the integration of EdDSA for signing Ethereum transactions. Section "Security and Performance" focuses on the security of EdDSA, accompanied by a performance comparison. Finally, we conclude this document in section "Conclusion".

Preliminaries

Notation:

- p indicates the prime number defining the underlying field.
- GF_p Finite field with p elements
- EdDSA Edwards-curve Digital Signature Algorithm
- **ENC(N)** The encoding of N in little-endian form as a b-bit string.

Elliptic Curve Cryptography

Over the past decade, we have witnessed a gradual transition in the field of digital signatures. It began with the shift from RSA signatures to DSA signatures, ultimately leading to elliptic curve-based signatures. Over time, these advancements have primarily aimed at optimizing the performance of these techniques. Modern cryptography currently favors the use of ECDSA due to several factors such as reduced key and signature sizes, the level of security it provides, and improved performance [11]. Elliptic curves have largely supplanted previous methods, notably DSA. In this lineage, Edwards curves have emerged as a successor to elliptic curves, bringing notable enhancements to this field.

Harold Edwards, in 2007, delved deeply into the family of elliptic curves and introduced a new variant known as Edwards curves. This innovation laid the foundation for the Edwards-curve Digital Signature Algorithm (EdDSA). The EdDSA offers standardized performance and successfully addresses many security issues that traditional digital signature systems faced.

Table 16.1 Comparison of EdDSA vs ECDSA

Attributes	Edwards curve DSA	Elliptic curve DSA
Curves	$ax^2 + yx^2 = 1 + dx^2y^2$	$y^2 = x^3 + ax + b$
Signature scheme	Schnorr signature scheme	ElGamalsignature scheme
Performance	Faster	Slower
Order	Not prime order	Prime order possible
Key recovery	Not possible	Possible
Curve safety	More	Less
Curve form	General	Subset
Curve arithmetic	Faster addition	Slower
Group law	Complete	Exception

Elliptic Curve and Edwards

EdDSA is a deterministic signature scheme that uses the elliptic curves Ed25519 and Ed448 [12]. Unlike ECDSA, which relies on cyclic groups over the finite field of the curve and the discrete logarithm problem, as well as a variant of the El Gamal signature, EdDSA is based on the Schnorr signature scheme, making it simple, secure, and faster compared to ECDSA.

Comparison of EdDSA vs ECDSA

Basic arithmetic operations, the group law, and the prime order are optimized in EdDSA. EdDSA offers a high level of curve security and high performance, preventing security vulnerabilities. In the event of key loss or theft, recovery is impossible in EdDSA. Table 16.1 provides a comparison between EdDSA and ECDSA, based on sources [12, 13].

Ethereum Transaction

Transactions are interactions among participants within a blockchain. In Ethereum, there are two types of transactions: user transactions and contract transactions. User transactions are used to transfer cryptocurrency (Ether in Ethereum) to a specific address. Contract transactions allow users, such as application developers, to execute predefined functions in smart contracts. These contracts are immutable computer programs deployed on the Ethereum blockchain [14]. This section examines in detail the essential components of an Ethereum transaction and the complex signing process that ensures the security and integrity of operations.

Ethereum Transaction Format

In the standard transaction format (Nonce, Gas Price, Gas Limit, To, Value, Data, v,r,s), only six elements are used, i.e., without v, r, and s, but with the new format, there are nine elements.

If this new format is used, then the v of the signature must be set to {0,1}+ CHAIN_ID * 2 + 35, where {0,1} is the parity of the y-value of the curve point for which r is the x-value in the secp256k1 signature process. If you choose to hash only six values, then v continues to be defined as {0,1} + 27 as before.

This value is included to prevent simple replay attacks [16].

Practical Ethereum Transaction Signature

To create a valid transaction, the sender must perform a digital signature of the message using the elliptic curve signature algorithm. However, this "transaction signature" actually pertains to the Keccak-256 hash of the transaction data serialized according to the RLP scheme. It is important to note that the signature is applied to the hash of the transaction data, not the transaction itself. To sign a transaction in Ethereum, the initiator must:

1. Develop a data structure for a transaction that includes nine elements: nonce, gasPrice, gasLimit, to, value, data, chainID, 0, 0.
2. Generate an RLP-encoded message from the transaction data structure, creating a serialization of the message.
3. Calculate the Keccak-256 hash (SHA3 family) of this serialized message.
4. Sign the hash of the message with its ECDSA private key. The ECDSA [17] signature process, summarized, proceeds as follows:
    - Computation of a random number called nonce (k)
    - Computation of r = k.G, where G is the base point of the secp256k1 curve.
    - Computation of s = k^{-1}(H(M)+r.sk), where H is the SHA-256 hash function and sk is the private key
5. Use "v": The "v" value in Ethereum is added to indicate which of the two public keys (compressed or uncompressed) should be used to verify the signature. The value of "v" depends on the Ethereum network (mainnet, Goerli, Sepolia, etc.) and whether the public key is compressed or uncompressed.
6. Add the values "v," "r," and "s" to the transaction.
7. Broadcast the signed transaction on the Ethereum network. Miners on the network receive the transaction, verify it, and add it to the next block if they deem it valid.

Ed25519 Signature

The signature system is defined on the elliptic curve group:

$$E = \{(x,y) \in F_q * F_q : -x^2 + y^2 = 1 + dx^2 y^2\}$$

where $d = -\frac{121665}{121666} \in F_q$ and $q = 2^{255} - 19$. The neutral element of the group is 0 =(0,1), and the complete addition law of the Edwards curve is

$$(x_1,y_1) + (x_2,y_2) = \left(\frac{x_1 y_2 + x_2 y_1}{1 + d_{x_1 x_2 y_1 y_2}}, \frac{y_1 y_2 + x_2 x_1}{1 - d_{x_1 x_2 y_1 y_2}}\right)$$

The number of points on the elliptic curve is |E|=8*L, where $L = 2^{255} + 27742317777372353535851937790883648493$ is prime. The base point B, as defined in RFC [5], has an order equal to L. The base point was selected as the point with the smallest "u" coordinate in the Montgomery representation (u = 9) [18].

Algorithm 1 Key generation

Require: k(An EdDSA secret key is a string of b bits)

1. $H(k) \leftarrow (h_0, h_1, ..., h_{2b-1})$
2. $a \leftarrow 2^{b-2} + \sum_{3 \leqslant i \leqslant b-3} 2^i h_i$
3. $A \leftarrow a \cdot B$
4. **return** ENC(A)

Algorithm 2 Signature generation

Require: M,$(h_0, h_1, ..., h_{2b-1})$, B and A

1. $a \leftarrow 2^{b-2} + \sum_{3 \leqslant i \leqslant b-3} 2^i h_i$
2. $h \leftarrow H(h_b, ..., h_{2b-1}, M)$
3. $r \leftarrow$ h mod L
4. $R \leftarrow r \cdot B$
5. $h \leftarrow H(R, A, M)$
6. $S \leftarrow (r + ah)$ mod L
7. **return** (R,S) ;

Algorithm 3 Signature verification

Require: M,A

if R ∈ E and S ∈ $\{0, 1, ..., L-1\}$ and $8S \cdot B = 8 \cdot R + 8H(R, A, M) \cdot A \in E$ **then**
  return True ;
else
  return False ;
end if

Integration

Description

In this section, we propose the integration of the EdDSA signature algorithm into the Ethereum transaction signing process for enhanced security and improved performance.

Integration of EdDSA into Ethereum Transaction Signing

ECDSA uses elliptic curves over finite fields to perform signature computations. In ECDSA, signature generation involves selecting a random nonce, followed by complex computations based on hash functions and mathematical operations on the elliptic curve. Additionally, the "v" value is added to the signature to indicate which public key is used for verification. EdDSA simplifies the signature process by using Edwards curves. In EdDSA, signature generation is done deterministically, using a hash function to generate a nonce from the private key and the message. Unlike ECDSA, EdDSA directly generates a single signature value, eliminating the need for the "v" value. This simplified approach makes EdDSA more efficient and less prone to human errors, and most importantly, it significantly speeds up transaction signing because by removing the "v" value, transactions consume less data. In practice, when using EdDSA, the transaction signing process follows the steps outlined in section Practical Ethereum Transaction Signature, with the following replacements for steps 4, 5, and 6:

4 Sign the hash of the message with the EdDSA private key. The EdDSA signature process involves the following steps:

  a Generate a random number called nonce (r).
  b Calculate the Edwards curve point: $R = r \cdot B$, where B is the base point of the Ed25519 curve.
  c Calculate the scalar "s" using $s = r + H(R|| A || M) \cdot sk$, where H is the SHA-512 hash function, A is the public key, and sk is the private key.

5 Do not use "v": Unlike ECDSA, EdDSA does not need the "v" value to indicate which public key to use for signature verification because the EdDSA signing process directly generates a single signature value.
6 Add the "r" and "s" values to the transaction.

Security and Performance

Security Analysis of Ed25519

EdDSA is appreciated for its level of security and resistance against certain well-known attacks, mainly due to its deterministic construction. The primary distinction between deterministic and random signature schemes lies in nonce generation. In random schemes, this relies on the use of a random number generator, while in deterministic schemes the nonce is typically a function of the input message and the private key. Despite its deterministic construction, EdDSA can still be vulnerable to fault attacks.

The first fault injection attack against elliptic curve cryptography was proposed by Biehl et al. [19]. They demonstrated that if the base point B is tampered with during signature operations and if the ECC implementation does not check whether this point resides on the curve or not, then computations will be performed on another curve where the Discrete Logarithm Problem (DLP) can be easily solved using the Pohlig–Hellman algorithm. This attack was later improved upon by Ciet et al. [20]. They showed that the altered base point value can be recovered from the erroneous output. Furthermore, they demonstrated that a fault occurring in system parameters such as field definition or curve coefficient values can lead to the recovery of the secret key.

Malleability We see no relevance of "malleability" to the standard definition of signature security. For instance, consider a slight modification to the system where S would be replaced with -S and A with -A. This would turn one valid signature into another valid signature for the same message, but with a new public key. However, this modification still would not achieve the attacker's goal, which is to forge a signature for a new message under a target public key.

Fault Attack Y. Romailler et al. [21] used the deterministic construction of EdDSA to perform a fault attack based on disrupting step 5 of Algorithm 2 when calculating the signature. They demonstrated that if the output of the hash is altered and changed to a value $h' \neq h$, this will lead to a faulty signature (R,S'), so only the second part of the signature is modified, and the value of a can be recovered as follows:

$$a = (S - S')(h - h')^{-1} \bmod L$$

The values R, A, and M are known, so the value of h can be computed, but the value of h' must be known or guessed. This limitation can be circumvented if the fault model is adequately characterized and if the altered value can be estimated during a post-processing phase. For example, they used the fault model as a random byte $e \in \{1, 2, \ldots, 255\}$ injected at a random offset i after the hash computation. The complete fault verification algorithm for their fault attack model proceeds as follows:

If the algorithm returns ERROR, it means that the injected fault did not correspond to a single random byte error. A solution has been proposed in [21] to counteract this attack. Our implementation is resistant to these types of attacks

Algorithm 4 Ed25519 fault post-processing

Require: M,A,(R,S) and (R,S$'$)
  $h \leftarrow H(R, A, M)$
  $i \leftarrow 0$
  for $i \leftarrow 32$ **do**
    $e \leftarrow 1$
    for $i \leftarrow 256$ **do**
      $h' \leftarrow 2^{8i} e \oplus h$
      $a \leftarrow (S - S')(h - h')^{-1}$ mod L
      if a $\cdot$ B == A **then**
        return a
      end if
      $e \leftarrow e + 1$
    end for
    $i \leftarrow i + 1$
  end for
  return ERROR

because it is based on RFC 8032 [5], in which several implementation methods have been proposed to protect against these kinds of attacks.

Side-Channel Attack Samwel et al. [22] demonstrated the possibility of side-channel attacks on the SHA512 hash function utilized in EdDSA. As a protective measure, they propose introducing randomness into the hash output. Their attack relies on Ed25519 as implemented in WolfSSL on an STM32F4 microcontroller. In the generation of the Ed25519 signature, it is known that the public key is computed as follows: R = r.B. Samwel et al. extracted the ephemeral key r from its scalar multiplication with the base point of the curve.

Performance and Comparison

ECDSA vs. EdDSA Here, we present a comparison of the EdDSA signature algorithm with the ECDSA signature algorithm in Ethereum transaction signing. The implementation is done in Python on an Intel Core i3 processor with 8 GB of RAM and a 64-bit operating system. The comparison is based on the execution time of various algorithms (key generation, signature generation, and signature verification) with a key size of 256 bits.

The execution times are listed in Table 16.2. We measure the speed of our implementation (the number of CPU cycles required to process 1 byte) as

$$\frac{execution\ time(s) * processor\ frequency(Hz)}{size\ of\ the\ test\ file\ (bytes)}.$$

Table 16.2 Execution times for Ed25519 and ECDSA

Schemas	Cycles			Execution time (in seconds)		
	Key generation	Signature generation	Signature verification	Key generation	Signature generation	Signature verification
Ed25519	18732	15838	46455	0.05	0.04	0.13
ECDSA	190493	182239	164711	0.54	0.51	0.46

Discussion As we know, EdDSA (Ed25519) and ECDSA have a similar key size of 256 bits and offer the same level of security at 128 bits.

With our implementation, we can observe that the key generation, signature generation, and signature verification processes are much faster in Ed25519 than in ECDSA. This is highly advantageous in a distributed environment like the blockchain.

Conclusion

As a first contribution, we proposed the use of EdDSA in Ethereum transaction signatures instead of ECDSA. Furthermore, we implemented both signature algorithms for signing Ethereum transactions in a simulated environment. As shown in Table 16.2, EdDSA is significantly faster than ECDSA.

References

1. W Mehuron. Digital signature standard (DSS). US department of commerce, national institute of standards and technology (NIST). *Information Technology Laboratory (ITL). FIPS PEB*, 186, 1994.
2. Diego F Aranha, Felipe Rodrigues Novaes, Akira Takahashi, Mehdi Tibouchi, and Yuval Yarom. LadderLeak: Breaking ECDSA with less than one bit of nonce leakage. In *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*, pages 225–242, 2020.
3. Marcan Bushing. Segher, and sven, PS3 epic fail. In *27th Chaos Communication Congress*, 2010.
4. Phong Q Nguyen and Igor E Shparlinski. The insecurity of the elliptic curve digital signature algorithm with partially known nonces. *Designs, codes and cryptography*, 30:201–217, 2003.
5. Simon Josefsson and Ilari Liusvaara. Edwards-curve digital signature algorithm (EdDSA). Technical report, 2017.
6. Daniel J Bernstein, Niels Duif, Tanja Lange, Peter Schwabe, and Bo-Yin Yang. High-speed high-security signatures. *Journal of cryptographic engineering*, 2(2):77–89, 2012.
7. George Barwood. Digital signatures using elliptic curves, message 32f519ad. 19609226@ news. dial. pipex. com posted to sci. crypt (1997). *URL: http://groups.google.com/group/sci.crypt/msg/b28aba37180dd6c6.Citations in this document*, 2.
8. J Guruprakash and Srinivas Koppu. An empirical study to demonstrate that EdDSA can be used as a performance improvement alternative to ECDSA in blockchain and IoT. *Informatica*, 46(2), 2022.

9. Shaik Johny Basha, Venkata Srinivasu Veesam, Tamminina Ammannamma, Sirisha Navudu, and MVVS Subrahmanyam. Security enhancement of digital signatures for blockchain using EdDSA algorithm. In *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, pages 274–278. IEEE, 2021.
10. Alessandro Barenghi and Gerardo Pelosi. A note on fault attacks against deterministic signature schemes (short paper). In *Advances in Information and Computer Security: 11th International Workshop on Security, IWSEC 2016, Tokyo, Japan, September 12–14, 2016, Proceedings 11*, pages 182–192. Springer, 2016.
11. Thomas Pornin. Deterministic usage of the digital signature algorithm (DSA) and elliptic curve digital signature algorithm (ECDSA). Technical report, 2013.
12. Benjamin Black, Joppe W Bos, Craig Costello, Patrick Longa, and Michael Naehrig. Elliptic curve cryptography (ECC) nothing up my sleeve (NUMS) curves and curve generation (2014).
13. Joppe W Bos, Craig Costello, Patrick Longa, and Michael Naehrig. Selecting elliptic curves for cryptography: an efficiency and security analysis. *Journal of Cryptographic Engineering*, 6:259–286, 2016.
14. Michael Pacheco, Gustavo A Oliva, Gopi Krishnan Rajbahadur, and Ahmed E Hassan. What makes Ethereum blockchain transactions be processed fast or slow? An empirical study. *Empirical Software Engineering*, 28(2):39, 2023.
15. *Mastering Ethereum: Building Smart Contracts and DApps*, chapter Transactions. O'Reilly Media, 2018.
16. Vitalik Buterin (@vbuterin). "EIP-155: Simple replay attack protection," Ethereum improvement proposals, no. 155, October 2016. Technical report, https://eips.ethereum.org, 2016.
17. Don Johnson, Alfred Menezes, and Scott Vanstone. The elliptic curve digital signature algorithm (ECDSA). *International journal of information security*, 1:36–63, 2001.
18. Adam Langley, Mike Hamburg, and Sean Turner. RFC 7748: Elliptic curves for security. *Internet Research Task Force (IRTF)*, 2016.
19. Ingrid Biehl, Bernd Meyer, and Volker Müller. Differential fault attacks on elliptic curve cryptosystems. In *Advances in Cryptology—CRYPTO 2000: 20th Annual International Cryptology Conference Santa Barbara, California, USA, August 20–24, 2000 Proceedings 20*, pages 131–146. Springer, 2000.
20. Mathieu Ciet and Marc Joye. Elliptic curve cryptosystems in the presence of permanent and transient faults. *Designs, codes and cryptography*, 36(1):33–43, 2005.
21. Yolan Romailler and Sylvain Pelissier. Practical fault attack against the ed25519 and EdDSA signature schemes. In *2017 Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC)*, pages 17–24. IEEE, 2017.
22. Niels Samwel, Lejla Batina, Guido Bertoni, Joan Daemen, and Ruggero Susella. Breaking ed25519 in WolfSSL. In *Topics in Cryptology–CT-RSA 2018: The Cryptographers' Track at the RSA Conference 2018, San Francisco, CA, USA, April 16–20, 2018, Proceedings*, pages 1–20. Springer, 2018.
23. Damian Poddebniak, Juraj Somorovsky, Sebastian Schinzel, Manfred Lochter, and Paul Rösler. Attacking deterministic signature schemes using fault attacks. In *2018 IEEE European Symposium on Security and Privacy (EuroS&P)*, pages 338–352. IEEE, 2018.

Chapter 17
Vulnerability Prediction of Web Applications from Source Code Based on Machine Learning and Deep Learning: Where Are At?

Mawulikplimi Florent Gnadjro and Samba Diaw

Abstract With the rise of new information technologies around the world, many distributed applications and web applications have emerged, so it is important to make them secure. Despite the emphasis placed by software security experts on the need to build secure web applications, the number of new vulnerabilities found in web applications is growing. Machine Learning (ML) and Deep Learning (DL) through their vulnerability prediction approach are increasingly being offered for source code analysis, providing a powerful way to make web applications less vulnerable. Many ML- and DL-based approaches have been proposed to automatically detect, locate, and repair software vulnerabilities. Although ML-based are more effective than vulnerability analysis tools based on static source code analysis by security experts, accurately identifying types of vulnerabilities and estimations severity remains challenging. The graphical representation of source code, the best vulnerability differentiation, and the support of a large corpus of vulnerabilities are not at least. This thesis aims to study the prediction of vulnerabilities in web applications from source codes using ML and DL techniques. A comprehensive review of the literature on the different approaches proposed for the prediction of vulnerabilities of web applications will allow us to identify the current state of research and challenges in this field, thus positioning us well to make a significant contribution in the prediction of vulnerabilities of web applications using the techniques of ML and DL.

Keywords Machine learning · Deep learning · Vulnerabilities · Security · Source code · Vulnerability prediction

M. F. Gnadjro (✉) · S. Diaw
Medical Imaging and Bioinformatics Laboratory, Higher Polytechnic School, Cheikh Anta DIOP University, Dakar, Senegal

C. T. Gueye et al. (eds.), *Mathematics of Computer Science, Cybersecurity and Artificial Intelligence*, Springer Proceedings in Mathematics & Statistics 468,
https://doi.org/10.1007/978-3-031-66222-5_17

Introduction

The growth of web applications has been significant over the last decade. According to figure provided by Kepios, the number of Internet users increased from 2.177 billion in January 2012 to 4.95 billion in January 2022, an average annual growth of 8.6% over the past decade as a whole. These figures show that the use of web applications is widespread and continues to grow. By most estimates, more than three quarters of cybercrime attacks target applications and their vulnerabilities. However, this growth has also led to an increase in the number and complexity of software vulnerabilities. Software vulnerabilities can be exploited by attackers to access confidential information and cause significant damage. These flaws are often due to programming or design errors, but can also result from insufficient security practices.

Faced with this challenge, it is therefore crucial to integrate automatic detection and correction of vulnerabilities into the web applications development process. Nowadays, instead of using static source code analysis tools, researchers and software engineers use machine learning (ML) and deep learning (DL) for predicting vulnerabilities in source code. Examples: vulnerabilities detection in source code using machine learning [2] and source code representation [3]. The increased availability of data (source code) is driving growing interest in deep learning. Researchers are particularly interested in deep learning because it promises good results without requiring an extraction process as required by traditional machine learning models [5]. By leveraging machine learning capabilities, developers and security experts can identify and correct potential vulnerabilities in the early stages of development, which can significantly reduce the time and costs associated with correcting security issues at a later stage.

The remainder of this chapter is structured as follows. First, we present the background and issues of this work. In section "Purposes", we present the purpose, while section "Related Works" deals with the related works on vulnerability prediction of web applications from source code using ML and DL. In section "Research Direction", we discuss the research direction. Finally, we conclude this work by recalling the problem and objectives of this work while indicating our research direction.

Background and Issues

Developed to meet a multitude of our needs, web applications are now essential tools in our daily lives. Web applications have grown considerably in recent years, just like technologies in general. The complexity of the technologies used today to create web applications (Java, C, C++, Groovy, JavaScript, PHP, Ruby, J2E, etc.) makes it particularly difficult to prevent the introduction of vulnerabilities in these applications and to estimate or predict their presence. In addition, network security and firewall installation do not provide adequate protection against attacks because these applications are by definition public and accessible to everyone, and most vulnerabilities are due to programming errors in the application itself [6].

According to statistics from Common Vulnerabilities and Exposures (CV) [7, 8], in 2020 and the first six months of 2021, the world saw a record number of exploited software security vulnerabilities. Thanks to these statistics, we can see the threats faced by web applications. Web application vulnerabilities can be severe, ranging from loss of information or disclosure of secret information to manipulation and system failure. All can have a serious impact on businesses, governments, society, and individuals. An exploit such as ransomware, for example, has shut down hospitals, telecommunications services, and transportation systems and caused damage in the hundreds of millions of dollars [9].

Many vulnerabilities are caused by subtle code flaws, covering a few or even a single line of code [10] in web applications. Most security problems in web applications are related to the exploitation of source code vulnerabilities. Manual detection of these vulnerabilities is very difficult and costly in terms of time and budget. Consequently, the use of tools that help developers and decision-makers to automatically predict vulnerable components is necessary, because it minimizes the search effort and helps to minimize the costs involved in remediation of vulnerabilities. This is how source code flaws have been discovered using static and dynamic analysis techniques. However, static analysis techniques are known to generate a high number of false positive results, whereas dynamic analysis tools are designed to underestimate the number of flaws in a program and are therefore prone to false negatives. In addition, tactic analysis techniques can require significant computing resources, while dynamic analysis tools increase the size and execution time of a program.

As a result, these techniques cannot yet be seamlessly integrated into continuous delivery pipelines of today's web applications. In this respect, machine learning (ML) techniques seem to have become a very attractive alternative to traditional software defect detection and correction techniques. Given the growing interest in ML applications in source code, several studies have begun to apply ML for bug prediction.

Purposes

The general goal of our study is the prediction of vulnerabilities in web applications based on machine learning and deep learning. To achieve this, we propose an innovative approach the takes into account the following criteria:

Accuracy in Vulnerability Detection Static analysis tools can only detect generic errors using a list of static rules and predefined vulnerability models. They do not allow for in-depth precision of vulnerabilities and their causes.

Granularity of Detection Vulnerability within source code must be identified at a finer granularity to locate the vulnerability and facilitate correction.

Graphical Representation of Code Source A method of detecting vulnerabilities automatically and intelligently, using a tools operation at source code level to provide an intermediate graphical representation of source code and a graphical neural-network-based model for vulnerability prediction. A major challenge is the representation of code source in a form that effectively captures the syntax and semantics of the source code, given that each developer has his or her own style of code writing, choice of variable names, etc... In addition, different programming languages are used by developers.

A Vulnerability Classification Model Using Deep Learning Exploring possible ways of leveraging learning techniques to apply a model trained on a certain language to other languages.

Related Works

Vulnerability Datasets

Data plays a crucial role in the creation and evaluation of software vulnerability detection models based on machine learning (ML) and deep learning (DL) [1, 19–21]. Vulnerability detection models rely on datasets to learn how to identify security flaws. Over the last decade, various vulnerability-related test suites and datasets have emerged, each with specific objectives. These sets were then adopted by other security researchers looking for relevant data to train or evaluate their techniques. The quality of datasets can be assessed by various factors such as data source, data size and scale, data types, and preprocessing steps performed on the data [4]. In this section, we review the data used in vulnerability detection studies and proceed to an in-depth analysis of the data processing steps, as well as an origin of the data, data types, and in the following section their representation. Consequently, there is a gap in research on how to obtain sufficient datasets to facilitate the training of ML/DL models for source code vulnerability detection. Our preliminary analysis dealing with vulnerability datasets reveals that these can be classified into three categories, namely reference sources, collected sources, and hybrid sources. These include open-source projects, public vulnerability databases, and bug repositories [19]. The limitations identified in the existing datasets are related to obsolete library dependencies: Some vulnerabilities in the dataset from [19] are no longer reproducible due to obsolete library dependencies. This may affect the relevance of some vulnerabilities for current research: data size; although the dataset contains information, it does not necessarily cover all existing vulnerabilities. This may limit the graphical representation for certain categories of vulnerabilities. One of the challenges of detecting and correcting vulnerabilities in source code is the insufficient amount of data available for training operations. Our aim is to aggregate a diverse set of programs or code extracts covering different application domains, programming languages, and vulnerability types. From the data collected, we will carefully select the programs to be included in our reference dataset. This selection process will take into account elements such as program complexity, vulnerability diversity, and code quality. Our aim will be to create a comprehensive set that addresses a wide range of maintainable vulnerabilities and reflects real-life situations.

Source Code Representation

Source code representation plays an important role in vulnerability prediction using ML/DL. In order to carry out our research, we investigated the following two questions:

Q1: What studies address source code representation models?
Q2: Which model offers the best source code representation for machine learning?

The more features we capture from the data, the better the modeling we can perform. Related studies such as [3, 14–17] inform us that methods such as fastText BERT, word2vec are effective for source code representation. Our results show that all code presentation methods are suitable for representation, but the BERT for code model is more promising because it takes less time to implement.

According to the studies carried out at article level [3], the Long Short-Term Memory (LSTM) model achieved an exceptional overall accuracy of 93.8% in predicting Python source code vulnerabilities. Analyzing the different types of vulnerabilities, they noted slight variations in model performance depending on source code integration. However, the prediction model based on the word2vec representation of the code clearly outperformed the models using fastText and BERT to detect SQL injections. In the case of command injection, cross-site request forgery (XSRF), remote code execution, and graph disclosure, BERT-based models outperformed models using the other two encoding methods.

The challenges identified by these articles include programming complexity, code representation, improving semantic representations, and the use of different classifiers. Thus, in our work we will study the representation of source code in a form that effectively captures the syntax and semantics of source code, given that each developer has his or her own style of writing code, choosing variable names...

Machine Learning Detection

Article [18] presents an empirical study on the use of machine learning (ML) and statistical techniques to predict software vulnerabilities. The authors evaluate the performance of these techniques using data from ten open-source software projects and according to various criteria such as accuracy. The study evaluates nine machine learning methods (feed-forward neural network with backpropagation (FFBPNN), cascading feed-forward neural network with backpropagation (CFBPNN), adaptive neuro fuzzy inference system (ANFIS), multilayer perceptron (MLP), support vector machine (SVM), bagging, M5Rrule, and M5P. Reduced error pruning and three statistical techniques (Alhazmi–Malaiya Model, Linear Regression, Logistic Regression Model). The authors conclude that machine learning techniques show a remarkable improvement in software vulnerability prediction over statistical vulnerability prediction models. The applicability of the techniques is examined in two approaches, namely suitability and predictive capability. However, the authors highlight some limitations in the application of machine learning techniques, namely insufficient training data: The article points out that access to large and diverse datasets is often limited, which can lead to bias and generalization errors; model interpretability: Some models such as deep neural networks lack interpretability. It is therefore essential to understand how models make their decisions in order to apply them effectively in real-life scenarios.

The referenced article [13] reviews papers that have studied machine learning approaches to detecting and correcting vulnerabilities in source code. Their studies have raised a number of challenges in the detection and correction of vulnerabilities in source code, namely the definition of a repository: It is essential to establish a set of benchmarks for evaluating the performance of models; language coverage and error types: Models must be able to detect vulnerabilities in different programming languages and for different error types, nonlinear code structure. With this in mind, our research will focus on the definition of a standard repository and a study of language coverage.

Deep Learning Detection

We are now presenting some articles on vulnerability prediction using deep learning. Previous peer-reviewed investigations have been carried out on similar topics, and we will discuss them in the following lines. Indeed, we carried out a systematic survey to see the various works carried out by researchers over the last seven (07) years in vulnerability prediction in source code with an ML/DL-based approach in order to better understand the challenges. In [11], the authors studied the use of heterogeneous graphs to represent both the syntactic and semantic structure of source code. They use neural networks on heterogeneous graphs to learn to reason about these complex structures. These models propagate information on the nodes and edges of the graph. The article evaluates this approach on two tasks: Code comment generation: predicting relevant comments for code snippets; method naming: finding appropriate names for methods in code. Their work has shown that using heterogeneous graphs to represent the structure of source code, and exploiting this information to improve the semantic understanding of programs, is a very effective way to improve the quality of code.

Ref. [12] carried out their work on vulnerability detection through source code, also proposing a tool called JCPG (Java Code Property Graph). This tool allows source code to be represented in the form of graphs fed by a pretrained GNN (Graph Neural Network) model to perform classification tasks. Experimental results show that the model outperforms the static analyzers and GNN models previously used for dataset. So the combination of a tool that operates at source code level to generate an intermediate graphical representation with a GNN model can be used as a vulnerability prediction tool.

To deepen their study, future work could focus on improving graph representations (explore other types of graphs such as data dependency graphs to capture more information about code structure, investigate methods for integrating vulnerability-specific contextual information into graph representations), learning transfer (study how to transfer knowledge learned from one vulnerability dataset to other programming languages or vulnerability types), a diversified dataset (create larger and more diversified vulnerability datasets, covering different vulnerability categories and programming languages, explore data from open-source and proprietary projects), and assessing the robustness of models in the face of adverse attacks. Our study covers the analysis of the source code but above all of the vulnerabilities written by a developer rather than focusing on the malicious code injected by a hacker or hacker and also deals with the points related to precision in the prediction of vulnerability in the source code.

Discussion

Studies diverge according to the specific types of faults they target. This diversity leads to the creation of a variety of fault models, with implications for representation choices. These factors seem to determine whether a tool can automatically correct bugs or simply detect them. In general, syntactic errors are easier to spot, while vulnerabilities pose more complex challenges. Given this complexity, most studies focus more on defects than on vulnerabilities. This hypothesis is justified by examining studies on the datasets used by researchers. Synthetic datasets are mainly used by vulnerability detection tools. It should be noted that vulnerability prediction tools use real projects and manage to identify vulnerabilities, but this cannot be done effectively on a large scale and without a large number of misclassifications. We find that defect correction tools can achieve their intended purpose using simpler representations, while defect detection tools use more advanced and combined representations. Defect prediction tools use more advanced and combined representations. This also shows that tackling vulnerabilities and semantic defects is probably more difficult, so large-scale automatic correction is not yet possible. ML/DL approaches have been widely explored by researchers, but they also have their limitations:

Lack of Training Data For ML/DL models to be effective, high-quality training data is required. It should be noted that vulnerabilities are often rare and difficult to find in datasets.

Code Characteristics Extracting meaningful characteristics from source code is a challenge. Traditional methods focus on basic metrics such as code length, but this may not be enough.

Class Imbalance Vulnerabilities are often in the minority compared to healthy code. This can lead to class imbalance and affect model performance.

Generalization Models may overlearn the specifics of training datasets and fail to generalize correctly to new codes.

Although ML/DL approaches have shown promising results, it is essential to combine them with other methods and consider their limitations for accurate vulnerability prediction in source code. This chapter is motivated by the need to discover models in the rapidly evolving field of ML for source code. Among the challenges to finding effective solutions are high-quality datasets of real, representative, and correctly labeled data. Effective source code representation is capable of semantic understanding in terms of goals and relationships.

Solution Architecture

Recent years have seen the emergence of ML for vulnerability scanning. A typical ML pipeline comprises several important stages: data collection and preparation, model training, and finally evaluation and deployment. As a rule, the tool is monitored, maintained, and improved after deployment. For implementation, we will follow the steps below:

Data Collection Collect a set of source code data containing code examples with and without vulnerabilities. The data will come from code repositories, vulnerability databases, or other sources.

Data Preprocessing Clean and normalize data. Extract relevant features from the code, such as tokens, n-grams, embeddings, etc.

Model Selection Select an appropriate ML/DL model for the vulnerability prediction task. Examples of models: neural networks, convolution networks, recurrent networks, etc. An in-depth study will be carried out in this respect.

Model Training Divide data into training and test sets. Train the model on the training set using backpropagation and optimization techniques.

Model Evaluation Evaluate model performance on the test set. Use metrics such as precision, recall, F-measurement, etc.

Model Optimization Adjust model hyperparameters to improve performance. Use techniques such as grid search, cross-validation, etc.

Model Deployment Integrate the model into a real-time vulnerability detection system. Monitor model performance in production. Integrate machine learning into development tools. Analysis results can be used to enhance existing vulnerability detection tools ; integrating machine learning models into IDEs (integrated development environments) could help developers detect and correct vulnerabilities more effectively.

Continuous Updating Regularly update the model with new data to maintain accuracy. Retrain the model as required.

Research Direction

This research aims to explore new approaches for efficiently detecting vulnerabilities in source code. We propose to use heterogeneous graphs to represent both the syntactic and semantic structure of the code, taking into account the different relationships between elements. We will then apply deep learning techniques to learn how to reason about these complex structures. Our aim is to improve the accuracy of vulnerability detection, while taking into account the specific challenges associated with these domains. Although datasets contain information, they do not necessarily cover all existing vulnerabilities. This may limit the graphical representation for certain categories of vulnerability. To improve the relevance of the dataset, it is essential to broaden the analysis corpus. Apart from supporting vulnerability prediction and automated remediation, research work can be directed into other application areas, such as vulnerability classification or analysis of vulnerability-related code changes. The results of the analysis can be used to enhance existing vulnerability detection tools by integrating machine learning models into IDEs (integrated development environments) to help developers detect and correct vulnerabilities more effectively.

Conclusion

Web applications are vulnerable to a wide range of security threats. Other than system vulnerabilities, most vulnerabilities are due to application bugs. For this reason, many techniques have been proposed in the literature to detect and prevent attacks on web applications due to poor security practices. The aim of this chapter is to position ourselves in according with the subject of prediction vulnerabilities in web application source code based on machine learning and deep learning.

To clarify our topic, we first conducted a search and selected a few journals, surveys, and conference papers based on vulnerability prediction in the source code of web applications based on machine learning and deep learning over the last seven (07) years. As a research direction, our next paper will focus on data collection. These will come from code repositories, vulnerability databases, or other sources. The aim will be to provide quality data on which to base the application of the representation model.

References

1. Raducu, Razvan, Gonzalo Esteban, Francisco J. Rodríguez Lera, and Camino Fernández. 2020. "Collecting Vulnerable Source Code from Open-Source Repositories for Dataset Generation" Applied Sciences 10, no. 4: 1270. https://doi.org/10.3390/app10041270
2. T. Marjanov, I. Pashchenko, and F. Massacci, "Machine Learning for Source Code Vulnerability Detection: What Works and What Isn't There Yet," IEEE Secur Priv, vol. 20, no. 5, pp. 60–76, 2022, doi: https://doi.org/10.1109/MSEC.2022.3176058.
3. A. Bagheri and P. Hegedűs, "A Comparison of Different Source Code Representation Methods for Vulnerability Prediction in Python," Aug. 2021, Accessed: Feb. 17, 2024. [Online]. Available: https://doi.org/10.5281/zenodo.4703996
4. Z. Bilgin, M. A. Ersoy, E. U. Soykan, E. Tomur, P. Comak, and L. Karacay, "Vulnerability Prediction from Source Code Using Machine Learning," IEEE Access, vol. 8, pp. 150672–150684, 2020, doi: https://doi.org/10.1109/ACCESS.2020.3016774.
5. Sarker, I.H. Deep Cybersecurity: A Comprehensive Overview from Neural Network and Deep Learning Perspective. SN COMPUT. SCI. 2, 154 (2021). https://doi.org/10.1007/s42979-021-00535-6
6. https://www.vaadata.com/blog/fr/comment-renforcer-la-securite-de-vos-applications-web-pour-contrerles-attaques-les-plus-courantes/
7. The State of Open-Source Vulnerabilities 2021, https://www.mend.io/wp-content/media/2021/03/Thestate-of-open-source-vulnerabilities-2021-annual-report.pdf, (25 December 2021)
8. Cho, D. X., Son, V. N., & Duc, D. (2022). Automatically Detect Software Security Vulnerabilities Based on Natural Language Processing Techniques and Machine Learning Algorithms. Journal of ICT Research and Applications, 16(1), 70–87. https://doi.org/10.5614/itbj.ict.res.appl.2022.16.1.5.
9. Dan Goodin, An NSA-derived ransomware worm is shutting down computers worldwide (2017) [cited 21.12.2020]. URL https://arstechnica.com/information-technology/2017/05/an-nsa-derived-ransomware-worm-is-shutting-down-computers-world
10. F. Yamaguchi, M. Lottmann, K. Rieck, Generalized Vulnerability extrapolation using abstract syntax trees, in: Proceedings of the 28th Annual Computer Security Applications Conference, 2012, pp. 359–368
11. K. Zhang, W. Wang, H. Zhang, G. Li and Z. Jin, "Learning to Represent Programs with Heterogeneous Graphs," 2022 IEEE/ACM 30th International Conference on Program Comprehension (ICPC), Pittsburgh, PA, USA, 2022, pp. 378–389, doi: https://doi.org/10.1145/3524610.3527905.
12. S. Singh, "CyberSecurity (CybSec) Automated Vulnerability Detection in Java Source Code using J-CPG and Graph Neural Network," 2021.
13. T. Marjanov, I. Pashchenko, and F. Massacci, "Machine Learning for Source Code Vulnerability Detection: What Works and What Isn't There Yet," IEEE Secur Priv, vol. 20, n o5, p. 60–76, 2022, doi: https://doi.org/10.1109/MSEC.2022.3176058
14. R. Jenni, "Better Code Representation for Machine Learning," 2022.
15. Y. Wainakh, M. Rauf, and M. Pradel, "EVALUATING SEMANTIC REPRESENTATIONS OF SOURCE CODE", Accessed: Sep. 29, 2023. [Online]. Available: https://github.com/sola-st/IdBench

16. R. Halepmollası, K. Hanifi, R. F. Fouladi, and A. Tosun, “A Comparison of Source Code Representation Methods to Predict Vulnerability Inducing Code Changes,” in International Conference on Evaluation of Novel Approaches to Software Engineering, ENASE - Proceedings, Science and Technology Publications, Lda, 2023, pp. 469–478. doi: https://doi.org/10.5220/0011859300003464.
17. R. L. Alaoui and E. H. Nfaoui, “Deep Learning for Vulnerability and Attack Detection on Web Applications: A Systematic Literature Review,” Future Internet, vol. 14, n o4. MDPI, 1 avril 2022. doi: https://doi.org/10.3390/fi14040118.
18. Jabeen, G., Rahim, S., Afzal, W. et al. Machine learning techniques for software vulnerability prediction: a comparative study. Appl Intell 52, 17614–17635 (2022). https://doi.org/10.1007/s10489-022-03350-5
19. G. Bhandari, A. Naseer, and L. Moonen, “CVEfixes: Automated collection of vulnerabilities and their fixes from open-source software,” in PROMISE 2021 - Proceedings of the 17th International Conference on Predictive Models and Data Analytics in Software Engineering, co-located with ESEC/FSE 2021, Association for Computing Machinery, Inc., Aug. 2021, pp. 30–39. doi: https://doi.org/10.1145/3475960.3475985.
20. D. Grahn and J. Zhang, “An Analysis of C/C++ Datasets for Machine Learning-Assisted Software Vulnerability Detection,” 2021.
21. Quang-Cuong Bui, Riccardo Scandariato, and Nicolás E. Díaz Ferreyra, “Vul4J,” 2022.

Chapter 18
Business Process Management and Process Mining on the Large: Overview, Challenges, and Research Directions

Mouhameth Fadal M. Aidara, Samba Diaw, and Mamadou Lakhassane Cisse

Abstract This review addresses the integration of Business Process Management (BPM) and Process Mining (PM) in the context of Industry 4.0's digital transformation. It highlights how BPM enhances business process efficiency, reducing costs and boosting profit, while PM, leveraging data from event logs, offers insights into actual process execution, bridging data and process science. Despite their complementary nature, the extent of BPM and PM integration remains underexplored. This review synthesizes existing literature and identifies future research directions, aiming to inform researchers and practitioners in these evolving fields.

Keywords Business process management · Process mining · Business processes

Introduction

The advent of Industry 4.0, marked by a significant digital transformation, has reshaped the landscape of business processes toward unprecedented efficiency [1], necessitating a deeper understanding of Business Process Management (BPM) and Process Mining (PM). BPM, a discipline focused on the discovery, modeling, analysis, improvement, and optimization of business processes, plays a crucial role in reducing production costs and enhancing organizational profit [2, 3]. Complementing BPM, PM emerges as a field concentrating on the discovery, monitoring, and improvement of real processes through the extraction of knowledge from event logs, bridging data science and process science [4]. This review aims to

All authors have equally contributed to this work

M. F. M. Aidara (✉) · S. Diaw · M. L. Cisse
UMMISCO Laboratory, Cheikh Anta Diop University of Dakar, Dakar, Senegal
e-mail: mouhamethfadalmaraby.aidara@ucad.edu.sn; samba.diaw@ucad.edu.sn; lakhassane.cisse@ucad.edu.sn

C. T. Gueye et al. (eds.), *Mathematics of Computer Science, Cybersecurity and Artificial Intelligence*, Springer Proceedings in Mathematics & Statistics 468,
https://doi.org/10.1007/978-3-031-66222-5_18

provide a comprehensive analysis of BPM and PM, exploring their integration and intersection and highlighting potential research directions.

Motivations for this research are grounded in the rapid advancement of Industry 4.0 and the digitalization of business operations [5], emphasizing the significance of BPM and PM in this context. The review intends to explore the synergies between these disciplines, aiming to enhance the digital transformation journey of businesses. Key challenges and opportunities within BPM and PM, including their role in addressing Industry 4.0 challenges, enhancing productivity, and their evolving relationship with new technologies, are discussed.

This review aims to conduct an overview on BPM and PM, covering their definition, scope, and methodologies. The research also aims to explore the integration of BPM and PM with AI and Machine Learning techniques, addressing key challenges such as event data extraction and cleaning, business process model automatic (re)design, augmenting Process Mining with domain expertise, and common sense.

The organization of the review is structured in the following manner: section "" provides an exploration of Business Process Management (BPM), covering its definition, fundamental concepts, and challenges. Section "Comprehensive Overview of Process Mining: Definitions, Scope, and Fundamental Concepts" shifts the focus to Process Mining (PM), where it examines its definition, fundamental concepts, and challenges faced in this domain. Section "Research Directions" is dedicated to discussing the research directions that have been identified. Conclusively, section "Conclusion" presents a comprehensive summary of the key findings and delineates the implications for future research endeavors.

Business Process Management (BPM): Definition, Evolution, Scope, Challenges

Definition

The concept of BPM, well known in the process experts community, has known a plethora of definitions in the literature. All those definitions essentially converge toward the same intellection of BPM as a management science of business processes. Simply put, Business Process Management is a best practice management principle embraced and applied by organizations with the aim to uphold their competitive advantage [6]. It encloses a set of techniques, methods, and tools to efficiently govern operational business processes with the ultimate goal of maximizing their performance [2]. As mentioned, BPM basically operates on business processes. In a nutshell, a business process is a series of tasks performed within an organization for the purpose of creating an output [7].

Exploring the Dynamics of Business Processes: Taxonomy, Components, and Essentials

Operational Business Processes Taxonomy The scope of Business Process Management (BPM), as defined in section "Definition", focuses primarily on operational business processes, excluding strategic organizational processes and implicit processes [8]. Weske places business goals and strategies at BPM's highest abstraction layer [9]. Operational business processes are concrete sequences of activities aimed at achieving organizational goals and delivering client-specific outcomes [10]. These contrast with strategic processes, which guide an organization's future direction (ABPMP, 2013). Zur Muehlen et al. categorize operational processes into *core (or identity) processes*, which generate the organization's primary products or services, and *Support processes*, which internally aid *core processes'* execution [11].

Business Process Components A *business process* is a structured combination of events, activities, and decision points aimed at achieving specific outcomes. *Events* are defined as instantaneous occurrences without duration, marking significant points within a process. *Activities* constitute concrete work performed in a process and can be further divided into simpler units termed *tasks*. *Decision points* within a process represent critical moments where choices are made, influencing the process's direction [2, 9].

Key elements of a business process also encompass *actors*, who can be internal (such as employees within the organization) or external (such as clients or partners). *Physical objects* and *informational objects* also play roles in the process, the former being tangible items involved in the process and the latter referring to data or documents used or generated. The execution of a business process leads to *outcomes*, which are of value to at least one customer, defining the process's success or failure [2].

With reference to the preceding and simply put, we define a *business process as a series of organized tasks, conducted by actors, influenced by events and decisions, culminating in outcomes beneficial to customers*.

Business Process Management (BPM) is then redefined here as an overarching framework that employs various methodologies, techniques, and tools for enhancing the efficiency and effectiveness of these business processes. The focus of BPM is on identifying, analyzing, redesigning, executing, and monitoring business processes to optimize performance [2, 9].

This approach underscores the importance of business processes as central elements within BPM.

Essential Business Processes Corporations rely on fundamental business processes for value creation in product or service delivery. Key processes include:

– **Order-to-Cash (O2C):** This process begins with a customer order and ends with product/service delivery and payment collection [2].

– **Quote-to-Order (Q2O):** Starting with a client's quote request and concluding with order placement, this process often precedes O2C [2].
– **Issue-to-Resolution (I2R):** Initiated by a user issue with a product or service, it concludes with a mutually agreed solution between client and supplier [2].
– **Application-to-Approval (A2A):** Relevant in contexts like grant or program applications, this process spans from application submission to its approval or rejection [2].

BPM Life Cycle Overview Business Process Management (BPM) focuses on continually and incrementally improving business processes to enhance operational business processes efficiency [3]. A critical aspect is the *BPM life cycle*, which is a series of phases forming a loop for ongoing process improvement [12]. This concept, extensively detailed by Van der Aalst [3] and others such as Netjes [8], Hallerbach [12], Houy [13], Weske [9], and Zur Muehlen [11], provides a framework for process analysis, design, management, and improvement. The BPM lifecycle model by Van der Aalst [3], which is the focus of this review, aligns closely with Workflow Management (WFM) and is structured into four stages: *Design, Configuration, Enactment, and Diagnosis*. These stages cover respectively the (re)design of processes, their implementation in a Workflow Management System (WFMS), execution and stakeholder participation, and finally, using execution data for process improvement and issue identification [3]. Van der Aalst highlights the integral role of IT and process automation in BPM life cycle [3].

Critical Challenges in Business Process Management to Address

In advancing from the foundational understanding of Business Process Management (BPM), it is imperative to address the multifaceted challenges it encounters, particularly in the context of digital innovation and the evolving business landscape:

– **BPM-Driven Value Creation from Data:** The surge in data and technological advancements significantly alters business operations, demanding profound changes in BPM. Organizations face socio-technical barriers in harnessing data for value creation, posing a challenge in integrating data-driven approaches both technically and culturally [14].
– **Expansive BPM:** Despite significant investments in BPM, many organizations experience fragmented processes, lacking a holistic view [14]. This limitation became pronounced during the COVID-19 pandemic, as organizations struggled with uncoordinated process changes [15]. Addressing "big processes," those extending beyond organizational boundaries and intertwined with various disciplines, remains a critical challenge [14].
– **Automated Process (Re-)Design:** The trend toward "hyperautomation" has heightened the expectation of automating process operations [16]. However, process (re-)design largely remains manual, cognitively demanding, and susceptible

to errors. The creativity required for (re-)design and the complex socio-technical nature of processes pose significant hurdles to automation [14].

- **Constructing Digital Twins:** Business processes frequently undergo changes, both organic and planned, such as task reordering or automation [17]. These interventions impact various performance measures, creating a challenge in constructing digital twins that accurately reflect these dynamic changes [14].
- **Lack of Objectivity in Process Descriptions:** Process models are central to BPM [18] but often lack objectivity in terms of terminology, perspectives, and granularity [14]. This issue extends to manual model creation and process discovery algorithms, impacting the effective utilization of models [19].
- **Augmenting Process Mining with Common Sense and Domain Knowledge:** Event logs of medium quality present challenges in process mining, requiring the application of human common sense and domain expertise, which often remains unutilized by algorithms [14].
- **Mining Processes Using Stochastic Data:** The rise in event data from diverse sources, including sensors with varying quality, leads to challenges in creating accurate process logs. This results in uncertain sensor data, complicating process mining tasks when dealing with stochastic rather than deterministic data [14].

Comprehensive Overview of Process Mining: Definitions, Scope, and Fundamental Concepts

Introduction

Over the preceding twenty years, the field of process mining has emerged as an innovative research domain, focusing on the analysis of business processes using event data as a primary source [20]. Distinct from conventional data mining approaches, which primarily concentrate on extracting relationships from the attributes of data, process mining uniquely prioritizes the exploration of business process models through the utilization of event data [21]. This transition to a comprehensive, end-to-end process viewpoint is enabled by the growing accessibility of event data, in conjunction with the development of methodologies in process discovery and conformance checking [20].

Business process models (BPMs) are crucial representations of organizational workflows, facilitating their analysis, simulation, verification, and implementation via specialized software systems [21]. Traditionally, BPMs were crafted manually, devoid of any empirical grounding [9]. This approach has been rendered obsolete by the pervasive availability of event logs, which capture detailed traces of process execution by humans, machines, and software systems [4]. Process mining techniques leverage these event logs to automate the discovery, analysis, and improvement of business processes, offering significant advantages over manual approaches [20].

In essence, Process Mining acts as a nexus between Process Science and Data Science, employing data (specifically, event data) to discover process models [4].

Definition

As defined by Wil van der Aalst, often referred to as the "Godfather of Process Mining," Process Mining is a field that aims to *discover, monitor, and enhance* actual processes [21]. This is achieved by extracting knowledge from event logs, which are abundantly available in contemporary information systems [4, 20].

Event Logs: The Foundation for Process Mining Initiatives

The ubiquity of digital event data spans all sectors, economies, organizations, and homes, and its volume is expected to increase exponentially [22]. This pervasive presence of data enables novel approaches to process analysis, relying on empirical observations rather than manually constructed models [21].

Event logs, fundamentally, are records generated within information systems during the execution of business processes [4]. These logs, however, are not immediately accessible in a usable format. They need to be extracted, cleaned, and converted into XES (Extensible Event Stream), the data format understood by classical process mining techniques [20]. The process of extracting and purifying event logs constitutes the initial phase in process mining and often demands considerable effort and time. In fact, when initiating process mining projects, the tasks of event data extraction and cleansing can take approximately 80% of the project's duration [20].

Exploring the Various Types of Process Mining

Building upon the foundational role of event logs in process mining, it is crucial to examine the spectrum of methodologies employed in this domain. These diverse types of process mining transform event log data into actionable insights, each method providing unique approaches to analyze and improve business processes:

- **Process Discovery:** This involves creating a process model based on observed example behaviors from event logs. The model should avoid overfitting (merely replicating observed traces) and underfitting (allowing behavior unrelated to what was observed) [4].
- **Conformance Checking:** This type requires both an event log and a process model. The aim is to identify discrepancies between the log and the model,

effectively measuring how well the actual process executions conform to the predefined model [4].

- **Performance Analysis:** This type of process mining aims to improve processes by identifying performance-related issues, such as delays, limited production, missed deadlines, and quality problems, by analyzing frequency and time information in event logs [20].
- **Comparative Process Mining:** It involves the analysis of multiple event logs, which could be from different locations, time periods, or case categories. This method is used to identify differences and commonalities, facilitating benchmarking and root cause analysis [20].
- **Predictive Process Mining:** This forward-looking approach uses discovered process models to predict future states of the process, employing machine learning techniques to forecast potential bottlenecks, deviations, or compliance issues [20].
- **Action-Oriented Process Mining:** This type focuses on converting process mining diagnostics into actionable improvements. It involves using process mining insights to initiate direct actions or process changes, often assisted by low-code automation platforms or Robotic Process Automation (RPA) for task automation [20].

Critical Challenges in the Field of Process Mining

Process mining, an integral component in the analysis of business processes, faces a spectrum of challenges that are pivotal to its effective application and advancement. The following enumerates and elaborates on these key challenges:

- **Finding, Merging, and Cleaning Event Data:** The challenge involves addressing issues such as distributed data sources, incomplete data, the presence of outliers, and varying granularity levels in event logs [21].
- **Dealing with Complex Event Logs Having Diverse Characteristics:** This challenge focuses on managing event logs with varying characteristics, including handling extremely large logs and deriving reliable conclusions from smaller logs [21].
- **Creating Representative Benchmarks:** The objective is to develop benchmarks comprising example datasets and quality criteria for comparing and enhancing tools and algorithms in process mining [21].
- **Dealing with Concept Drift:** This involves understanding and managing changes in the process during analysis, which is crucial for process management [21].
- **Improving the Representational Bias Used for Process Discovery:** The challenge here is to refine the selection of representational biases to achieve high-quality process mining results [21].

- **Balancing Between Quality Criteria such as Fitness, Simplicity, Precision, and Generalization:** The task is to develop models that perform well across four quality dimensions: fitness, simplicity, precision, and generalization [21].
- **Cross-organizational Mining:** This challenge addresses the analysis of event logs from multiple organizations, including cooperative processes like supply chains and shared infrastructure [21].
- **Providing Operational Support:** The focus here is on using process mining for real-time operational support, encompassing detection, prediction, and recommendation activities [21].
- **Combining Process Mining with Other Types of Analysis:** The challenge is integrating process mining with various analytical approaches like optimization, data mining, simulation, and visual analytics [21].
- **Improving Usability for Non-experts:** This involves designing user-friendly interfaces for process mining tools that automatically adjust settings and suggest analysis types [21].
- **Improving Understandability for Non-experts:** The goal is to present results in an easily understandable manner and clearly indicate the trustworthiness of the findings to prevent misinterpretations [21].

Research Directions

In light of the evolving complexities and emerging challenges in Process Mining and Business Process Management (BPM), we propose the following future research directions that hold significant potential for advancing these fields. These directions, which will be the main focus of our future works, aim to meticulously address the critical issues identified and pave the way for more efficient, accurate, and user-friendly process analysis techniques. Our dedicated efforts in these areas will seek to transform the theoretical and practical aspects of process mining and BPM, responding dynamically to the needs of an ever-changing business landscape:

- **Advanced Frameworks for Data Integration and Cleaning:** Research should focus on developing sophisticated frameworks for efficiently finding, integrating, and cleaning data from varied and distributed sources. Automated data cleaning algorithms, real-time data quality assessment tools, and meta-models for data standardization are pivotal areas that can substantially enhance the quality and usability of event data in process mining.
- **Hybrid Analytical Approaches in Process Mining:** The integration of process mining with other analytical methods such as predictive analytics, simulation, and artificial intelligence, including machine learning and natural language processing, offers a promising avenue for comprehensive process analysis. This integration can lead to the development of hybrid analytical frameworks that provide deeper insights into business processes.

– **Innovations in Automated Process (Re-)Design:** Automating the process of redesigning business processes is a challenging yet crucial area. Future research should explore AI-driven process optimization algorithms and simulation-based design testing to enable effective and efficient process (re-)design.
– **Augmenting Process Mining with Domain Expertise and Common Sense:** Incorporating domain-specific knowledge and common-sense reasoning into process mining algorithms is essential for enhancing their applicability and relevance. Developing models for contextual data interpretation and collaborative mining techniques, where domain experts can contribute to and refine the analysis, will significantly improve the outcomes of process mining.

  Each of the aforementioned research directions presents a unique opportunity to delve deeply into the specific challenges faced by Process Mining and BPM. In our future works, we are committed to exploring and thoroughly investigating these areas, with the aim of unlocking new avenues for innovation and development in these fields. By rigorously studying these challenges, we intend to contribute significantly to a more nuanced understanding of business processes. Our efforts will be geared toward developing robust and effective process management strategies, particularly critical in the context of our increasingly data-driven world. This focused exploration in our forthcoming research endeavors promises to yield valuable insights and advancements in Process Mining and BPM.

Conclusion

In conclusion, this review extensively explores the integration of Business Process Management (BPM) and Process Mining (PM) within the digital transformation landscape of Industry 4.0. It delves into the definitions, scope, and methodologies of BPM and PM, highlighting their critical roles in enhancing business process efficiency and providing insights through data analysis. The challenges and future research directions identified in this review underscore the need for advanced frameworks in data integration, hybrid analytical approaches, automated process (re-)design, and the incorporation of domain expertise in process mining. This comprehensive analysis not only provides a deeper understanding of BPM and PM but also opens avenues for future research, aiming to improve and innovate business processes in the digital era.

References

1. V. Kondarevych, K. Andriushchenko, N. Pokotylska, G. Ortina, O. Zborovska, and L. Budnyak, "Digital Transformation of Business Processes of an Enterprise," *TEM Journal*, vol. 9, no. 4, 2020.

2. M. Dumas, M. La Rosa, J. Mendling, and H. A. Reijers, *Fundamentals of business process management*. Springer, 2 ed., 2018.
3. W. M. Van Der Aalst, "Business process management: A personal view," *Business Process Management Journal*, vol. 10, no. 2, 2004.
4. W. Van der Aalst, *Process mining: Data science in action*. 2016.
5. Heiner Las, Peter Fettke, Hans Georg Kemper, Thomas Feld, and Michael Hoffmann, "Industry 4.0," *Business and Information Systems Engineering*, 2014.
6. R. Y. Y. Hung, "Business Process Management as competitive advantage: A review and empirical study," *Total Quality Management and Business Excellence*, vol. 17, no. 1, 2006.
7. A.-W. Scheer, *ARIS — Business Process Frameworks*. 1999.
8. M. Netjes, H. A. Reijers, and W. M. Van Der Aalst, "Supporting the BPM life-cycle with FileNet," in *CEUR Workshop Proceedings*, vol. 364, 2006.
9. M. Weske, *Business process management: Concepts, languages, architectures, second edition*. 2012.
10. ABPMP, *Guide to the business process management common body of knowledge (BPM CBOK)*. 2013.
11. M. Zur Muehlen and D. T. Y. Ho, "Risk management in the BPM lifecycle," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 3812 LNCS, 2005.
12. A. Hallerbach, T. Bauer, and M. Reichert, "Managing process variants in the process life cycle," in *ICEIS 2008 - Proceedings of the 10th International Conference on Enterprise Information Systems*, vol. 2 ISAS, 2008.
13. C. Houy, P. Fettke, and P. Loos, "Empirical research in business process management - analysis of an emerging field of research," *Business Process Management Journal*, vol. 16, no. 4, 2010.
14. I. Beerepoot, C. Di Ciccio, H. A. Reijers, S. Rinderle-Ma, W. Bandara, A. Burattin, D. Calvanese, T. Chen, I. Cohen, B. Depaire, G. Di Federico, M. Dumas, C. van Dun, T. Fehrer, D. A. Fischer, A. Gal, M. Indulska, V. Isahagian, C. Klinkmüller, W. Kratsch, H. Leopold, A. Van Looy, H. Lopez, S. Lukumbuzya, J. Mendling, L. Meyers, L. Moder, M. Montali, V. Muthusamy, M. Reichert, Y. Rizk, M. Rosemann, M. Röglinger, S. Sadiq, R. Seiger, T. Slaats, M. Simkus, I. A. Someh, B. Weber, I. Weber, M. Weske, and F. Zerbato, "The biggest business process management problems to solve before we die," *Computers in Industry*, vol. 146, 2023.
15. A. Van Looy, "How the COVID-19 pandemic can stimulate more radical business process improvements: Using the metaphor of a tree," *Knowledge and Process Management*, vol. 28, no. 2, 2021.
16. Kasey Panetta, "Gartner top strategic technology trends for 2021.," 2020.
17. M. Reichert and B. Weber, *Enabling flexibility in process-aware information systems: Challenges, methods, technologies*. 2012.
18. M. Malinova and J. Mendling, "Identifying do's and don'ts using the integrated business process management framework," *Business Process Management Journal*, vol. 24, no. 4, 2018.
19. C. Klinkmüller and I. Weber, "Every apprentice needs a master: Feedback-based effectiveness improvements for process model matching," *Information Systems*, vol. 95, 2021.
20. W. M. van der Aalst, "Process Mining: A 360 Degree Overview," in *Lecture Notes in Business Information Processing*, vol. 448, 2022.
21. W. V. D. Aalst, "Process Mining: Overview and Opportunities," *ACM Transactions on Management Information Systems*, vol. 3, no. 2, 2012.
22. J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. H. Byers, "Big data: The next frontier for innovation, competition, and productivity," *McKinsey Global Institute*, no. June, 2011.