Ronny Ramlau, Otmar Scherzer (Eds.)
The Radon Transform

Radon Series on Computational and Applied Mathematics

Volume 22

The Radon Transform

The First 100 Years and Beyond

Edited by
Ronny Ramlau and Otmar Scherzer

DE GRUYTER

Editors
Prof. Dr. Ronny Ramlau
University of Linz
Industrial Mathematics Institute
Altenbergerstr. 69
4040 Linz
Austria
ronny.ramlau@ricam.oeaw.ac.at

Prof. Dr. Otmar Scherzer
University of Vienna
Computational Science Center
Oskar-Morgenstern-Platz 1
1090 Vienna
Austria
otmar.scherzer@univie.ac.at

ISBN 978-3-11-055941-5
e-ISBN (PDF) 978-3-11-056085-5
e-ISBN (EPUB) 978-3-11-055951-4
ISSN 1865-3707

Library of Congress Control Number: 2019935656

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

Typesetting: VTeX UAB, Lithuania
Printing and binding: CPI books GmbH, Leck

www.degruyter.com

Contents

Ronny Ramlau and Otmar Scherzer

100 years of Mathematical Tomography

Abstract: This volume is honoring the 100th anniversary of the publication of the famous paper of Johann Radon: "Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten," which appeared in *Berichte über die Verhandlungen der Königlich-Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-Physische Klasse 69* (1917), pp. 262–277 [2].

Exactly 100 years later, in 2017, the Johann Radon Institute for Computational and Applied Mathematics (RICAM) jointly with the Johannes Kepler University Linz (JKU) held a conference in honor of the publication of Johann Radon's paper. The conference took place from March 27th to 31st in Linz, Austria https://www.ricam.oeaw.ac.at/events/conferences/radon100/ and was attended by about 170 participants who reported on the status of the still growing field of mathematical tomography.

About this volume

In 1917, Johann Radon published his fundamental paper [2], wherein he introduced what is nowadays called the Radon transform. Today, this paper is considered to be the foundation of the area of *Mathematical Tomography*, which is a booming area in applied sciences.

For a long time, this paper did not get the credit which it deserves.

Johann Radon himself appears to have not have commented on these results, beyond their initial publication. In an obituary of Radon [1], written by his colleague Paul Funk (himself a distinguished mathematician at the University of Vienna), many mathematical achievements of Johann Radon were outlined, but Paul Funk left out the Radon transform.

Only much later, driven by applications, was the paper of Johann Radon rediscovered, and today it gets the broad credit it deserves.

This volume has been collected exactly 100 years after the publication of Radon's fundamental paper [2]. It collects surveys and original papers on Mathematical Tomography and shows the remarkable advancement of the field. These new achievements are also driven by fundamental developments in imaging and industry, where computational methods based on mathematical inversion techniques allow us to vi-

Ronny Ramlau, Institute for Industrial Mathematics, Johannes Kepler University, Altenbergerstraße 69, A-4040 Linz, Austria; and Johann Radon Institute for Computational and Applied Mathematics (RICAM), Altenbergerstraße 69, A-4040 Linz, Austria, e-mail: ronny.ramlau@ricam.oeaw.ac.at
Otmar Scherzer, Johann Radon Institute for Computational and Applied Mathematics (RICAM), Altenbergerstraße 69, A-4040 Linz, Austria; and Faculty of Mathematics, University of Vienna, Oskar-Morgenstern Platz 1, A-1090 Vienna, Austria, e-mail: otmar.scherzer@univie.ac.at

https://doi.org/10.1515/9783110560855-201

sualize underlying information which is not directly accessable [3]. The interplay between mathematics and applications was not always as close, demonstrated by the well-known fact that Allan M. Cormack and Godfrey Hounsfield, who received the Nobel-prize in Physiology or Medicine for the development of the first medical CT-scanner, developed their algorithms for image reconstruction not only independently from each other but also without knowledge of Radon's work.

About this book

This book provides an overview on mathematical tomography and the Radon transform, with thirteen papers by well-known experts in the field. All articles were peer reviewed.

Karl Sigmund gives an intriguing insight in the life and work of Johann Radon. He outlines his groundbreaking mathematical achievements and his life in the most turbulent time of the last century.

Michel Defrise and Christine De Mol present applications for PET imaging and blind deconvolution. This paper is devoted to Mario Bertero, a pioneer in Inverse Problems.

Simon Gindikin studies deep results for the Radon transform in a very general setting in a historic context from Leray to Radon.

Joonas Ilmavirta and François Monard consider the problem of inversion of integral transforms on manifolds.

Roman Novikov presents recent and new results on the non-Abelian Radon transform and its inversion.

Victor Palamodov discusses the developments of the Funk–Radon theory over the last centuries.

Gaik Ambartsoumian shows recent results on V-line and Conical Radon transforms and presents applications in imaging.

Alfred Louis discusses and explains ghost artefacts in Radon problems.

Thomas Schuster discusses applications of the Radon transform in vector field tomography. His paper is devoted to Armin Lechleiter, a leading researcher in Inverse Problems, who passed away much too young, at the age of 36.

Gabor Herman applies different algorithms for x-ray inversion and compares the results for medical applications.

Tanja Tarvainen reports about active research fields in photoacoustical imaging and Bayesian inversion.

Shari Moskow and John C. Schotland report on the inverse Born series for solving imaging problems.

Andreas Alpers and Peter Gritzmann represent the field of discrete tomography in this volume.

The editors would like to dedicate the volume to the memory of Joyce McLaughlin, †2018 who was a life long supporter of RICAM, and an expert in Inverse Problems and, in particular, elastography.

Bibliography

[1] P. Funk. Nachruf auf Prof. Johann Radon. In: *Mathematische Nachrichten*, volume 62, pages 191–199, 1958. issn: 0025-584X.

[2] J. Radon. Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. In: *Berichte über die Verhandlungen der KöniglichSächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-Physische Klasse*, volume 69, pages 262–277, 1917.

[3] R. West. In industry, seeing is believing. In: *Physics world*, pages 27–30, 2003.

Karl Sigmund

1 Johann Radon 1887–1956

Abstract: This paper provides a short biography of the famous Austrian mathematician, Johann Radon, and it outlines the fundamental mathematical achievements.

Keywords: Radon transform, Radon measure, Radon numbers, Radon–Helly theorem, Radon–Riesz theorem

MSC 2010: 01A-60, 01A-70

His name is a household word in mathematics: next to the Radon transform, students learn about the theorems of Lebesgue–Radon–Nikodym, of Helly–Radon, and of Radon–Riesz; they are familiar with the concepts of Radon measure and Radon integral, and may have heard of the Radon curve and the Radon sequence. Johann Radon has certainly left an enduring mark and significantly contributed to the mathematics of the first half of the twentieth century.

By all accounts, Radon was a modest, quiet, and unassuming person, who did not make much ado about his mathematical fame. It is quite probable that the tremendous growth of the field of inverse problems would have surprised him. His colleague and near-contemporary, Paul Funk (1886–1969), a former student of Hilbert and a first-rate mathematician on his own right, did not even mention the Radon transform when he wrote Radon's obituary for the Austrian Academy of Science.

Johann Radon was born on December 16, 1987, in Tetschen on the Elbe River. The small provincial town in Bohemia lay close to Saxony. Today, it is Decin, a border town of the Czech republic. His father, a Sudeten German, was head accountant at a local bank. The mother came from Thuringia. The parents sent their only child to the gymnasium of Leitmeritz (today Litomerice). Actually, they also moved there, after Johann's father had retired. The school-boy's health was frail, but he showed great promise. He did well in mathematics and science, as well as in Latin and Greek. Moreover, he played several instruments, and had a beautiful singing voice. The intention of the gifted young man, at that time, was to become a philosopher (Figure 1.1).

After finishing secondary school with brilliant marks in 1905, Johann Radon enrolled in mathematics and physics at the University of Vienna, and also took up music

Acknowledgement: The author thanks Dr. Sienell from the Austrian Academy of Science and Dr. Maisel from the Archive of the University of Vienna. Special thanks go to Mag. Brigitte Bukovics for her advice and her kind permission to use the photographs.

Karl Sigmund, Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria, e-mail: Karl.Sigmund@univie.ac.at

https://doi.org/10.1515/9783110560855-001

Figure 1.1: Johann Radon as a schoolboy.

classes. Again, his parents followed him. Among Radon's teachers were Franz Mertens (1840–1927), an eminent number theorist; the brilliant Wilhelm Wirtinger (1865–1945), who contributed to many fields of analysis; and the associate professor of geometry, Gustav Kohn (1859–1921). Among the lecturers, or docents, were Josef Plemelj (1873–1967), who made essential contributions to the Riemann–Hilbert problem and introduced Radon to potential theory, and Alfred Tauber (1866–1942). By then, the field of what soon became known as Tauberian theorems was taking roots. Among the younger post-docs, Hans Hahn (1879–1934) and Heinrich Tietze (1880–1964) stood out. Their future paths would often cross with Radon's. At the time, Hahn mentored Radon in the calculus of variations, and introduced him to the foundations of arithmetic and geometry. There were some brilliant other students, too, senior to Radon: for instance, Roland Weitzenböck (1885–1955), Wilhelm Blaschke (1885–1962), and Eduard Helly (1884–1943).

Obviously, mathematics was in a good shape in Vienna at that time. The main responsible for this development was Gustav von Escherich (1849–1935). Escherich's principal contributions were in the calculus of variations. His scientific work alone would not have guaranteed him an outstanding place. But he was instrumental in putting Vienna on the map, in what concerned mathematics, after centuries of near-insignificance. Escherich had studied under Weierstrass in Berlin. He introduced first-rate standards of precision and clarity into the curriculum at the University of Vienna and managed to link up with international contemporary research.

Among Escherich's many brilliant students, Johann Radon stood out. He obtained his PhD on February 18, 1910, with a thesis on the calculus of variations. In that same year, he also obtained the teacher's diploma. But it was already becoming clear that he was cut out for an academic career. The next step was almost preordained. Like so many other mathematical "rookies" in his generation, Radon traveled with a stipend

to study at Hilbert's institute in Göttingen, at that time undoubtedly the mecca of mathematics.

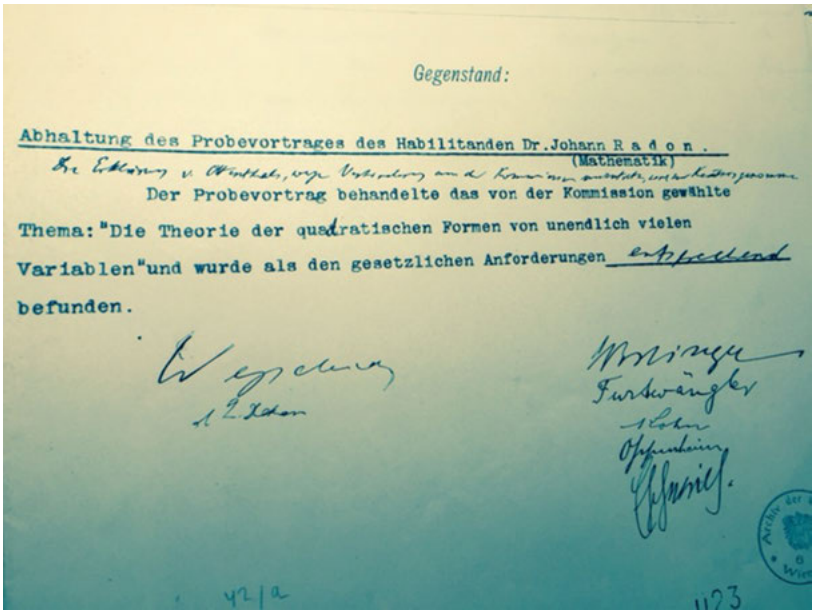
During these years, analysis experienced a profound metamorphosis. The tools developed by Henri Lebesgue's theory of integration and the problems raised by Fredholm's theory of functional operators came together in various, often unexpected ways. This had lasting influence on Radon's development.

After his post-doc semester in Göttingen, Johann Radon was appointed at the University of Brünn (today Brno), as an assistant to Emmanuel Czuber, an insurance mathematician and statistician well known for his work on probability theory. Soon afterward, in 1912, Professor Czuber was appointed at the Technische Hochschule (today: Technical University) in Vienna. Radon followed him there as his assistant, and thus returned to the buzzing capital of the Austro-Hungarian empire.

In 1913, Radon published in the *Sitzungsberichte der Akademie der Wissenschaften zu Wien* a truly seminal work of 144 pages, entitled "Theorie und Anwendungen der absolut additive Mengenfunktionen." It was a masterpiece, combining the measure theoretical ideas of the French school centering around Lebesgue and Borel with the functional analytic work pioneered by Hilbert's famous *Mitteilungen* from 1906 and the work of Fredholm and Hellinger. In Radon's own words, the aim of his paper was "to encompass in a general theory, on the one hand the theory of linear integral equations and on the other hand the theory of linear and bi-linear forms of infinitely many variables." This development of a general measure theory as the basis for functional analysis proved enormously successful. It must have been in the air: A closely parallel undertaking was due to the Hungarian Friedrich (or Frigyes) Riesz, in particular, with the latter's contribution "Über Systeme integrierbarer Funktionen."

Radon's seminal paper was strikingly well rounded and complete. It contains the decomposition theorems for signed measures, and relates measures (now called Radon-measures) with linear forms on spaces of continuous functions. It extends to the n-dimensional case a fundamental theorem which Lebesgue had derived for absolutely continuous measures on the real line. (Otton Nikodym (1887–1974) would generalize it further, and define what is now known as the Radon–Nikodym property for locally convex spaces). In addition, Radon's opus magnum also contains a highly polished duality theory of L^p-spaces, and a theorem on weak versus strong convergence which today is known as Radon–Riesz theorem. (It yields the basic motivation for the study of Radon–Riesz spaces). A final step was taken in 1915 by Maurice Fréchet, who developed measure theory on abstract spaces, rather than on topological spaces.

In 1914, Radon applied at the University of Vienna for the degree of lecturer, and submitted his 1913 paper as a habilitation thesis. The commission was duly impressed. Escherich called it "one of the most important treatises that appeared in the last years," Wirtinger viewed it as "a most important achievement" and on Furtwängler (who had succeeded Mertens after the latter's retirement), it made "a very good impression." After the probationary lecture in June on "The theory of quadratic forms


Gegenstand:

Abhaltung des Probevortrages des Habilitanden Dr. Johann R a d o n .
(Mathematik)

Der Probevortrag behandelte das von der Kommission gewählte
Thema: "Die Theorie der quadratischen Formen von unendlich vielen
Variablen" und wurde als den gesetzlichen Anforderungen entsprechend
befunden.

42/a

1123

Figure 1.2: Radon's probationary lectures satisfies the requirements for his habilitation: he can give lectures at the university.

with infinitely many variables," Radon obtained the title of Privatdozent on August 26, 1914 (Figure 1.2).

By then, the First World War was already in full swing. Johann Radon, however, was exempted from military service, thanks to his extreme myopia. During the war years, he stayed as assistant at the Technische Hochschule, and also lectured at the Universität für Bodenkultur (today, the University for Natural Resources and Life Sciences). In August 1916, he married Maria Rigele, named Mizzi, a cousin of his friend, Weitzenböck. During the next three years, three sons were born: in 1917, Wolfgang, who died very young; in 1918, Hermann; and in 1919, Ludwig. A daughter, Brigitte, followed in 1924 (Figure 1.3).

For Radon, these years were extremely productive from the scientific view-point, too. In particular, in 1917, he published "Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten," little dreaming what stellar fate this note would encounter half a century later. In addition to working on boundary value problems for logarithmic potentials, Radon did important work on convex geometry and also, influenced by the research program of Wilhelm Blaschke, on affine geometry.

This last would soon have decisive consequences for Johann Radon's career. After the defeat of 1918, the town-fathers of Hamburg courageously decided to establish a university in their town. The creation of a mathematical center for this university was

Figure 1.3: Conjugal bliss: Johan Radon and his wife Maria, née Rigele.

entrusted to Wilhelm Blaschke, then in his early thirties. Blaschke combined mathematical talents with inspired leadership: within a few years, he built up a vibrant and dynamic mathematical hot spot (Figure 1.4). Among the earliest appointees were Erich Hecke, Kurt Rademacher, Kurt Reidemeister, and a string of Austrians including Johann Radon, Emil Artin, Otto Schreier, and Wolfgang Pauli. (This migration was not all one-sided: in 1923, Kurt Reidemeister moved from Hamburg to Vienna, where he became associate professor in geometry).

When Johann Radon left Vienna with his young family, heading for an associate professorship in Hamburg, he may have hoped to return very soon from the waterkant to the shores of the Danube. Indeed, in 1919, his former professor, Gustav Escherich, had retired. The competition for his succession was intense. In the end, it was Hans Hahn who was appointed, while Johann Radon shared second place with Heinrich Tietze. It was a brilliant list, made up of three former students of Escherich who, by now, all were professors in Germany. Hahn's return from Bonn to Vienna would in due time lead to the creation of the Vienna Circle.

As for Radon, he accepted in 1922 a position as full professor in Greifswald, as the successor of Felix Hausdorff, who had moved to Frankfurt. In 1925, Radon moved to Erlangen, this time as the successor of Tietze; and in 1928, Radon moved to Breslau as the successor of Kneser. In the 1920s, such frequent displacements from one university to another were by no means uncommon. Actually, in 1929, Radon was offered a chair in Leipzig, but this time he declined it.

His scientific work continued at the highest level, and was remarkable for its diversity. Thus in 1920, Radon offered a beautiful proof of a result nowadays known as the theorem of Helly–Radon: If a family of convex bodies in n-dimensional space has the property that each $n+1$ of them intersect, then the whole family has nonempty intersection. This result had already been obtained by Eduard Helly shortly before WWI. At the outbreak of the war, Helly had volunteered for the Austro-Hungarian army. Soon

Figure 1.4: Wilhelm Blaschke as a friend of the young Radon family.

after, he was wounded on the Eastern front and had to spend, not only the rest of the war but several years afterward, in Siberian camps. Since by 1920 it was still unclear whether he would ever manage to return, Radon published his own proof (with full acknowledgement, of course, of Helly's claim to priority). Radon's proof was based on a reformulation of Helly's result: Any subset of R_n with more than $n + 2$ points carries two measures with disjoint support but same centers of gravity.

Radon kept his active interest in the calculus of variations though all his life. In 1927, in "Über die Oszillationstheoreme der konjugierten Punkte beim Problem von Lagrange," he published a solution to a problem which had dogged his former teacher Escherich throughout his career, and later returned to it in a series of lectures presented in Hamburg (Figure 1.5).

Of particular interest are Radon's contributions to the composition of quadratic forms, inspired by seminal work of Hurwitz which had shed deep insights into the study of division algebras. Every natural number n can be written in the form $2^{4k+l}m$, where m is an odd number and $0 \leq l \leq 3$. The Radon number is $\rho(n)$ defined as $2^l + 8k$. The sequence of Radon numbers

$$1, 2, 1, 4, 1, 8, 1, 9, 1, 10, 1, 12, 1, 16, 1, 17, 1, 18, 1, 20 \ldots$$

has the remarkable property that there exist $\rho(n) - 1$ linearly independent vector fields on the sphere S^{n-1}. In fact, Frank Adams proved in 1963 that this is the precise upper bound.

On several occasions, in the 1930s, it seemed that Radon would be offered to return to the University of Vienna, due to a string of vacancies. Thus in 1934, Hahn had unexpectedly died. But the ministry abolished his chair, apparently as a gesture directed against the Vienna Circle. Next, in 1936, Wirtinger retired. But again, the job was neither offered to Johann Radon, nor to Karl Menger, nor to Emil Artin, eminent candidates all, but to Wirtinger's assistant, Karl Mayrhofer, a mathematician of much lesser caliber who later turned out to have been an illegal National Socialist all along.

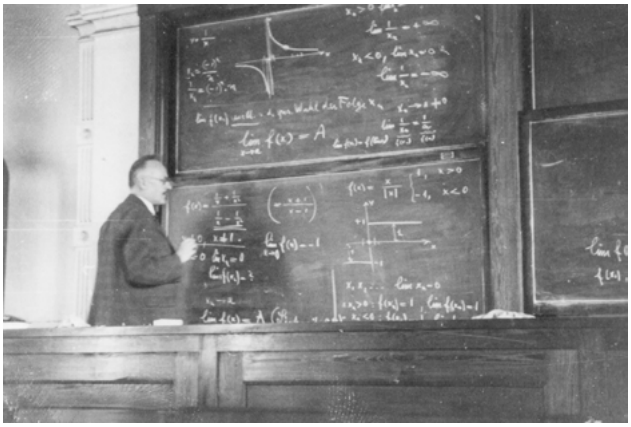


Figure 1.5: Radon as professor in Breslau (today, Wroclaw).

Finally, in 1938, when Furtwängler retired, the mathematical nonentity, Anton Huber, another long-standing party member, was appointed in his place. By then, the "Anschluss" of Austria to the Third Reich had already taken place.

In 1939, the Radon family lost their second son, Hermann, who had been grievously ill for many years. Radon learned the tragic news while he himself was hospitalized and recovering from surgery.

World War II broke out, and in 1943, Radon's third son, Ludwig, was fatally wounded on the Russian front. In 1945, the Red Army fought its way into Breslau. Johann Radon, his wife and his daughter had to flee, and found refuge in far-off Innsbruck. Breslau became Wroclav, in Poland. It soon became clear that Radon could never hope to return to his home or to his position; but Vienna became an option again.

In the chaotic aftermath of the war, the two professors of mathematics of the University of Vienna, Karl Mayrhofer and Anton Huber, had to step down from office, due to their National Socialist past. A committee was instructed to look for two candidates. When it was found out that Radon lived in Innsbruck, and actually lectured at the University there, the problem seemed half-solved. But due to intentional or unintentional misinformation, rumors spread that Radon was unwilling to move to Vienna. Fortunately, Hans Thirring, who had returned in 1945 to his chair in theoretical physics, was able to dispel these doubts, and after an exchange of hectic telegrams, Johann Radon was finally appointed to a chair at the University of Vienna (Figure 1.6).

In 1946, Radon had also been offered chairs in Greifswald, Jena, and Leipzig, thanks not only to his scientific reputation, but also to the fact that he was among the few scientists in Germany whose political past was without blemish. However, these universities were all beyond the line that was soon to mark the Iron Curtain. Not that life promised to be easy in bombed-out Vienna, divided as it was into four occupation zones.

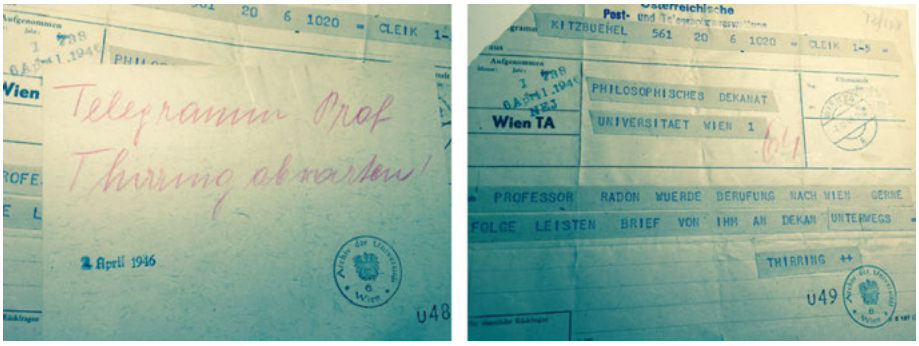


Figure 1.6: A telegram confirms that Radon is ready to move to Vienna.

On October 1, 1946, Johann Radon took up his duties at the University of Vienna. At first, he had to run the mathematical department almost single-handed. The situation improved only in 1948 with the appointment of the young Edmund Hlawka to the second chair of mathematics, and with the able assistance of Nikolaus Hofreiter and Leopold Schmetterer. Thus Radon was able to concentrate again on his favorite fields: real analysis, calculus of variations, and differential geometry. His lectures were widely praised as mathematical gems. More and more, he embodied the ideal of the world-famous scientist, a kind grandfatherly figure. He enjoyed house concerts and family life. His daughter, Brigitte, had acquired the teachers' diploma in mathematics. In 1950, she married Erich Bukovics, an up-and-coming young mathematician who would later become professor at the Technical University. The young couple had two sons.

However, it proved impossible to avoid the administrative burdens. From 1947 onward, Radon was editor-in-chief of the "Monatshefte für Mathematik," which had been founded, generations ago, by his teacher Gustav von Escherich. In 1947, he also became full member of the Austrian Academy of Science. In 1951/52, Radon held office as dean of the huge philosophical faculty. From 1950 to 1952, he headed the Austrian Mathematical Society. In 1953, he became secretary general of the Austrian Academy of Science, and in 1954, he was elected rector of the University (Figure 1.7). The various inaugural lectures gave him some opportunity to return to his early philosophical interests. But he had to write more and more obituaries.

On May 26, 1956, Johann Radon died, aged sixty-nine. He had long suffered from a weak heart and a probably congenital lung-ailment. In that same year, 1956, Alan Cormack discovered tomography, in complete ignorance of the work that Radon had published some 40 years ago in an obscure journal. In 1971, Houndsfield constructed the first CT scanner. In 1979, Houndsfield and McCormack were rewarded with the Nobel Prize. By then, the mathematical roots of medical imaging had been rediscovered, and inverse problems were well on their way of becoming a leading branch of mathematics.

Figure 1.7: Radon as rector of the University of Vienna.

Bibliography

[1] Über das Minimum des Integrals $\int F(x, y, \rho, k)ds$. *Sitzungsber. Kaiserl. Akad. Wiss. Wien, Math.-Nat. Kl.*, 1257–1326, 1910.

[2] Zur Theorie der Mayerschen Felder beim Lagrangeschen Variationsproblem. *Sitzungsber. Kaiserl. Akad. Wiss. Wien, Math.-Nat. Kl.*, 1337–1360, 1911.

[3] Über einige Fragen betreffend die Theorie der Maxima und Minima mehrfacher Integrale. *Montash. Math. Phys.*, 53–63, 1911.

[4] Lineare Scharen orthogonaler Matrizen. *Abh. Math. Sem. Univ. Hamburg*, 1–14, 1912.

[5] Theorie und Anwendungen der absolut additiven Mengenfunktionen. *Kaiserl. Akad. Wiss. Wien, Math.-Nat. Kl.*, 1295–1438, 1913.

[6] Berichtigung, Die Kettenlinie bei allgemeinster Massenverteilung. *Sitzungsber. Kaiserl. Akad. Wiss. Wien, Math.-Nat. Kl.*, 339, 1916.

[7] Die Kettenlinie bei allgemeinster Massenverteilung. *Sitzungsber. Kaiserl. Akad. Wiss. Wien, Math.-Nat. Kl.*, 221–240, 1916.

[8] Über eine besondere Art ebener konvexer Kurven. *Ber. Math.-Phys. Kl. Sächs. Akad. Wiss. Leipzig*, 123–128, 1916.

[9] Über eine Erweiterung des Begriffs der konvexen Funktionen mit einer Anwendung auf die Theorie der konvexen Körper. *Sitzungsber. Kaiserl. Akad. Wiss. Wien, Math.-Nat. Kl.*, 241–258, 1916.

[10] Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. *Ber. Math.-Phys. Kl. Sächs. Akad. Wiss. Leipzig*, 262–277, 1917.

[11] Über affine Geometrie XVI: Die Grundgleichungen der affinen Flächentheorie. *Ber. Math.-Phys. Kl. Sächs. Akad. Wiss. Leipzig*, 91–107, 1918.

[12] Über affine Geometrie XVII: Zur Affingeometrie der Regelflächen. *Ber. Math.-Phys. Kl. Sächs. Akad. Wiss. Leipzig*, 147–155, 1918.

[13] Über die Randwertaufgaben beim logarithmischen Potential. *Sitzungsber. Akad. Wiss. Wien, Math.-Nat. Kl.*, 1123–1167, 1919.

[14] Kurzfassung, Über die Randwertaufgaben beim logarithmischen Potential. *Anz. Akad. Wiss. Math.-Nat. Kl.*, 190, 1919.

[15] Über lineare Funktionaltransformationen und Funktionalgleichungen. *Sitzungsber. Akad. Wiss. Wien, Math.-Nat. Kl.*, 1083–1121, 1919.

[16] Kurzfassung, Über lineare Funktionaltransformationen und Funktionalgleichungen. *Anz. Akad. Wiss. Math.-Nat. Kl.*, 189, 1919.

[17] Kurzbericht, Bestimmung einer Riemannschen Metrik durch Krümmungseigenschaften. *Jber. Deutsch. Math. Ver.*, 76, 1921.
[18] Mengen konvexer Körper, die einen gemeinsamen Punkt enthalten. *Math. Ann.*, 113–115, 1921.
[19] Über statische Gravitationsfelder. *Abh. Math. Sem. Univ. Hamburg*, 268–280, 1922.
[20] Geometrie und Raumvorstellung. *Naturwissenschaften*, 510–511, 1924.
[21] Über konforme Geometrie VI: Kurvennetze auf Flächen und im Raume von Riemann. *Abh. Math. Sem. Univ. Hamburg*, 45–53, 1924.
[22] Berichtigung, Zur Behandlung geschlossener Extremalen in der Variationsrechnung. *Abh. Math. Sem. Univ. Hamburg*, 13–14, 1926.
[23] Hermann Rothe. *Jber. Deutsche Math. Ver.*, 172–175, 1926.
[24] Mechanische Herleitung des Parallelismus von T. Levi-Civita. In F. Klein and W. Blaschke, editors, *Vorlesungen über höhere Geometrie*, pages 331–346. Springer, Berlin, 1926.
[25] Zur Behandlung geschlossener Extremalen in der Variationsrechnung. *Abh. Math. Sem. Univ. Hamburg*, 195–205, 1926.
[26] Über die Differentialgleichungen von Monge. Ihre Beziehungen zur Theorie der partiellen Differentialgleichungen erster Ordnung und zur Variationsrechung. In F. Klein and W. Blaschke, editors, *Vorlesungen über höhere Geometrie*, pages 367–379. Springer, Berlin, 1926.
[27] Über konforme Geometrie V: Neue Kennzeichnung der zyklischen Kurvennetze. *Abh. Math. Sem. Univ. Hamburg*, 313–320, 1926.
[28] Mathematik und Wirklichkeit. *Sitzungsber. Phys.-Med. Soz. Erlangen.*, 181–190, 1926/27.
[29] Über die Oszillationstheoreme der konjugierten Punkte beim Problem von Lagrange. *Sitzungsber. Bayer. Akad. Wiss. Math.-Nat. Abt.*, 243–257, 1927.
[30] Bestimmung einer Riemannschen Metrik durch Krümmungseigenschaften. *Monatsh. Math. Phys.*, 9–24, 1928.
[31] Zum Problem von Lagrange. *Abh. Math. Sem. Univ. Hamburg*, 273–299, 1928.
[32] Lösung einer Aufgabe, gestellt im Jber. Deutsch. Math.-Verein 43, 20-20 (1933). *Jber. Deutsch. Math. Ver.*, 20–22, 1933.
[33] Restausdrücke bei Interpolations- und Quadraturformeln durch bestimmte Integrale. *Monatsh. Math. Phys.*, 389–396, 1935.
[34] Annäherung konvexer Körper durch analytisch begrenzte. *Monatsh. Math. Phys.*, 340–344, 1936.
[35] Bewegungsinvariante Variationsprobleme betreffend Kurvenscharen. *Abh. Math. Sem. Univ. Hamburg*, 70–82, 1937.
[36] Singuläre Variationsprobleme. *Jber. Deutsch. Math. Ver.*, 220–232, 1937.
[37] Ein Satz der Matrizenrechnung und seine Bedeutung für die Analysis. *Monatsh. Math. Phys.*, 198–204, 1939.
[38] Über Tschebyscheff-Netze auf Drehflächen und eine Aufgabe aus der Variationsrechnung. *Mitt. Math. Ges.*, 147–151, 1940.
[39] Ein einfacher Beweis für die Halbstetigkeit der Integrale der Variationsrechnung. *Math. Ann.*, 205–209, 1944.
[40] Wilhelm Wirtinger. Nachruf. *Alm. Akad. Wiss. Wien*, 336–346, 1947.
[41] Zur mechanischen Kubatur. *Monatsh. Math.*, 286–300, 1948.
[42] Godfrey Harold Hardy. Nachruf. *Alm. Öster. Akad. Wiss.*, 313–316, 1949.
[43] Über geschlossene Extremalen und eine einfache Herleitung der isoperimetrischen Ungleichung. *Ann. Mat. Pura Appl.*, 315–320, 1949.
[44] Bemerkungen zu, Über geschlossene Extremalen und eine einfache Herleitung der isoperimetrischen Ungleichung. *Ann. Mat. Pura Appl.*, 309, 1950.
[45] Constantin Caratheodory. *Nachruf. Alm. Öster. Akad. Wiss.*, 323–328, 1950.
[46] Nachruf für R. Suppantschitsch. *Nachr. Österr. Math. Ges.*, 7–8, 1950.

[47] Zur Polynomentwicklung analytischer Funktionen. *Math. Nachr.*, 156–157, 1950/51.
[48] Emanuel Czuber zum Gedächtnis. *Nachr. Österr. Math. Ges.*, 11–13, 1951.
[49] Gleichgewicht und Stabilität gespannter Netze. *Arch. Math.*, 309–316, 1954.
[50] Heinrich Tietze. *Nachr. Österr. Math. Ges.*, 74–78, 1954.
[51] Mathematik und Naturerkenntnis. Inaugurationsvortrag Universität Wien, 1954.
[52] P. M. Gruber, E. Hlawka, W. Nöbauer, and L. Schmetterer, editors. *Collected Works (2 vols)*. Verlag der Österr. Akad. Wiss. Wien, 1987.

Michel Defrise and Christine De Mol

2 On blind imaging, NMF and PET

Abstract: After reviewing the multiplicative iterative algorithms EMML and ISRA which allow to naturally enforce, at each iteration, a positivity constraint on the solution of linear inverse problems, we discuss their extension to blind imaging with positivity constraints and to the related problem of Nonnegative Matrix Factorization (NMF). We then establish a connection between this framework and two problems of interest in medical imaging, namely dynamic Positron Emission Tomography (PET) and the joint estimation of activity and attenuation in time-of-flight PET.

Keywords: Multiplicative algorithms, blind imaging, positivity, nonnegative matrix factorization, dynamic positron emission tomography, time-of-flight positron emission tomography

MSC 2010: 35R30, 44A12

2.1 The multiplicative algorithms EMML and ISRA

In many applications, positivity has proved a very useful property to be enforced on the solution of linear inverse problems, which consist in solving the equation

$$\boldsymbol{K}\boldsymbol{x} = \boldsymbol{y} \tag{1}$$

formulated here in a discrete setting, i. e., where $\boldsymbol{x} \in \mathbb{R}^M$ is a vector of coefficients describing the unknown object, $\boldsymbol{y} \in \mathbb{R}^N$ the vector of (noisy) data and $\boldsymbol{K}$ the linear operator (here a $N \times M$ matrix) modeling the link between the two.

In the presence of noise, instead of trying to solve exactly the above equation, the standard strategy is to minimize some contrast or cost function such as a least-squares distance between $\boldsymbol{K}\boldsymbol{x}$ and $\boldsymbol{y}$ in the case of Gaussian noise or a Kullback–Leibler discrepancy in the case of Poisson noise. When the problem is ill-conditioned, as typical for discretized ill-posed problems, one can regularize the problem by adding to the contrast function one or more penalties on the solution, such as the classical quadratic squared L^2-norm $\|\boldsymbol{x}\|_2^2 = \sum_m |\boldsymbol{x}_m|^2$ or the sparsity-enforcing L^1-norm $\|\boldsymbol{x}\|_1 = \sum_m |\boldsymbol{x}_m|$.

Note: Dedicated to Professor Mario Bertero on the occasion of his 80th birthday.
Acknowledgement: Michel Defrise acknowledges support by the FWO Project G027514N and the Strategic Research Program SRP10 of the VUB, and Christine De Mol by the network IAP P7/06 StUDys.

***Corresponding author: Christine De Mol,** Department of Mathematics and ECARES, Université libre de Bruxelles, Brussels, Belgium, e-mail: demol@ulb.ac.be
Michel Defrise, Department of Nuclear Medicine, Vrije Universiteit Brussel, Brussels, Belgium, e-mail: Michel.Defrise@vub.be

https://doi.org/10.1515/9783110560855-002

Positivity naturally arises in the case of photon-counting measurements, i. e., in the case of Poisson noise, where the cost function to be minimized is the following generalized Kullback–Leibler divergence (which turns out to be the negative log-likelihood for this case)

$$F(\boldsymbol{x}) = \mathrm{KL}(\boldsymbol{y}, \boldsymbol{K}\boldsymbol{x}) \equiv \sum_{n=0}^{N-1} \left[\boldsymbol{y}_n \ln\left(\frac{\boldsymbol{y}_n}{(\boldsymbol{K}\boldsymbol{x})_n} \right) - \boldsymbol{y}_n + (\boldsymbol{K}\boldsymbol{x})_n \right] \tag{2}$$

subject to the constraint $\boldsymbol{x} \geq \boldsymbol{0}$. It is assumed that $\boldsymbol{K} \geq \boldsymbol{0}$ and $\boldsymbol{y} \geq \boldsymbol{0}$. This means that all elements of the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ and of the matrix $\boldsymbol{K}$ are nonnegative. Notice that it is such entrywise notion of nonnegativity (or else of positivity) which will be used throughout the paper. To find the minimum of this convex cost function, a celebrated iterative algorithm has been proposed and goes under the name Expectation Maximization Maximum Likelihood (EMML) in medical imaging [42] and Richardson [41]—Lucy [30] in astronomy. The successive iterates are given by

$$\boldsymbol{x}^{(k+1)} = \frac{\boldsymbol{x}^{(k)}}{\boldsymbol{K}^T \boldsymbol{1}_N} \circ \boldsymbol{K}^T \frac{\boldsymbol{y}}{\boldsymbol{K}\boldsymbol{x}^{(k)}} \quad (k = 0, 1, \ldots) \tag{3}$$

using the Hadamard (entrywise) product $\circ$ and division ($\boldsymbol{K}^T$ denotes the transpose of $\boldsymbol{K}$ and $\boldsymbol{1}_N$ is a vector of ones in $\mathbb{R}^N$). When initialized by a strictly positive vector, $\boldsymbol{x}^{(0)} > \boldsymbol{0}$, the algorithm automatically preserves the positivity of the iterates $\boldsymbol{x}^{(k)}$. Another nice feature is that it ensures a monotonic decrease of the cost function, a property that is most easily established by viewing it as a Majorization–Minimization (MM) algorithm [25]. Indeed, as noticed by De Pierro [15], it can be rewritten as

$$\boldsymbol{x}^{(k+1)} = \arg\min_{\boldsymbol{x}} G(\boldsymbol{x}; \boldsymbol{x}^{(k)}) \tag{4}$$

where $G(\boldsymbol{x}; \boldsymbol{a})$ is the following "*surrogate*" cost function for $F(\boldsymbol{x})$

$$\begin{aligned} G(\boldsymbol{x}; \boldsymbol{a}) = \sum_{n=0}^{N-1} \Bigg[& \boldsymbol{y}_n \ln \boldsymbol{y}_n - \boldsymbol{y}_n + (\boldsymbol{K}\boldsymbol{x})_n \\ & - \frac{\boldsymbol{y}_n}{(\boldsymbol{K}\boldsymbol{a})_n} \sum_{m=0}^{M-1} \boldsymbol{K}_{n,m} \boldsymbol{a}_m \ln\left(\frac{\boldsymbol{x}_m}{\boldsymbol{a}_m} (\boldsymbol{K}\boldsymbol{a})_n \right) \Bigg] \end{aligned} \tag{5}$$

satisfying the surrogating properties

$$G(\boldsymbol{x}; \boldsymbol{a}) \geq F(\boldsymbol{x}) \quad \text{and} \quad G(\boldsymbol{a}; \boldsymbol{a}) = F(\boldsymbol{a}) \tag{6}$$

for all $\boldsymbol{x}$ and $\boldsymbol{a}$, with the restriction in this case that $\boldsymbol{x} \geq \boldsymbol{0}$ and $\boldsymbol{a} \geq \boldsymbol{0}$. Since this surrogate is separable (i. e., it can be written as a sum of terms, where each term depends only on a single unknown component $\boldsymbol{x}_m$), it can be easily minimized explicitly, yielding the EMML algorithm. The surrogating properties ensure that $F(\boldsymbol{x}^{(k+1)}) \leq F(\boldsymbol{x}^{(k)})$,

which is often considered as a desirable property for a good numerical behavior of the algorithm. A possible drawback, however, is slow convergence to a minimizer of the cost function. The convergence proof for the EMML algorithm can be found in the literature (see, e. g., [33] and the references therein).

An analogue of the EMML iterative algorithm for the case of Gaussian noise, i. e., of the minimization of the least-squares (negative log-likelihood) cost function

$$F(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{K}\boldsymbol{x} - \boldsymbol{y}\|_2^2 \tag{7}$$

subject to $\boldsymbol{x} \geq \boldsymbol{0}$, assuming $\boldsymbol{K} \geq \boldsymbol{0}$ and $\boldsymbol{y} \geq \boldsymbol{0}$, is called Image Space Reconstruction Algorithm (ISRA) and has been proposed in [12] and [14].

The successive multiplicative updates are given by

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} \circ \frac{\boldsymbol{K}^T\boldsymbol{y}}{\boldsymbol{K}^T\boldsymbol{K}\boldsymbol{x}^{(k)}} \tag{8}$$

and, again, positivity is automatically preserved if $\boldsymbol{x}^{(0)} > \boldsymbol{0}$. The algorithm is easily derived through the separable surrogates, for $\boldsymbol{x} \geq \boldsymbol{0}$, $\boldsymbol{a} \geq \boldsymbol{0}$,

$$G(\boldsymbol{x};\boldsymbol{a}) = \frac{1}{2}\sum_{n=0}^{N-1}\frac{1}{(\boldsymbol{K}\boldsymbol{a})_n}\sum_{m=0}^{M-1}\boldsymbol{K}_{n,m}\boldsymbol{a}_m\left[\boldsymbol{y}_n - \frac{\boldsymbol{x}_m}{\boldsymbol{a}_m}(\boldsymbol{K}\boldsymbol{a})_n\right]^2. \tag{9}$$

Hence the cost function decreases monotonically throughout the iteration process and can be shown to converge to its minimal value. Moreover, the iterates themselves have been shown to converge to the minimizer of (7) when it is unique [14] and to a minimizer of (7) when $\boldsymbol{K}$ is singular [19].

In the case of an ill-conditioned problem, however, the minimizers are unstable with respect to the noise on the data. The algorithms, being unregularized, typically exhibit a semi-convergence behavior. On numerical simulations involving a "true" solution $\boldsymbol{x}$, one can observe that the reconstruction error first decreases with the number of iterations, then passes through a minimum and increases due to noise amplification. In practice, regularization is usually achieved by early stopping of the iterative process, before instabilities occur (see [5] for a thorough discussion of this semi-convergence phenomenon).

2.2 Blind imaging with positivity and NMF

In the previous section, the operator or matrix $\boldsymbol{K}$ modeling the imaging process was supposed to be known. In many instances, however, it is unknown, in which case the inverse problem is said to be "*blind*." The blind problem can be formulated as the minimization of (2) or (7) for both unknowns $\boldsymbol{x}$ and $\boldsymbol{K}$. A major difficulty then arises from the

fact that, although the resulting functional is convex with respect to $\boldsymbol{x}$ or $\boldsymbol{K}$ separately, it is not jointly convex, leading to all drawbacks of a nonconvex optimization setting (local minima, saddle points, etc.). Due to such bi-convexity property, however, a natural strategy appears to be an alternate minimization on $\boldsymbol{x}$ (with $\boldsymbol{K}$ fixed) and $\boldsymbol{K}$ (with $\boldsymbol{x}$ fixed). Such alternating minimization approaches date back to [3, 10, 11, 43] for alternating least squares or to [20] for alternating Richardson–Lucy. Notice that when the imaging operator $\boldsymbol{K}$ is translation-invariant, the problem is also referred to as "*Blind Deconvolution*." There is a vast literature on the subject, which we will not review here. Let us just single out the paper [8] for its convergence result on regularized alternating least squares, as well as the noniterative and nonlinear inversion method of [23], with an interesting uniqueness result, but with some drawbacks analyzed in [9].

In this chapter, we will focus on the case where positivity constraints apply, and in particular on generalizations of the EMML and ISRA algorithms to the blind case.

Let us first remark that the formulation of the problem can be easily extended to include multiple inputs/unknowns ($\boldsymbol{x}$ becomes a $M \times P$ matrix $\boldsymbol{X}$) and multiple outputs/measurements ($\boldsymbol{y}$ becomes a $N \times P$ matrix $\boldsymbol{Y}$). The blind inverse problem is then equivalent to solving the equation $\boldsymbol{KX} = \boldsymbol{Y}$ for both $\boldsymbol{K}$ and $\boldsymbol{X}$ or else to "*Nonnegative Matrix Factorization*" (NMF) when all elements of these three matrices are assumed to be nonnegative. NMF is used as a data dimension reduction method analogous to Singular Value Decomposition (SVD) or Principal Component Analysis (PCA) [18, 28]. In the case of noisy data, exact factorization should be replaced by the minimization of a discrepancy cost function. The inclusion of multiple outputs can contribute to somehow compensate for the difficulty of the blind imaging problem, by making use of the information contained in several images produced by the same instrument (e. g., a telescope). Such formulation is also typically met in so-called "*Hyperspectral Imaging*" problems where the same object is observed with different wavelengths (see, e. g., the review papers [31] and [21]).

Another point of attention is the ill-conditioning of the linear inverse problem which induces to include some regularizing penalties in the cost function. To fix the ideas, we introduce penalties on the Frobenius norm of $\|\boldsymbol{K}\|_F^2 = \sum_{n,m} \boldsymbol{K}_{n,m}^2$ of $\boldsymbol{K}$, as well as on the Frobenius and on the L^1-norm $\|\boldsymbol{X}\|_1 = \sum_{m,p} |\boldsymbol{X}_{m,p}|$ of $\boldsymbol{X}$ and tune them with the (positive) regularization parameters μ, ν, and λ, respectively. The following algorithm, however, can be easily generalized to include other separable penalties. Using at each iteration step, the surrogate cost function (5) for the Kullback–Leibler discrepancy, the alternating minimization, for $\boldsymbol{K}$, $\boldsymbol{X}$ nonnegative (assuming $\boldsymbol{Y}$ nonnegative, too) of the resulting (bi-convex) cost function

$$F(\boldsymbol{K}, \boldsymbol{X}) = \mathrm{KL}(\boldsymbol{Y}, \boldsymbol{KX}) + \frac{\mu}{2} \|\boldsymbol{K}\|_F^2 + \lambda \|\boldsymbol{X}\|_1 + \frac{\nu}{2} \|\boldsymbol{X}\|_F^2 \tag{10}$$

with

$$\mathrm{KL}(\boldsymbol{Y}, \boldsymbol{KX}) = \sum_{n=0}^{N-1} \sum_{p=0}^{P-1} \left[(\boldsymbol{Y})_{n,p} \ln\left(\frac{(\boldsymbol{Y})_{n,p}}{(\boldsymbol{KX})_{n,p}} \right) - (\boldsymbol{Y})_{n,p} + (\boldsymbol{KX})_{n,p} \right] \tag{11}$$

can be done column by column for $\boldsymbol{X}$ and row by row for $\boldsymbol{K}$. The successive multiplicative updates for $\boldsymbol{K}$ and $\boldsymbol{X}$ are then explicitly given by

$$\boldsymbol{K}^{(k+1)} = \frac{2\boldsymbol{A}^{(k)}}{\boldsymbol{B}^{(k)} + \sqrt{\boldsymbol{B}^{(k)} \circ \boldsymbol{B}^{(k)} + 4\mu\boldsymbol{A}^{(k)}}} \tag{12}$$

where

$$\boldsymbol{A}^{(k)} = \boldsymbol{K}^{(k)} \circ \frac{\boldsymbol{Y}}{\boldsymbol{K}^{(k)}\boldsymbol{X}^{(k)}} \left(\boldsymbol{X}^{(k)}\right)^T \tag{13}$$

$$\boldsymbol{B}^{(k)} = \mathbf{1}_{N\times P} \left(\boldsymbol{X}^{(k)}\right)^T \tag{14}$$

($\mathbf{1}_{N\times P}$ denotes the $N \times P$ matrix of ones) and by

$$\boldsymbol{X}^{(k+1)} = \frac{2\boldsymbol{C}^{(k+1)}}{\boldsymbol{D}^{(k+1)} + \sqrt{\boldsymbol{D}^{(k+1)} \circ \boldsymbol{D}^{(k+1)} + 4\nu\boldsymbol{C}^{(k+1)}}} \tag{15}$$

where

$$\boldsymbol{C}^{(k+1)} = \boldsymbol{X}^{(k)} \circ \left(\boldsymbol{K}^{(k+1)}\right)^T \frac{\boldsymbol{Y}}{\boldsymbol{K}^{(k+1)}\boldsymbol{X}^{(k)}} \tag{16}$$

$$\boldsymbol{D}^{(k+1)} = \lambda\mathbf{1}_{M\times P} + \left(\boldsymbol{K}^{(k+1)}\right)^T \mathbf{1}_{N\times P}. \tag{17}$$

The algorithm is to be initialized with strictly positive but otherwise arbitrary $\boldsymbol{K}^{(0)}$ and $\boldsymbol{X}^{(0)}$.

As a special case for $\lambda = \mu = \nu = 0$, we recover the blind algorithm proposed by Lee and Seung [28], which in turn reduces to the EMML/Richardson–Lucy algorithm for $\boldsymbol{K}$ fixed.

Similarly, for the case of Gaussian noise, we can minimize the following bi-convex cost function alternately row by row for $\boldsymbol{K}$, and column by column for $\boldsymbol{X}$ (assumed to be nonnegative, as well as $\boldsymbol{Y}$),

$$F(\boldsymbol{K}, \boldsymbol{X}) = \frac{1}{2}\|\boldsymbol{Y} - \boldsymbol{K}\boldsymbol{X}\|_F^2 + \frac{\mu}{2}\|\boldsymbol{K}\|_F^2 + \lambda\|\boldsymbol{X}\|_1 + \frac{\nu}{2}\|\boldsymbol{X}\|_F^2 \tag{18}$$

using the surrogate (9), and derive the successive multiplicative updates

$$\boldsymbol{K}^{(k+1)} = \boldsymbol{K}^{(k)} \circ \frac{\boldsymbol{Y}(\boldsymbol{X}^{(k)})^T}{\boldsymbol{K}^{(k)}\boldsymbol{X}^{(k)}(\boldsymbol{X}^{(k)})^T + \mu\boldsymbol{K}^{(k)}} \tag{19}$$

$$\boldsymbol{X}^{(k+1)} = \boldsymbol{X}^{(k)} \circ \frac{(\boldsymbol{K}^{(k+1)})^T\boldsymbol{Y}}{(\boldsymbol{K}^{(k+1)})^T\boldsymbol{K}^{(k+1)}\boldsymbol{X}^{(k)} + \nu\boldsymbol{X}^{(k)} + \lambda\mathbf{1}_{M\times P}} \tag{20}$$

again to be initialized with strictly positive but otherwise arbitrary $\boldsymbol{K}^{(0)}$ and $\boldsymbol{X}^{(0)}$. We recover as special cases blind algorithms proposed in [22] for $\mu = 0$, $\nu = 0$, in [29] for $\lambda = 0$, $\mu = 0$, $\nu = 0$, and ISRA for $\boldsymbol{K}$ fixed and $\lambda = \mu = \nu = 0$.

As concerns the convergence of the iterates in both the Poisson and Gaussian cases, the following results can be established [26, 27] under the assumption that μ and either ν or λ are strictly positive and that Y has at least one strictly positive element in each row and each column:

1. The cost function decreases monotonically due the surrogate properties and monotonicity is strict iff $(\boldsymbol{K}^{(k+1)}, \boldsymbol{X}^{(k+1)}) \neq (\boldsymbol{K}^{(k)}, \boldsymbol{X}^{(k)})$.
2. The sequence of the values of the cost function $F(\boldsymbol{K}^{(k)}, \boldsymbol{X}^{(k)})$ converges.
3. An asymptotic regularity property holds for the sequence of iterates: $\forall n, m, p$, $\lim_{k\to+\infty}(\boldsymbol{K}^{(k+1)}_{n,m} - \boldsymbol{K}^{(k)}_{n,m}) = 0; \lim_{k\to+\infty}(\boldsymbol{X}^{(k+1)}_{m,p} - \boldsymbol{X}^{(k)}_{m,p}) = 0$.
4. As a consequence, by Ostrowski's theorem ([35], Theorem 26.1), the set of accumulation points of the sequence of iterates $(\boldsymbol{K}^{(k)}, \boldsymbol{X}^{(k)})$ is compact and connected.
5. Hence, if this set is finite, the iterates $(\boldsymbol{K}^{(k)}, \boldsymbol{X}^{(k)})$ converge. Moreover, it can be shown that they converge to a stationary point $(\boldsymbol{K}^{*}, \boldsymbol{X}^{*})$, i. e., a point satisfying the first-order Karush–Kuhn–Tucker conditions.

Notice that the algorithms above, as well as these convergence results, can be generalized to the case where an additional normalization constraint is implemented at each iteration [26], namely $\sum_m \boldsymbol{K}_{n,m} = 1$, a natural constraint met in several applications such as NMF for hyperspectral imaging.

Some applications lead to "*semi-blind*" or "*myopic*" problems, where partial knowledge is available on the structure of $\boldsymbol{K}$. Examples include deconvolution problems, where $\boldsymbol{K}$ might be a Toeplitz matrix, and the joint estimation of activity and attenuation in PET discussed in Section 2.4. Alternate minimization is still applicable to these semi-blind problems provided the known structure of $\boldsymbol{K}$ is enforced at each iteration.

Unfortunately, for the time being, a complete convergence proof of the iterates to a stationary point is not available and does not seem so easy to obtain despite several attempts in that direction. This does not preclude other convergence results obtained with different types of algorithms such as proximal alternating minimization and projection methods for nonconvex problems [2, 6, 7] or constrained gradient methods based on the Scaled Gradient Projection (SGP) algorithm [36].

2.3 Nonnegative matrix factorization in dynamic PET

We briefly describe an application of the previous algorithms to medical imaging. Positron emission tomography allows estimating the spatial distribution in the body of a radioactive tracer, which decays by emitting positrons. An emitted positron annihilates with an electron to produce two 511 keV photons, which are emitted back to back. Pairs of annihilation photons are detected in a ring of detectors around the

patient, and the number of photon pairs (*coincident events*) detected by a pair of detectors is related to the integral of the tracer concentration along the line of response (LOR) linking the two detectors [32]. These coincidence data are histogrammed in L lines of response (LOR), each identified by a pair of detectors in the scanner. Using the EMML algorithm, these data are reconstructed to yield an estimate of the tracer concentration in each voxel of a discretized image matrix.

Most clinical protocols in PET are static: they involve a single measurement taken typically one hour after injecting the radiolabeled tracer, when its biodistribution can be considered as stabilized. Additional diagnostic information is potentially provided by the time evolution of the tracer concentration. In dynamic PET, data are measured for a set of T so-called *time frames* starting directly after injecting the tracer. Image reconstruction—typically via the EMML algorithm—from these data leads to a set of T activity images, which can be organized as a nonnegative $N \times T$ matrix $\boldsymbol{Y}$, with $\boldsymbol{Y}_{n,t}$ the estimated activity in voxel n at time t. Note that in this section we consider the reconstructed activity images as the data, hence the use of the symbol $\boldsymbol{Y}$.

A large variety of methods has been developed to reconstruct and analyze dynamic PET data (see a review in [37]), and in a number of cases it has been shown that the time evolution of the tracer activity allows for an improved discrimination between tissues, leading to an improved diagnostic or prognostic accuracy compared to static imaging.

A first category of methods describes the time evolution of the activity by means of a system of differential equations, which model the metabolism of the specific tracer under study. These equations depend on a small set of kinetic parameters (typically 3 or 4), which can be estimated from the time activity curve $\boldsymbol{Y}_{n,t}$ of each voxel n, and convey meaningful information such as metabolic rates for a biochemical reaction or a transport rate. These methods often require the knowledge of the arterial concentration of the tracer as a function of time.

A second category includes generic methods, which model the time evolution in each voxel as a linear combination of $M < T$ basis time-activity vectors. These vectors can be organized as a $M \times T$ matrix $\boldsymbol{X}$, with $\boldsymbol{X}_{m,t}$ the basis time-activity vector m at time t. The dynamic image data are then modeled as

$$\boldsymbol{Y} = \boldsymbol{K}\boldsymbol{X} \tag{21}$$

with a $N \times M$ matrix $\boldsymbol{K}$, with elements $\boldsymbol{K}_{n,m}$ equal to the weight of the time basis function m in voxel n. The problem consists in jointly estimating $\boldsymbol{K}$ and $\boldsymbol{X}$ from the reconstructed dynamic images $\boldsymbol{Y}$. In contrast with the methods based on tracer-specific kinetic models, the estimated time-activity vectors $\boldsymbol{X}$ cannot in general be interpreted in terms of tissue types or metabolic processes but the achieved dimensionality reduction ($M < T$) is useful as a preprocessing step for an automated clustering of the image voxels.

The matrices $\boldsymbol{X}$ and $\boldsymbol{K}$ are determined up to a nonsingular $M \times M$ matrix. Several penalties are well adapted to dynamic PET and improve identifiability and stability. Viewing $\boldsymbol{X}$ as the time evolution of the tracer in M compartments (associated to tissue types or biochemical processes), and $\boldsymbol{K}$ as weighing the fraction of each compartment in a voxel, naturally suggests imposing nonnegativity of all elements of $\boldsymbol{X}$ and $\boldsymbol{K}$. Moreover, penalties can enforce sparsity of $\boldsymbol{K}$ (assuming that most voxels contain few tissue types), smoothness of the time evolution in each compartment, and spatial smoothness of $\boldsymbol{K}$. Assuming Gaussian noise in the reconstructed images $\boldsymbol{Y}$, and a unit covariance matrix, one is led to the same nonnegative matrix factorization problem as in Section 2.2.

Note that similar factorization methods can be applied before rather than after image reconstruction, in which case $\boldsymbol{Y}$ in (21) represents the raw dynamic data measured by the scanner, and a Poisson noise model is appropriate (see [24, 37]). Also, PCA data reduction methods have been used instead of NMF [1].

2.4 Joint estimation of activity and attenuation in time-of-flight PET

This section describes a second problem in positron emission tomography, which also leads to a bi-convex minimization problem similar to those described in Section 2.2. We consider the static case with a single measurement but for a time-of-flight PET scanner. A time-of-flight (TOF) PET scanner collects data in the same way as described in Section 2.3, but in addition the detectors measure the arrival time of the two detected photons emitted by the annihilation of a positron. The difference between the arrival times of the two photons localizes the origin of the annihilation along the line of response (LOR), with an accuracy determined by the timing resolution of the detectors (typically 300 ps). The time difference is histogrammed into T time-of-flight bins. As a result, TOF-PET data are stored in a vector $\boldsymbol{y} = \{\boldsymbol{y}_n, n = 0, \ldots, N-1\}$, where $N = LT$ and $\boldsymbol{y}_{n=\ell T+t}$ is the number of events detected for LOR $\ell = 0, \ldots, L-1$ and time bin $t = 0, \ldots, T-1$. The activity image is parameterized as a vector $\boldsymbol{x} = \{\boldsymbol{x}_m \geq 0, m = 0, \ldots, M-1\}$, $\boldsymbol{x}_m$ being the tracer concentration in voxel m. The goal of PET is to estimate $\boldsymbol{x}$ from the data $\boldsymbol{y}$.

The noise-free data is related to the activity image by

$$\boldsymbol{y} = \boldsymbol{K}\boldsymbol{x} + \boldsymbol{b} \tag{22}$$

with $\boldsymbol{b} \in \mathbb{R}^N$ a known nonnegative background, and a $N \times M$ matrix $\boldsymbol{K}$. The matrix element $\boldsymbol{K}_{n,m}$ is the probability that a pair of photons emitted in voxel m is detected in the LOR $\ell = n/T$ with time-of-flight difference bin $t = n\%T$ (n modulo T). This matrix is modeled as a product $\boldsymbol{K} = \boldsymbol{Q}\boldsymbol{P}$. The $N \times M$ matrix $\boldsymbol{P}$ is the detection probability

calculated assuming that the photons only interact in the detectors. The $N \times N$ diagonal matrix $\boldsymbol{Q}$ models the data attenuation caused by photon interactions within the object. The attenuation is independent of the time difference and, therefore,

$$\boldsymbol{Q}_{n,n'} = \boldsymbol{\delta}_{n,n'} \boldsymbol{q}_{n/T}, \tag{23}$$

with $1 \geq \boldsymbol{q}_\ell > 0$, $\ell = 0, \ldots, L-1$, equal to the probability that a photon pair emitted along LOR ℓ escapes the body without interacting. The matrix $\boldsymbol{P}$ only depends on the scanner and is known, but the attenuation $\boldsymbol{Q}$ depends on the object.

In practice, PET scans are performed with a hybrid PET/CT scanner and the attenuation is estimated by the Beer–Lambert law,

$$\boldsymbol{q}_\ell = \exp\left\{ -\sum_{m=0}^{M-1} \boldsymbol{s}_{\ell,m} \boldsymbol{\mu}_m \right\}, \quad \ell = 0, \ldots, L-1 \tag{24}$$

where $\boldsymbol{s}_{\ell,m} \geq 0$ is the intersection length of LOR ℓ with voxel m, and $\boldsymbol{\mu}_m$, the linear attenuation coefficient in voxel m, is measured using a CT scan of the patient. This section deals with more challenging situations where the CT scan is not available or is unreliable due to patient motion between the PET and CT scans. The matrix $\boldsymbol{Q}$ (or equivalently the vector $\boldsymbol{q} \in \mathbb{R}^L$) is then unknown and must be estimated jointly with the activity $\boldsymbol{x}$.

The data $\boldsymbol{y}$ in PET are modeled as independent Poisson variables with expectation given by (22). Given $\boldsymbol{y}$ and the background $\boldsymbol{b}$, maximum likelihood estimation then leads to the minimization of the generalized Kullback–Leibler divergence

$$F(\boldsymbol{x}, \boldsymbol{q}) = \mathrm{KL}(\boldsymbol{y}, \boldsymbol{QPx} + \boldsymbol{b}), \tag{25}$$

with the diagonal matrix $\boldsymbol{Q}$ defined by (23). We consider here a nonpenalized likelihood, assuming that the data parameterization (voxel size) and noise level (number of detected coincident events $\sum_n \boldsymbol{y}_n$) are sufficient to guarantee a good stability. Note that $F(\boldsymbol{x}, \boldsymbol{q}) = F(a\boldsymbol{x}, \frac{\boldsymbol{q}}{a})$ for any $a > 0$; therefore, the solution is determined at best up to a global scale factor. Up to this scale factor, the solution to the equation $\boldsymbol{y} = \boldsymbol{QPx} + \boldsymbol{b}$ has been shown to be unique for a continuous-continuous model of the joint estimation problem [16] with noise-free data. Unfortunately, no general result on the identifiability of $\boldsymbol{x}$, $\boldsymbol{q}$ is known in the discrete setting considered here. See [4] for a general overview on this joint estimation problem in TOF-PET.

Following [39], we minimize (25) by alternate minimization as in the previous sections, noting that fixing either $\boldsymbol{Q}$ or $\boldsymbol{x}$ reduces the bi-convex problem to KL minimization for a linear inverse problem. The activity and the attenuation updates are therefore given by the Richardson–Lucy formula:

$$\boldsymbol{x}^{(k+1)} = \frac{\boldsymbol{x}^{(k)}}{\boldsymbol{P}^T \boldsymbol{Q}^{(k)} \mathbf{1}_N} \circ \boldsymbol{P}^T \boldsymbol{Q}^{(k)} \frac{\boldsymbol{y}}{\boldsymbol{Q}^{(k)} \boldsymbol{P} \boldsymbol{x}^{(k)} + \boldsymbol{b}} \tag{26}$$

$$\boldsymbol{q}^{(k+1)} = \frac{\boldsymbol{q}^{(k)}}{\boldsymbol{V}\boldsymbol{P}\boldsymbol{x}^{(k+1)}} \circ \boldsymbol{V}\frac{(\boldsymbol{P}\boldsymbol{x}^{(k+1)}) \circ \boldsymbol{y}}{\boldsymbol{Q}^{(k)}\boldsymbol{P}\boldsymbol{x}^{(k+1)} + \boldsymbol{b}} \tag{27}$$

where the $L \times N$ matrix $\boldsymbol{V}$ sums over all time bins for a given LOR: $\boldsymbol{V}_{\ell,n} = \boldsymbol{\delta}_{\ell,n/T}$.

Several remarks are in order:

1. In the absence of background ($\boldsymbol{b} = \boldsymbol{0}$) the attenuation update (27) simplifies and yields a closed-form expression for the minimizer of the convex cost function at fixed activity $\boldsymbol{x}$:

$$\boldsymbol{q}^*(\boldsymbol{x}) = \arg\min_{\boldsymbol{q}} F(\boldsymbol{x}, \boldsymbol{q}) = \frac{\boldsymbol{V}\boldsymbol{y}}{\boldsymbol{V}\boldsymbol{P}\boldsymbol{x}}. \tag{28}$$

   The parameter $\boldsymbol{q}$ can be eliminated by inserting (28) into (25) (with $\boldsymbol{b} = \boldsymbol{0}$). The reduced nonconvex cost function $F(\boldsymbol{x}, \boldsymbol{q}^*(\boldsymbol{x}))$ is then minimized using a majorization–minimization algorithm:

$$\boldsymbol{x}^{(k+1)} = \frac{\boldsymbol{x}^{(k)}}{\boldsymbol{P}^T\boldsymbol{V}^T\frac{\boldsymbol{V}\boldsymbol{y}}{\boldsymbol{V}\boldsymbol{P}\boldsymbol{x}^{(k)}}} \circ \boldsymbol{P}^T\frac{\boldsymbol{y}}{\boldsymbol{P}\boldsymbol{x}^{(k)}}. \tag{29}$$

   Properties of this algorithm are analyzed in [17]. In particular, the KL divergence $F(\boldsymbol{x}^{(k)}, \boldsymbol{q}^{(k)})$ is nonincreasing and converges, the activity estimates $\boldsymbol{x}^{(k)}$ are nonnegative, and the rescaled activity estimates $\tilde{\boldsymbol{x}}^{(k)} = \boldsymbol{x}^{(k)}/\|\boldsymbol{x}^{(k)}\|$ are asymptotically regular. It can also be shown that the bi-convex cost function $F(\boldsymbol{x}, \boldsymbol{q})$ has no local minimum if the data are consistent, that is if there is some nonnegative $\boldsymbol{x}_0$, $\boldsymbol{q}_0$ such that $\boldsymbol{y} = \boldsymbol{Q}_0\boldsymbol{P}\boldsymbol{x}_0$. This property has little impact because measured data are inconsistent and one easily finds small-scale toy examples where several local minima can be identified.
2. The algorithm (26), (27) can be easily adapted to include a quadratic penalty on $\boldsymbol{x}$ and $\boldsymbol{q}$.
3. An alternative approach jointly estimates the attenuation coefficient vector $\boldsymbol{\mu} \in \mathbb{R}^M$ instead of the attenuation factors $\boldsymbol{q} \in \mathbb{R}^L$, using (24). This approach facilitates regularization because prior information on the attenuation coefficient $\boldsymbol{\mu}$ is easily derived knowing a typical material composition of the object; for example, $\boldsymbol{\mu} \sim 0.1\,\mathrm{cm}^{-1}$ for water at 511 keV. In contrast the values of the attenuation factors $\boldsymbol{q}$ are more difficult to define a priori. Joint estimation of $\boldsymbol{\mu}$ and $\boldsymbol{x}$ by alternate minimization has been explored previously for non-TOF PET [13, 34], but in that case the data do not provide sufficient information for a reliable estimation. Thanks to the additional information provided by the time-of-flight measurement, the recent application to TOF-PET [38] leads to a reliable estimation and opens the way to practical implementations.
4. When a CT measurement of the attenuation coefficient $\boldsymbol{\mu}$ is available but corresponds to a different position of the patient than the TOF-PET data, joint estimation has also been proposed to estimate the activity $\boldsymbol{x}$ and a nonrigid geometric warping that maps the CT map $\boldsymbol{\mu}$ onto the deformed frame required for PET [40].

Bibliography

[1] Y. Anzai, S. Minoshima, G. T. Wolf, and R. L. Wahl. Head and neck cancer: detection of recurrence with three-dimensional principal components analysis at dynamic FDG PET. *Radiology*, 212(1):285–290, 1999.

[2] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran. Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka–Lojasiewicz inequality. *Math. Oper. Res.*, 35(2):438–457, 2010.

[3] G. R. Ayers and J. C. Dainty. Iterative blind deconvolution method and its applications. *Opt. Lett.*, 13(7):547–549, 1988.

[4] Y. Berker and Y. Li. Attenuation correction in emission tomography using the emission data—a review. *Med. Phys.*, 43(2):807–832, 2016.

[5] M. Bertero and P. Boccacci. *Introduction to Inverse Problems in Imaging*. Institute of Physics Publishing, 1998.

[6] J. Bolte, P. L. Combettes, and J. C. Pesquet. Alternating proximal algorithm for blind image recovery. In *2010 IEEE International Conference on Image Processing*, pages 1673–1676, 2010.

[7] J. Bolte, S. Sabach, and M. Teboulle. Proximal alternating linearized minimization for nonconvex and nonsmooth problems. *Math. Program.*, 146(1):459–494, 2014.

[8] M. Burger and O. Scherzer. Regularization methods for blind deconvolution and blind source separation problems. *Math. Control Signals Syst.*, 14(4):358–383, 2001.

[9] A. S. Carasso. False characteristic functions and other pathologies in variational blind deconvolution. A method of recovery. *SIAM J. Appl. Math.*, 70(4):1097–1119, 2009.

[10] T. F. Chan and C. K. Wong. Total variation blind deconvolution. *IEEE Trans. Image Process.*, 7(3):370–375, 1998.

[11] T. F. Chan and C. K. Wong. Convergence of the alternating minimization algorithm for blind deconvolution. *Linear Algebra Appl.*, 316(1–3):259–285, 2000.

[12] M. E. Daube-Witherspoon and G. Muehllehner. An iterative image space reconstruction algorithm suitable for volume ECT. *IEEE Trans. Med. Imaging*, 5(2):61–66, 1986.

[13] A. De Pierro and F. Crepaldi. Activity and attenuation recovery from activity data only in emission computed tomography. *Comput. Appl. Math.*, 25(2-3):205–227, 2006.

[14] A. R. De Pierro. On the convergence of the iterative image space reconstruction algorithm for volume ECT. *IEEE Trans. Med. Imaging*, 6(2):174–175, 1987.

[15] A. R. De Pierro. On the relation between the ISRA and the EM algorithm for positron emission tomography. *IEEE Trans. Med. Imaging*, 12(2):328–333, 1993.

[16] M. Defrise, A. Rezaei, and J. Nuyts. Time-of-flight PET data determine the attenuation sinogram up to a constant. *Phys. Med. Biol.*, 57(4):885–899, 2012.

[17] M. Defrise, A. Rezaei, and J. Nuyts. Transmission-less attenuation correction in time-of-flight PET: analysis of a discrete iterative algorithm. *Phys. Med. Biol.*, 59(4):1073–1096, 2014.

[18] D. Donoho and V. Stodden. When does non-negative matrix factorization give a correct decomposition into parts? In S. Thrun, L. K. Saul, and B. Schölkopf, editors, *Advances in Neural Information Processing Systems 16*, pages 1141–1148. MIT Press, 2004.

[19] P. Eggermont. Multiplicative iterative algorithms for convex programming. *Linear Algebra Appl.*, 130:25–42, 1990.

[20] D. A. Fish, A. M. Brinicombe, E. R. Pike, and J. G. Walker. Blind deconvolution by means of the Richardson–Lucy algorithm. *J. Opt. Soc. Am. A*, 12(1):58–65, 1995.

[21] M. Hong, M. Razaviyayn, Z. Q. Luo, and J. S. Pang. A unified algorithmic framework for block-structured optimization involving big data: with applications in machine learning and signal processing. *IEEE Signal Process. Mag.*, 33(1):57–77, 2016.

[22] P. O. Hoyer. Non-negative sparse coding. In *Proceedings of the 12th IEEE Workshop on Neural Networks for Signal Processing*, volume 1, pages 557–565, 2002.
[23] L. Justen and R. Ramlau. A non-iterative regularization approach to blind deconvolution. *Inverse Probl.*, 22(3):771–800, 2006.
[24] E. Krestyannikov, J. Tohka, and U. Ruotsalainen. Joint penalized-likelihood reconstruction of time-activity curves and regions-of-interest from projection data in brain PET. *Phys. Med. Biol.*, 53(11):2877–2896, 2008.
[25] K. Lange, D. R. Hunter, and I. Yang. Optimization transfer using surrogate objective functions. *J. Comput. Graph. Stat.*, 9(1):1–20, 2000.
[26] L. Lecharlier. *Blind inverse imaging with positivity constraints*. PhD Thesis, Université libre de Bruxelles, Faculté des Sciences, 2014.
[27] L. Lecharlier and C. De Mol. Regularized blind deconvolution with Poisson data. *J. Phys. Conf. Ser.*, 464(1):012003, 2013.
[28] D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401:788–791, 1999.
[29] D. D. Lee and H. S. Seung. Algorithms for non-negative matrix factorization. In T. K. Leen, T. G. Dietterich, and V. Tresp, editors, *Advances in Neural Information Processing Systems 13*, pages 556–562. MIT Press, 2001.
[30] L. B. Lucy. An iterative technique for the rectification of observed distributions. *Astron. J.*, 79:745–754, 1974.
[31] W.-K. Ma, J. Bioucas-Dias, T.-H. Chan, N. Gillis, P. Gader, A. Plaza, A. Ambikapathi, and C.-Y. Chi. A signal processing perspective on hyperspectral unmixing: insights from remote sensing. *IEEE Signal Process. Mag.*, 31(1):67–81, 2014.
[32] S. R. Meikle and R. D. Badawi. Quantitative techniques in PET. In D. L. Bailey, D. W. Townsend, P. E. Valk, and M. N. Maisey, editors, *Positron Emission Tomography, Basic Sciences*, pages 93–126. Springer Verlag, 2005.
[33] F. Natterer and F. Wübbeling. *Mathematical Methods in Image Reconstruction*. Society for Industrial and Applied Mathematics, 2001.
[34] J. Nuyts, P. Dupont, S. Stroobants, R. Benninck, L. Mortelmans, and P. Suetens. Simultaneous maximum a posteriori reconstruction of attenuation and activity distributions from emission sinograms. *IEEE Trans. Med. Imaging*, 18(5):393–403, 1999.
[35] A. Ostrowski. *Solution of Equations in Euclidean and Banach Spaces*. Academic Press, 1973.
[36] M. Prato, A. L. Camera, S. Bonettini, and M. Bertero. A convergent blind deconvolution method for post-adaptive-optics astronomical imaging. *Inverse Probl.*, 29(6):065017, 2013.
[37] A. Reader and J. Verhaeghe. 4D image reconstruction for emission tomography. *Phys. Med. Biol.*, 59(22):R371–R418, 2014.
[38] A. Rezaei, M. Defrise, G. Bal, C. Michel, M. Conti, C. Watson, and J. Nuyts. Simultaneous reconstruction of activity and attenuation in time-of-flight PET. *IEEE Trans. Med. Imaging*, 31(12):2224–2233, 2012.
[39] A. Rezaei, M. Defrise, and J. Nuyts. ML-reconstruction for TOF-PET with simultaneous estimation of the attenuation factors. *IEEE Trans. Med. Imaging*, 33(7):1563–1572, 2014.
[40] A. Rezaei, C. Michel, M. E. Casey, and J. Nuyts. Simultaneous reconstruction of the activity image and registration of the CT image in TOF-PET. *Phys. Med. Biol.*, 61(4):1852–1874, 2016.
[41] W. H. Richardson. Bayesian-based iterative method of image restoration. *J. Opt. Soc. Am.*, 62(1):55–59, 1972.
[42] L. A. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Trans. Med. Imaging*, 1(2):113–122, 1982.
[43] Y.-L. You and M. Kaveh. A regularization approach to joint blur identification and image restoration. *IEEE Trans. Image Process.*, 5(3):416–428, 1996.

Simon Gindikin

3 From Radon to Leray

Abstract: We discuss the influence of the Radon transform on multidimensional complex analysis. In the first turn, it is Leray's construction of an universal multidimensional analogue of the Cauchy formula—*the Cauchy–Fantappie formula*. It was extended by fundamental conceptions of analytic functionals, analytic duality, $\bar{\partial}$-cohomology, and different complex versions of the Radon transform. In our exposition, the conception of *the Cauchy–Radon transform* plays an essential role. It is the result of the replacement of δ-function in the integrand of the Radon transform by a Cauchy kernel. It appears as a modification of the Radon transform for which the inversion formula is independent of the evenness of the dimension. This then gives new possibilities to consider analogs of the Radon transform in non-Euclidean situation.

Keywords: Radon transform, Radon–Cauchy transform, inversion formula, hyperfunctions, Cauchy–Fantappie integral formula, Martineau's duality

MSC 2010: 44A12, 46F15, 32A26

Without doubts, tomography is the most spectacular application of the Radon transform. However, there also were a few absolutely fundamental results in pure mathematics, which appeared under the strong influence of the Radon transform. I would emphasize applications to differential equations of F. John, multidimensional Cauchy integral formulas of Leray and the horospherical transform on symmetric homogeneous spaces of Gelfand. In this exposition, we will talk about Leray's results and their development. Our exposition has a narrow focus on complex parallels to the Radon transform. We do not consider more broad possibilities, corresponding to the Radon–John transform for integration along planes of codimension higher than one, such as the Penrose transform.

3.1 The Radon–Cauchy transform

We will make a small, but essential, modification of the usual Radon transform at $\mathbb{R}^n$. Let $\mathbb{R}^n_x$, $\mathbb{R}^n_\xi$ be dual copies with the duality form

$$\langle \xi, x \rangle = \xi_1 x_1 + \cdots + \xi_n x_n$$

Simon Gindikin, Departm. of Math., Hill Center, Rutgers University, 110 Frelinghysen Road, Piscataway, NJ 08854, USA, e-mail: gindikin@math.rutgers.edu

https://doi.org/10.1515/9783110560855-003