Sampling the Multiple Facets of Light

THÈSE N^{O} 8795 (2018)

PRÉSENTÉE LE 19 OCTOBRE 2018

À LA FACULTÉ INFORMATIQUE ET COMMUNICATIONS

LABORATOIRE DE COMMUNICATIONS AUDIOVISUELLES

PROGRAMME DOCTORAL EN INFORMATIQUE ET COMMUNICATIONS

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Gilles BAECHLER

acceptée sur proposition du jury:

Prof. M. C. Gastpar, président du jury
Prof. M. Vetterli, Dr A. J. Scholefield, directeurs de thèse
Prof. J. Lasenby, rapporteuse
Prof. P. Milanfar, rapporteur
Prof. S. Süsstrunk, rapporteuse

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE
Suisse
2018

À mes parents

Acknowledgments

Physics isn't the most important thing. Love is.

Letter to Marcus Chown's mother
Richard Feynman

First and foremost, I have been very fortunate to work under the supervision of two fantastic and complementary PhD advisors. Martin, thanks for offering me (twice!) the opportunity to work in your laboratory. You are not only a researcher who has nothing left to prove to the scientific world, but also a great person with a big heart. I will also always be amazed by your instinctive long term vision and intuition that guided my steps countless times. Adam, thank you for your availability: no matter how busy you were, you could always find time to listen to me—even though you are literally carrying the whole lab on your shoulders since Martin took over his 'side job'! Throughout my PhD, you have always been an infinite source of knowledge, advice, support and help. With you, I not only had a great advisor, but I also found a friend with whom I could share my passion for mountains and photography. To both of you, many thanks for your trust and guidance. It has been a pleasure working and collaborating with you; I learned so much from you!

It was a great honor to have Professors Peyman Milanfar, Joan Lasenby, and Sabine Süsstrunk on my thesis committee. Thank you for reading my manuscript and for your thorough feedback that helped me refine the dissertation.

Many people think of a PhD as individual work. From my experience, I have been blessed to work and collaborate with lots of great and smart people, without whom this thesis would not exist. For example, the pluridisciplinarity of the Lippmann project involved several actors from

various fields. Thanks to my team at LCAV: Arnaud's skills in experimental physics and attention to details, combined with Michalina's strong mathematical knowledge have been instrumental to the success of this project, which was masterfully orchestrated by Adam. The realization of our own Lippmann plates would not have been possible without the help and contributions of the following people: Filipe Alves, who was the first to show us that Lippmann photography can be made practical and who shared some of his secrets on how to realize beautiful plates; Yann Pierson, who assisted us with our chemistry shopping list and lent us his chemistry laboratory; and the GR-CEL group and in particular Prof. De Alencastro and Sylvain Coudret, who let us use their chemistry laboratory and assisted us with our chemistry experiments. I am also grateful to Prof. Jean-Marc Fournier for sharing his infinite knowledge of the Lippmann process and thoroughly proofreading the Lippmann chapter, and to Yves Gentet, for his advice on the development and making of holographic plates. Finally, I would also like to thank Professors Yves Bellouard, Philippe Renaud, and Jacques Moser, who took the time to explain me numerous concepts that were beyond the scope of my knowledge, ranging from femtosecond lasers to the chemistry of photographic development.

Another effort that implicated numerous actors was the construction of our light domes, which have been the result of the contribution from the following people: Gaël Soudan and Vincent Namy for the early prototypes; Marc Jeanneret and Laurent Chevalley from the mechanical workshop for their work at improving the structure of the device; Adrien Hoffet for his support and help with the electronics; and Martial and Benjamin Baechler for sharing their knowledge in mechanics and helping me construct a few useful parts. I would also like to express my gratitude to Loïc Baboulaz and the startup Artmyn for their collaboration and for lending us one of their light domes for a few experiments.

Other very fruitful and memorable collaborations include the following. On the phase retrieval project, thanks to Juri and Miranda: it was a pleasure and a very rewarding experience working with the two of you. On the micropolarizer array project, I am thankful to Laurent, for his dedication to work and his versatility.

Hanjie, my office mate, it has been great working next to you. Thank you for all your mathematical advices, particularly on finite rate of innovation. Thanks to my present and former colleagues from the Audiovisual Communication laboratory: Alireza, Benjamin, Dalia, Eric, Frederike, Ivan, Karen, Luc, Marta, Mathieu, Mihailo, Mitra, Nicoletta, Niranjan, Paolo, Reza, Robin, Runwei, Zhou, ... and the many other brilliant people who crossed my road at EPFL and that I might have omitted here. I believe that the combination of all your different backgrounds, cultures, passions and mindsets contributed to make our lab a great and unique place to be. Thank you all for the lunches and 'coffee' breaks, for the conference dinners and parties, for the skiing and hiking week-ends, and for the fun we had together.

Thanks to my friends, who allowed me to sometimes disconnect from the academic world. I am especially grateful to Xavier, for the organization of countless outings and races in the mountains, which were literally breaths of fresh air during my work at EPFL. My special thoughts go to Stéphane, who tragically left us in the mountains 10 days before the final submission of this manuscript. May you rest in peace in your playground.

Last but not least, I would like to thank my extraordinary family for their love, care and support. In particular, I am infinitely grateful to my parents, Jean-Paul and Thérèse, who always encouraged me and supported all my career and life choices.

Abstract

The theme of this thesis revolves around three important manifestations of light, namely its corpuscular, wave and electromagnetic nature. Our goal is to exploit these principles to analyze, design and build imaging modalities by developing new signal processing and algorithmic tools, based in particular on sampling and sparsity concepts.

First, we introduce a new sampling scheme called *variable pulse width*, which is based on the *finite rate of innovation* (FRI) sampling paradigm. This new framework enables us to sample and perfectly reconstruct weighted sums of Lorentzians; perfect reconstruction from sampled signals is guaranteed by a set of theorems. Weighted sums of Lorentzians provide an interesting parametric model to describe signals that can be naturally represented by a sum of pulses. For instance, extensive experiments have shown that it is a particularly suitable model to tackle the problem of electrocardiogram compression.

Second, we return to the context of light and study its reflection, which is based on the corpuscular model of light. More precisely, we propose to use our FRI-based model to represent bidirectional reflectance distribution functions. Such functions are naturally defined on the sphere, which is also a motivation to generalize our sampling framework to the spherical domain. We develop dedicated *light domes* to acquire reflectance functions and use the measurements obtained to demonstrate the usefulness and versatility of our model. In particular, we concentrate on the representation of specularities, which are sharp and bright components generated by the direct reflection of light on surfaces.

Third, we explore the wave nature of light through Lippmann photography, a century-old photography technique that acquires the entire spectrum of visible light via the phenomenon of *interference*. This fascinating process captures interferences patterns created by the exposed scene inside the depth of a photosensitive plate. Following development, the colors of the original scene can be reproduced by illuminating the plate with a neutral light source. After interacting inside the plate, the illuminating light is reflected back and its spectrum corresponds to that of the exposed scene. Based on the original theory by Gabriel Lippmann, who won the 1908 Nobel Prize for this invention, we propose a mathematical model which precisely explains the technique. More importantly, we demonstrate that the spectrum reproduction suffers from a number of distortions due to the finite thickness of the plate and the choice of reflector. In addition to describing these artifacts, we propose an algorithm to invert them, essentially recovering the original spectrum of the exposed scene.

Next, the wave nature of light is further generalized to the electromagnetic theory, which we invoke to leverage the concept of polarization of light. We also return to the topic of the representation of reflectance functions and focus this time on the separation of the specular component from the other reflections. We exploit the fact that the polarization of light is preserved in specular reflections and investigate camera designs with polarizing micro-filters with different orientations placed just in front of the camera sensor; the different polarizations of the filters create a mosaic image, from which we propose to extract the specular component. We

apply our demosaicing method to several scenes and additionally demonstrate that our approach improves photometric stereo.

Finally, we delve into the problem of retrieving the phase information of a signal from the magnitude of its Fourier transform. We assume the original signal to be sparse; this is a natural assumption in many applications, such as X-ray crystallography, speckle imaging and blind channel estimation. We propose an algorithm that resolves the phase retrieval problem for sparse signals in three stages: first, we exploit the finite rate of innovation sampling theory to super-resolve the auto-correlation function of a sparse signal from a limited number of samples; second, we design a greedy algorithm that identifies the locations of a sparse solution given the super-resolved auto-correlation function; third, we recover the amplitudes of the atoms given their locations and the measured auto-correlation function. Unlike traditional approaches that recover a discrete approximation of the underlying signal, our algorithm estimates the signal on a continuous domain, which makes it the first of its kind.

The concluding chapter outlines several avenues for future research, like new optical devices such as displays and digital cameras, inspired by the topic of Lippmann photography.

Keywords: Color, electrocardiogram, finite rate of innovation, Hilbert transform, hyperspectral imaging, interference, inverse problem, Lippmann, phase retrieval, photography, polarizing filter, pulse compression methods, reflectance function, sampling, signal denoising, sparsity, spectral analysis, specularity, spherical harmonics, standing wave, variable width, X-ray crystallography.

Résumé

Le thème de cette thèse s'articule autour de trois manifestations importantes de la lumière, à savoir sa nature corpusculaire, ondulatoire et électromagnétique. Notre objectif est d'exploiter ces principes pour analyser, concevoir et construire des modalités d'imagerie en développant de nouveaux outils de traitement du signal et des algorithmes, basés notamment sur les concepts d'échantillonnage et de parcimonie.

Premièrement, nous introduisons un nouveau schéma d'échantillonnage appelé *variable pulse width*, qui est basé sur le paradigme de l'*échantillonnage à taux d'innovation fini* (FRI, pour *finite rate of innovation*). Ce nouveau concept permet d'échantillonner et de reconstruire parfaitement des sommes pondérées de Lorentziens ; la reconstruction parfaite à partir de signaux échantillonnés est garantie par un ensemble de théorèmes. Les sommes pondérées de Lorentziens présentent un modèle paramétrique intéressant pour décrire des signaux qui peuvent être naturellement représentés par une somme d'impulsions. Par exemple, nombre d'expériences ont montré qu'il s'agit d'un modèle particulièrement approprié pour s'attaquer au problème de la compression de signaux cardiaques.

Deuxièmement, nous revenons sur le thème de la lumière et étudions le principe de réflexion, basé sur le modèle corpusculaire de la lumière. Plus précisément, nous proposons d'utiliser notre modèle d'échantillonage basé sur la théorie FRI pour représenter les fonctions de distribution de la réflectance bidirectionnelle. De telles fonctions sont naturellement définies sur un domaine sphérique, ce qui est aussi une motivation pour généraliser notre technique d'échantillonnage aux signaux définis sur la sphère. Nous développons des *dômes de lumière* dédiés pour acquérir des fonctions de réflectance et utilisons les données obtenues pour démontrer l'utilité et la polyvalence de notre modèle. En particulier, nous nous concentrons sur la représentation des spécularités, qui sont des composantes effilées et brillantes générées par la réflexion directe de la lumière sur les surfaces.

Troisièmement, nous explorons la nature ondulatoire de la lumière grâce à la photographie Lippmann, une technique de photographie plus que centenaire qui capte l'ensemble du spectre de la lumière visible par le biais du phénomène d'interférence. Ce processus fascinant piège des motifs d'interférences créés par la scène exposée à l'intérieur de la profondeur d'une plaque photosensible. Après développement, les couleurs de la scène originale peuvent être reproduites en éclairant la plaque avec une source de lumière neutre. Suite à son interaction à l'intérieur de la plaque, la lumière est réfléchie et son spectre correspond à celui de la scène exposée. Basé sur la théorie originale de Gabriel Lippmann, qui a remporté le prix Nobel en 1908 pour cette invention, nous proposons un modèle mathématique qui décrit précisément le procédé. Plus important encore, nous démontrons que la reproduction du spectre souffre de plusieurs distorsions dues à l'épaisseur finie de la plaque et au choix du réflecteur. En plus de décrire ces artefacts, nous proposons également un algorithme pour les inverser, retrouvant ainsi le spectre original de la scène exposée.

Ensuite, la nature ondulatoire de la lumière est généralisée à la théorie électromagnétique,

que nous invoquons pour tirer parti du concept de polarisation de la lumière. Nous revenons également sur le sujet de la représentation des fonctions de réflectance et nous focalisons cette fois sur la séparation de la composante spéculaire des autres réflexions. Nous exploitons le fait que la polarisation de la lumière est préservée lors de réflexions spéculaires et étudions la conception de caméras avec des microfiltres polarisants ayant différentes orientations placées juste devant le capteur de la caméra ; les différentes polarisations des filtres créent une image en mosaïque, à partir de laquelle nous proposons d'extraire la composante spéculaire. Nous appliquons notre méthode de dématriçage à plusieurs scènes et démontrons entre autres que notre approche améliore le procédé stéréo photométrique.

Finalement, nous nous penchons sur le problème de l'extraction de l'information de phase d'un signal à partir de la magnitude de sa transformée de Fourier. Nous supposons que le signal mesuré est parcimonieux ; c'est une hypothèse réaliste dans de nombreuses applications, telles que la cristallographie aux rayons X, l'imagerie à granularité et l'estimation aveugle de canal. Nous proposons un algorithme qui résout le problème d'extraction de phase pour les signaux parcimonieux en trois étapes : premièrement, nous exploitons la théorie de l'échantillonnage à taux d'innovation fini pour améliorer la résolution de la fonction d'auto-corrélation du signal mesuré à partir d'un nombre limité d'échantillons ; deuxièmement, nous estimons les positions d'une solution parcimonieuse compte tenu de la fonction d'auto-corrélation à super résolution ; troisièmement, nous calculons les amplitudes des atomes à partir de leur position et de la fonction d'auto-corrélation mesurée. Contrairement aux approches traditionnelles qui fournissent une approximation discrète du signal sous-jacent, notre algorithme estime le signal sur un domaine continu, ce qui en fait le premier du genre.

Le chapitre de conclusion évoque plusieurs pistes pour de futures recherches, comme de nouveaux dispositifs optiques tels que des écrans et des appareils photo numériques, inspirés par le sujet de la photographie Lippmann.

Mots-Clés : Couleur, électrocardiogramme, taux d'innovation fini, transformée de Hilbert, imagerie hyperspectrale, interférence, problème inverse, Lippmann, extraction de phase, photographie, filtre polarisant, méthodes de compression d'impulsions, fonction de réflectance, échantillonnage, débruitage, parcimonie, analyse spectrale, spécularité, harmoniques sphériques, ondes stationnaires, largeur variable, cristallographie aux rayons X.

Contents

Acknowledgments v

Abstract vii

Résumé ix

Abbreviations and Notation xv

1 Introduction 1
- 1.1 A photographic perspective on light . . . 1
- 1.2 Understanding light . . . 3
  - 1.2.1 Corpuscular theory . . . 3
  - 1.2.2 Wave theory . . . 6
  - 1.2.3 Electromagnetic theory . . . 11
  - 1.2.4 Wave-particle duality . . . 13
- 1.3 Thesis outline and main contributions . . . 14

2 Sampling Diracs and Non-Bandlimited Shapes 19
- 2.1 Introduction . . . 19
- 2.2 Background . . . 21
  - 2.2.1 Signals with finite rate of innovation . . . 21
  - 2.2.2 From samples to Fourier series coefficients . . . 22
  - 2.2.3 From Fourier series coefficients to signal parameters . . . 22
- 2.3 Pulses with variable width: The 1D case . . . 24
  - 2.3.1 Variable width . . . 25
  - 2.3.2 Asymmetry . . . 25
  - 2.3.3 Recovery of VPW pulse parameters . . . 26
  - 2.3.4 Sampling kernels . . . 27
  - 2.3.5 Denoising techniques . . . 28
- 2.4 Pulses with variable width: The 2D case . . . 30
  - 2.4.1 Estimation of the parameters . . . 31
  - 2.4.2 Radial formulation . . . 33
  - 2.4.3 Anisotropic pulses . . . 33
- 2.5 Pulses with variable width: On the sphere . . . 34
  - 2.5.1 Preliminaries . . . 34
  - 2.5.2 Sampling the hemisphere . . . 34

2.5.3 From samples to spherical harmonic coefficients . . . 35
2.5.4 Retrieving Diracs from sectoral SH coefficients . . . 36
2.5.5 Retrieving VPW pulses from sectoral SH coefficients . . . 36
2.6 Example: ECG compression . . . 38
2.7 Conclusion . . . 38
2.A Proof of Lemma 2.1 . . . 40
2.B Time-domain formulae for VPW pulse . . . 40
2.C Computation of the Cramer-Rao bound for VPW signals . . . 41

3 Corpuscular Theory: Acquisition of Reflectance Functions 43
3.1 Introduction . . . 43
3.2 Background . . . 44
3.2.1 The rendering equation . . . 46
3.2.2 BRDF acquisition . . . 47
3.2.3 Light domes . . . 47
3.2.4 BRDF shape and models . . . 48
3.3 FRI with unknown sampling kernel . . . 49
3.3.1 Noisy case . . . 51
3.3.2 Application to specularity estimation . . . 52
3.4 VPW modeling of the BRDF . . . 54
3.4.1 Modeling glossy materials . . . 54
3.4.2 Specularity localization on the plane . . . 55
3.4.3 Specularity localization on the sphere . . . 55
3.4.4 Practical considerations . . . 56
3.5 Conclusion . . . 57

4 Wave Theory and Interferences: Lippmann Photography 59
4.1 Introduction . . . 59
4.2 Analysis: Recording a Lippmann plate . . . 63
4.2.1 Standing waves . . . 63
4.2.2 Development . . . 65
4.3 Synthesis: Viewing a Lippmann plate . . . 66
4.3.1 Analytical model . . . 66
4.3.2 Wave-transfer matrix model . . . 69
4.4 The 3D case: Oblique illumination and viewing angles . . . 71
4.4.1 Recording . . . 71
4.4.2 Viewing . . . 73
4.4.3 Experiment . . . 74
4.5 Color and spectrum reproduction . . . 75
4.5.1 Thickness of the plate . . . 77
4.5.2 Processing and windowing effect . . . 79
4.5.3 Dependance on the reflection coefficient . . . 80
4.6 An alternative: A refractive index based model . . . 85
4.6.1 Comparison with the reflection-based model . . . 87
4.7 Can we invert the Lippmann operations? . . . 87
4.7.1 Recording . . . 88
4.7.2 Viewing . . . 88

4.8 Conclusion . . . 94
4.A Practical aspects . . . 97
4.A.1 Lippmann plates . . . 97
4.A.2 Hurter and Driffield curve . . . 98
4.A.3 Development and processing . . . 98
4.A.4 Thickness of plate and choice of development technique . . . 100
4.A.5 Modern Lippmann photography . . . 101

5 Electromagnetic Theory: Micro-Polarizer Arrays 103
5.1 Introduction . . . 103
5.1.1 Related work . . . 104
5.1.2 Proposed approach . . . 105
5.2 Problem statement . . . 105
5.3 Algorithms . . . 106
5.3.1 Minimizing the ℓ^2 TV norm . . . 107
5.3.2 Minimizing the ℓ^1 TV norm . . . 107
5.3.3 Minimizing the Huber TV norm . . . 108
5.3.4 Unknown light polarization . . . 108
5.4 Practical considerations . . . 109
5.4.1 One-step vs. two-step algorithm . . . 109
5.4.2 Filter design patterns and number of orientations . . . 110
5.4.3 Cost of introducing the filter array . . . 110
5.5 Experiments . . . 111
5.5.1 Single image diffuse and specular separation . . . 111
5.5.2 Multiple image photometric stereo . . . 114
5.6 Conclusion . . . 115

6 Super-Resolution Phase Retrieval 117
6.1 Introduction . . . 117
6.1.1 Previous work . . . 118
6.1.2 Main contributions and outline . . . 120
6.2 Problem statement . . . 120
6.3 A three-stage approach . . . 121
6.4 Algorithms . . . 123
6.4.1 ACF super resolution . . . 123
6.4.2 Support recovery . . . 124
6.4.3 Amplitude recovery . . . 125
6.5 Complexity analysis . . . 127
6.6 Performance analysis . . . 128
6.6.1 Expected performance . . . 128
6.6.2 Probability of success . . . 128
6.6.3 Numerical simulations . . . 132
6.7 Improving noise resilience . . . 133
6.7.1 Deleting solutions from the set of differences . . . 133
6.7.2 Symmetric cost function . . . 133
6.7.3 Denoising of partial solutions . . . 133
6.7.4 Comparison of improvement strategies . . . 135

6.8 Influence of the point locations 135
6.9 Comparison with Charge Flipping 137
6.9.1 Charge Flipping 138
6.9.2 Experimental setup and results 138
6.10 Conclusion 139

Conclusion **141**

Bibliography **147**

Photographic Credits **161**

Curriculum Vitæ **163**

Abbreviations and Notation

Abbreviations

Abbreviation	Description
ACF	Autocorrelation function
AWGN	Additive white Gaussian noise
BRDF	Bidirectional reflectance distribution function
BTF	Bidirectional texture function
CIE	Commission internationale de l'éclairage (international commission on illumination)
CF	Charge flipping
CFA	Color filter array
CRB	Cramér-Rao bound
CRT	Cathode ray tube
DC	Direct current
DFT	Discrete Fourier transform
DSLR	Digital single-lens reflex camera
ECG	Electrocardiogram
ESPRIT	Estimation of signal parameters via rotational invariant techniques
FRI	Finite rate of innovation
FS	Fourier series
FT	Fourier transform
GESPAR	Greedy sparse phase retrieval
HSI	Hue, saturation, intensity
IMOD	Interferometric modulator
IQML	Iterative quadratic maximum likelihood
MERL	Mitsubishi Electric research laboratories
MSE	Mean squared error
MIT-BIH	Massachusetts Institute of Technology - Beth Israel hospital
PR	Phase retrieval
PSNR	Peak signal-to-noise ratio
RGB	Red, green, blue
SLM	Spatial light modulator

sRGB	Standard red, green, blue
SVBRDF	Spatially-varying bidirectional reflectance distribution function
SH	Spherical harmonics
SNR	Signal-to-noise ratio
SRR	Signal-to-residual ratio
TSPR	Two-stage sparse phase retrieval
TV	Total variation
VPW	Variable pulse width

List of symbols

Description	Symbol	Formula/details
discrete indexing	$[\cdot]$	or alternatively, subscript
continuous indexing	$(\cdot)$	
scalars	$a, b, c, \ldots$	
constant scalars	$A, B, C, \ldots$	
vectors	$\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \ldots$	
matrices	$\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \ldots$	
sets	$\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots$	
Dirac delta function	$\delta(t)$	$\int_{-\infty}^{\infty} x(t)\delta(t)dt = x(0)$ (for x continuous at 0)
ℓ^2 norm	$\|\cdot\|$	$(\|x_0\|^2 + \cdots + \|x_{N-1}\|^2)^{1/2}$
ℓ^p norm	$\|\cdot\|_p$	$(\|x_0\|^p + \cdots + \|x_{N-1}\|^p)^{1/p}$
total variation (TV) norm	$\|\cdot\|_{TV}$	
imaginary unit	j	$j^2 = -1$
set of natural numbers	$\mathbb{N}$	$0, 1, 2, \ldots$
set of integers	$\mathbb{Z}$	$\ldots, -1, 0, 1, \ldots$
set of real numbers	$\mathbb{R}$	$(-\infty, \infty)$
set of complex numbers	$\mathbb{C}$	$a + jb$ with $a, b \in \mathbb{R}$
unit sphere	$\mathbf{S}^2$	$\{\boldsymbol{x} \in \mathbb{R}^3 : \|\boldsymbol{x}\| = 1\}$
unit hemisphere	Ω	$\{\boldsymbol{x} \in \mathbb{R}^3 : \|\boldsymbol{x}\| = 1, x_2 \geq 0\}$
cardinal sine	$\mathrm{sinc}(x)$	$\frac{\sin x}{x}$
real part	$\mathfrak{Re}\{\cdot\}$	
imaginary part	$\mathfrak{Im}\{\cdot\}$	
sign function	$\mathrm{sgn}(x)$	-1 for $x < 0, 0$ for $x = 0, 1$ for $x > 0$
matrix transpose	$\cdot^\top$	
conjugate transpose	$\cdot^*$	
Moore-Penrose pseudoinverse	$\cdot^\dagger$	
phase or argument	$\angle$	
absolute value or magnitude	$\|\cdot\|$	
floor function	$\lfloor\cdot\rfloor$	
ceiling function	$\lceil\cdot\rceil$	
inner product	$\langle\cdot\,,\cdot\rangle$	$\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \boldsymbol{x}^*\boldsymbol{y}$
convolution	$*$	
gradient	∇	$\nabla f = \left(\frac{\partial f}{\partial x_0}, \ldots, \frac{\partial f}{\partial x_{N-1}}\right)$

divergence	$\nabla\cdot$	$\nabla \cdot \boldsymbol{f} = \frac{\partial f_0}{\partial x_0} + \cdots + \frac{\partial f_{N-1}}{\partial x_{N-1}}$
Laplacian	∇^2	$\nabla \cdot \nabla$
expected value	$\mathbb{E}[\cdot]$	
big O notation	$\mathcal{O}(\cdot)$	
Legendre polynomial of degree ℓ	P_ℓ	
associated Legendre polynomial	$P_{\ell,m}$	
Fourier transform	$\mathcal{F}\{x\}(\omega)$	$\int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt$
inverse Fourier transform	$\mathcal{F}^{-1}\{X\}(t)$	$\frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)e^{-j\omega t}d\omega$
Fourier cosine transform	$\mathcal{F}_c\{x\}(\omega)$	$\int_{-\infty}^{\infty} x(t)\cos(\omega t)dt$
Hilbert transform	$\mathcal{H}\{x\}(t)$	$\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau}d\tau$

Chapter 1

Introduction

> Newton thought that light was made up of particles, but then it was discovered that it behaves like a wave. Later, however (in the beginning of the twentieth century), it was found that light did indeed sometimes behave like a particle. [...] Now we have given up. We say: "It is like neither."
>
> *The Feynman Lectures on Physics*
> Richard Feynman

1.1 A photographic perspective on light

Since the dawn of civilization, people have tried to understand and explain light. In ancient Greece for example, Pythagoras believed that vision was due to light rays emitted from the eye and illuminating objects. Even though light seems a very familiar concept, it took several centuries for physicists and scientists to establish sound theories that fully characterize and explain the principles and laws that govern the propagation of light and its interaction with matter. In fact, a number of theories can explain different phenomena.

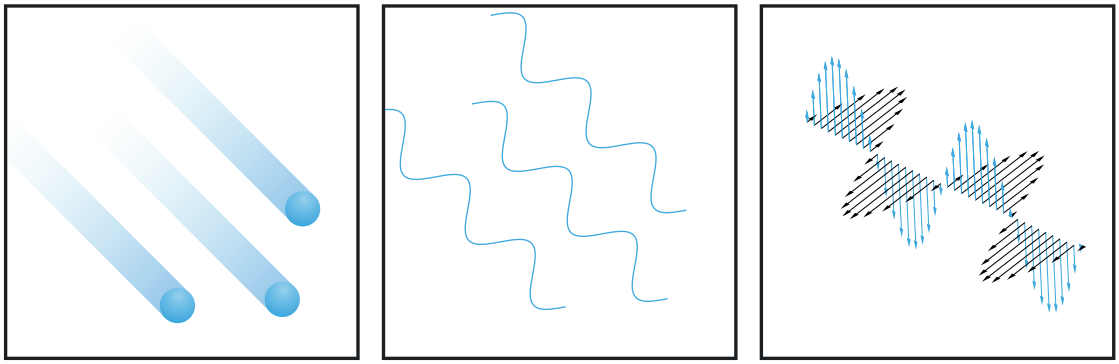
Light plays a central role in this thesis. Before diving into the theoretical details, we briefly review its main properties from a photographic perspective: to do so, let us have a closer look at Figure 1.1, which summarizes some of the effects and interactions that light can have in a scene. In vacuum, light propagates in straight lines. However, when traveling in different media or interacting with objects, light can be subject to various optical phenomena. The most obvious effect in Figure 1.1 is the *reflection* of light, embodied by the mountain that appears inverted

Figure 1.1: Mount Rundle and Vermilion Lakes near Banff, AB in Canada. This photograph is a composite of images that have been taken with a circular polarizing filter having two different orientations. The three theories of light can be observed in this single picture: the corpuscular theory can describe among other things the reflection of light on the objects of the scene. The wave theory is highlighted by the soap bubbles, whose thin layer creates interferences. The electromagnetic nature of light is embodied by the polarizing filter that almost fully suppresses the reflection from the surface of the lake on the left side of the image.

in the lake of the foreground. More generally, all the light reaching the camera has undergone some reflection of some sort as there are no light sources directly visible in the scene: the sun for instance is already hidden behind the surrounding mountains. A facet of light that usually goes hand in hand with reflection is *absorption*. Light is rarely fully reflected by materials; instead, a certain amount ends up being absorbed by it. This is the reason why the trees appear much darker than the snow in the scene: the trees absorb some of the light and use it to create energy via photosynthesis. On the other hand, almost all the light that hits snow patches is reflected.

Refraction is another important aspect of light: when light traverses different media, it can undergo changes of direction if the indices of refraction of these media are different. It can be observed for instance by looking through a glass of water. In Figure 1.1, refraction occurs at the interface between air and the lake.

Another optical phenomenon of interest is *interference*, which relies on the wave nature of light and occurs when several waves interact with each other. Water interferences can be created by throwing two stones into a still lake. The resulting waves either add up constructively or destructively, producing peaks and valleys on the surface of the water. The same happens with



(a) Corpuscular theory (b) Wave theory (c) Electromagnetic theory

Figure 1.2: Illustration of the three theories regarding the nature of light used in this thesis.

light, the difference is that light waves have extremely short wavelengths, which makes them more complicated to observe. A relatively easy way to indirectly visualize interference is with soap bubbles, which are made of a very thin film of soap. The light reflects back and forth at the interface between the soap and air and interferes with itself; this results in intriguing colors that depend on the thickness of the film, as showcased in Figure 1.1.

Lastly, we are interested in *polarization*. This phenomenon is based on the intuition that light traveling through space has an orientation. Using a polarizing filter, light can be filtered and selected according to its different orientations. For example, the photograph in Figure 1.1 has been realized with a polarizing filter in front of the camera. The difference between the left-hand side and the right-hand side is that the filter has been rotated by 90 degrees. As a result, the reflection from the surface of the lake is either preserved or diminished.

There are more properties that are not covered in this dissertation: for instance, the light can be *scattered* by objects; this is what happens in the sky and the clouds.

1.2 Understanding light

Now that we are familiar with the main properties of light, let us dive deeper into the mathematics behind them. We begin our journey by reviewing the main theories that model the different aspects of light; we explain here a few essential ideas from each of these theories and provide the necessary background for the later chapters. In order to illustrate these concepts, we pick a familiar light source—our sun—that will follow us throughout this introductory chapter.

1.2.1 Corpuscular theory

Chronologically, the first rigorous theory of light dates back to the 16th century and is directly influenced by the *mechanical philosophy*, a scientific movement that aimed at explaining the world with mechanical and deterministic principles. The first mathematical model was introduced by Descartes, who proposed the *corpuscular theory* in his famous *Discours de la Méthode* [50]. The corpuscular theory postulates that light is made of tiny particles that travel in straight lines and interact with the objects of a scene the same way a ball would interact when it hits other

Quantity	Expression	Unit
Flux	$\Phi(A)$	$\mathrm{J/s} = \mathrm{W}$
Irradiance	$E(\boldsymbol{x}) = \frac{d\Phi(A)}{dA(\boldsymbol{x})}$	$\mathrm{W/m^2}$
Radiosity	$B(\boldsymbol{x}) = \frac{d\Phi(A)}{dA(\boldsymbol{x})}$	$\mathrm{W/m^2}$
Intensity	$I(\boldsymbol{\omega}) = \frac{d\Phi}{d\boldsymbol{\omega}}$	W/sr
Radiance	$L(\boldsymbol{x}, \boldsymbol{\omega}) = \frac{d^2\Phi(A)}{\cos\theta dA(\boldsymbol{x}) d\boldsymbol{\omega}}$	$\mathrm{W/m^2 sr}$

Table 1.1: Basic radiometric quantities

objects. The theory can explain physical phenomena such as reflection and refraction and was the predominant doctrine in the seventeenth century, being notably supported by Newton [132].

In the corpuscular model, if we know the geometry of a scene and the position and intensity of the light sources, we can precisely compute how light is scattered in space by simulating a very large number of light particles bouncing in the environment. This is exactly what is done in rendering engines, which employ algorithms such as *ray tracing* to compute synthetic images [146]. To do so, we also need to precisely measure light; this is the science of *radiometry*.

Radiometry

Radiometry is of interest here as it allows us to precisely quantify how light is transported, reflected and scattered in a scene. In this section, we assume that light consists of particles—photons—with position $\boldsymbol{x}$ and direction of motion $\boldsymbol{\omega}$[1]; unless specified otherwise, $\boldsymbol{\omega}$ is measured with respect to the surface normal of the object hit by the photon. We briefly review below a few important radiometric quantities. These quantities are also summarized in Table 1.1.

Flux The *radiant flux* or *radiant power* $\Phi(A)$ is the total amount of energy passing through a surface or space A per unit time and its unit of measurement is the Watt (W) or Joule per second (J/s). Note that the term 'flux' is not only reserved to radiometry and is used to describe various other physical quantities such as the *volumetric flux* in fluid dynamics. In fact, we can make the analogy between photons passing through a surface and water flowing in a pipe.

The flux of the sun is the total energy carried by the photons it emits every second and is called the *solar luminosity*. Its value is estimated at $3.8 \cdot 10^{26}$ W [199].

Radiant intensity Given a pointwise light source, we can compute its *radiant intensity* $I(\boldsymbol{\omega})$, which is measured on the unit sphere and corresponds to the angular density of its emitted power [146]. For instance, computed over the entire unit sphere, the radiant intensity is simply $\Phi/4\pi$. More generally, we can calculate it in terms of flux along any differential cone of directions as

$$I(\boldsymbol{\omega}) = \lim_{\boldsymbol{\omega} \to 0} \frac{\Delta\Phi}{\Delta\boldsymbol{\omega}} = \frac{d\Phi}{d\boldsymbol{\omega}}.$$

The quantity $d\boldsymbol{\omega}$ is called a *solid angle* and is the 3-dimensional generalization of a planar angle; it is defined as a surface on the unit sphere and its unit is the *steradian* (sr). Conversely, the

[1] As it is standard notation in computer graphics, we use $\boldsymbol{\omega}$ for this quantity. This should not be confused with ω (non-bold), which represents the angular frequency in the context of wave optics.

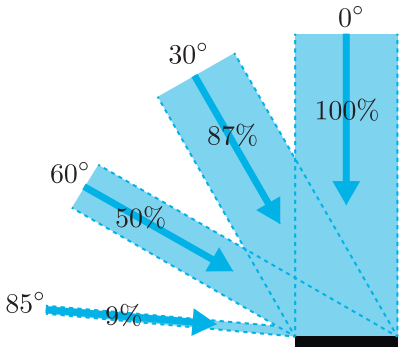


Figure 1.3: Lambert's cosine law: $E_\theta = E\cos\theta$ for different angles θ.

flux is connected to the radiant intensity via

$$\Phi = \int_{\mathbf{S}^2} I(\boldsymbol{\omega}) d\boldsymbol{\omega},$$

where $\mathbf{S}^2 = \{\boldsymbol{x} \in \mathbb{R}^3 : \|\boldsymbol{x}\| = 1\}$ is the unit sphere. Assuming our sun is an isotropic source—it emits equally in all directions—its radiant intensity is $\frac{\Phi}{4\pi}$ W/sr.

Irradiance The *irradiance* $E(\boldsymbol{x})$ describes the flux density—flux per surface area—*received* at point $\boldsymbol{x}$. Its is expressed in Watt per square meter (W/m^2) as a function of flux as

$$E(\boldsymbol{x}) = \frac{d\Phi(A)}{dA(\boldsymbol{x})}.$$

The irradiance follows the *inverse square law*, which states that the irradiance of a point source at distance d is inversely proportional to the square of d. Another relevant property of the irradiance is the *Lambert's cosine law*: the irradiance hitting any surface varies as the cosine of its incident angle θ. It directly follows from the fact that the flux is spread over an area proportional to $\cos\theta$; hence the larger the angle, the larger the area and the smaller the resulting irradiance (see Figure 1.3). The analogue of the irradiance for outgoing light is called *radiosity*.

The total solar irradiance incident on the upper atmosphere of the earth represents the solar flux that reaches our planet. In this context, it is called the *solar constant* and its value is about 1.36 kW/m^2 [95].

Radiance The radiance denotes the irradiance per solid angle (or equivalently the radiant intensity per unit area). It is the most comprehensive of the radiometric quantities that we expose here and it is given in Watts per steradian per square meter (W/sr/m^2). It can be expressed using the flux as

$$L(\boldsymbol{x}, \boldsymbol{\omega}) = \frac{d^2\Phi(A)}{\cos\theta dA(\boldsymbol{x}) d\boldsymbol{\omega}},$$

where $\cos\theta dA(\boldsymbol{x})$ represents the area projected onto the plane that is perpendicular to the direction of propagation. Other radiometric quantities can be expressed in terms of radiance, for

(a) Sun behind a pole (b) Soap bubbles (c) Cloud iridescence

Figure 1.4: Phenomena that cannot be explained by the corpuscular theory: (a) the sun rays do not seem to travel in straight light as light leaks through the pole (additionally, a halo can also be seen); (b) the thin layer of soap in the bubbles creates interferences which results in interesting colors; (c) cloud iridescence: the sunlight is scattered and diffracted by tiny droplets of water or ice crystals in the air.

instance the irradiance is given by

$$L(\boldsymbol{x}, \boldsymbol{\omega}) = \frac{dE(x)}{\cos\theta d\boldsymbol{\omega}}, \qquad E(\boldsymbol{x}) = \int_{\Omega} L(\boldsymbol{x}, \boldsymbol{\omega}) \cos\theta d\boldsymbol{\omega},$$

where the integral is computed over the unit hemisphere Ω. An important property of the radiance is that it remains constant along a ray. Furthermore, the sensor response of digital cameras is directly proportional to the radiance. This makes it very useful: as we will see in Chapter 3, coupled with the theory of bidirectional reflectance distribution functions (BRDFs) and the rendering equation, the radiance and other radiometric quantities provide a powerful tool to analyze and simulate the light transport in a scene.

1.2.2 Wave theory

The corpuscular theory cannot explain all the behaviors of light. For instance, in Figure 1.4, the sun's light creates some unique phenomena that cannot be predicted by tracing rays in the scene. More generally, the corpuscular model falls short of providing an explanation to diffraction, interference and polarization.

A new school of thought appeared when Christian Huygens hypothesized that light behaves like a wave [80]. His theory was based on the assumption that light moves at different speeds in different mediums, and provided among other things a simple explanation of refraction. The wave theory was further strengthened by Thomas Young's famous double-slit experiment, where he created interference patterns by splitting and combining the sun's light.

Augustin-Jean Fresnel was another scientist who played an important role in the acceptance of the wave theory. In 1817, he submitted a manuscript to the French Academy of Sciences to explain the phenomenon of interference using wave theory. Siméon Denis Poisson, who was a convinced corpuscular theorist, was determined to find a flaw in Fresnel's submission. He thought he succeeded when Fresnel's theory was predicting a bright spot in the middle of the shadow

cast by a circular object illuminated with a point light source. However, to his stupefaction, experiments showed that the now famous *Poisson spot* was indeed real.

We now briefly describe the basic concepts regarding the wave nature of light. For more details, we direct the reader to [35, 145, 160, 167].

The wave equation

An optical wave can be mathematically described by the real scalar function $u(\boldsymbol{r}, t)$, which is a function of position $\boldsymbol{r} = [x, y, z]^\top$ and time t. The optical wave $u(\boldsymbol{r}, t)$ satisfies the wave equation:

$$\nabla^2_{\boldsymbol{r}} u - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = 0, \tag{1.1}$$

where c is the speed of light in the medium, and $\nabla^2_{\boldsymbol{r}}$ is the Laplacian operator with respect to the position.

Monochromatic waves

A *monochromatic wave* $u(\boldsymbol{r}, t)$ represents an optical wave that is composed of a single frequency:

$$u(\boldsymbol{r}, t) = a(\boldsymbol{r}) \cos(\varphi(\boldsymbol{r}) + \omega t), \tag{1.2}$$

where $a(\boldsymbol{r})$ denotes the amplitude, $\varphi(\boldsymbol{r})$ the phase and ω the angular frequency (in rad/s). Additionally, we denote the *wavenumber* $k = \omega/c$, which represents the spatial frequency. It is often convenient to express optical waves in a complex form (or phasor notation):

$$U(\boldsymbol{r}, t) = a(\boldsymbol{r}) e^{j\varphi(\boldsymbol{r})} e^{j\omega t} = \hat{U}(\boldsymbol{r}) e^{j\omega t},$$

where $\hat{U}(\boldsymbol{r}) = a(\boldsymbol{r}) e^{j\varphi(\boldsymbol{r})}$ is called the *complex amplitude* and characterizes the spatial variation of the wave. Clearly, we have

$$u(\boldsymbol{r}, t) = \mathfrak{Re}\left\{U(\boldsymbol{r}, t)\right\}.$$

The Helmoltz equation If we substitute the expression for the monochromatic wave into the wave equation (1.1), we obtain the *Helmoltz equation*:

$$(\nabla^2_{\boldsymbol{r}} + k^2) U(\boldsymbol{r}) = 0.$$

Plane waves A *plane wave* is a wave whose wavefronts are planar and perpendicular to its direction of propagation $\boldsymbol{k}$; for $U(\boldsymbol{r}, t)$ to satisfy the wave equation, we need $\|\boldsymbol{k}\| = k$. The complex amplitude of a plane wave is given by

$$\hat{U}(\boldsymbol{r}) = A e^{-j\boldsymbol{k}^\top \boldsymbol{r}},$$

where $A = A_0 e^{j\varphi_0}$ is called the *complex envelope*. The wavefronts of a plane wave describe infinite parallel planes that are separated by a distance λ called the *wavelength* and which relates to the frequency and the wavenumber via the following relation:

$$\lambda = \frac{2\pi}{k} = \frac{2\pi c}{\omega}.$$

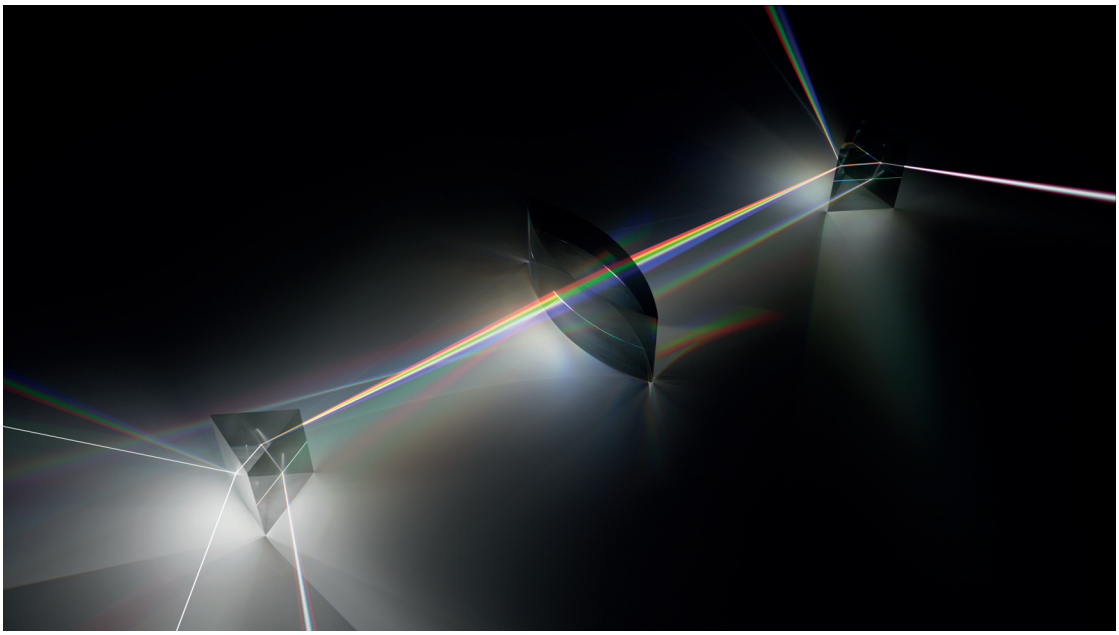


Figure 1.5: Newton's double prism experiment: white collimated light enters the first prism in the bottom left. The white light is dispersed and decomposed into the colors of the spectrum as it exits the prism. The spectrum is then focused with a convex lens and recombined into white light by the prism in the top right. Simulation created with the computer graphics software Blender [33] and the physically based renderer LuxCoreRender [107].

Each wavelength has an associated color, ranging from around 400 nm for violet to 700 nm for red. The decomposition of visible light into colors according to their wavelength is called the *spectrum*. Interestingly, the term *spectrum* was coined by Isaac Newton, a strong advocate of the corpuscular theory, who first discovered it with his famous *crucial experiment* [133], reproduced in Figure 1.5. In this experiment, he created a spectrum by decomposing white light with a prism. He also showed that white light can be recreated by passing the spectrum through a second prism.[2] This experiment is a beautiful illustration of Fourier analysis, in which signals are studied and analyzed from their spectral representation.

Spherical waves Other typical functions that satisfy the wave equation include *spherical waves*, whose complex amplitude is given by:

$$\hat{U}(\boldsymbol{r}) = \frac{A}{\|\boldsymbol{r}-\boldsymbol{c}\|} e^{-jk\|\boldsymbol{r}-\boldsymbol{c}\|},$$

where $\|\boldsymbol{r}-\boldsymbol{c}\|$ is the distance from the center $\boldsymbol{c}$ of the sphere.

[2] Newton's crucial experiment has also been replicated on the cover of [190].

Polychromatic waves

A polychromatic wave $u(\boldsymbol{r}, t)$ can be thought of as a superposition integral of several monochromatic waves (1.2):

$$\begin{aligned} u(\boldsymbol{r}, t) &= \int_0^\infty a(\boldsymbol{r}, \omega) \cos(\varphi(\boldsymbol{r}, \omega) + \omega t) d\omega \\ &= \mathfrak{Re}\left\{ \int_0^\infty \hat{U}(\boldsymbol{r}, \omega) e^{j\omega t} d\omega \right\}, \end{aligned}$$

where $\hat{U}(\boldsymbol{r}, \omega) = a(\boldsymbol{r}, \omega) e^{j\varphi(\boldsymbol{r}, \omega)}$ is the complex amplitude for frequency ω. As in the monochromatic case, we can express a polychromatic wave as a complex function, which is given by

$$U(\boldsymbol{r}, t) = \int_0^\infty \hat{U}(\boldsymbol{r}, \omega) e^{j\omega t} d\omega, \tag{1.3}$$

Expression (1.3) is also known as *Gabor's analytic signal*. Additionally,

$$\begin{aligned} u(\boldsymbol{r}, t) &= \int_0^\infty \frac{a(\boldsymbol{r}, \omega)}{2} \left(e^{j(\varphi(\boldsymbol{r}, \omega) + \omega t)} + e^{-j(\varphi(\boldsymbol{r}, \omega) + \omega t)} \right) d\omega \\ &= \frac{1}{2} \int_0^\infty \hat{U}(\boldsymbol{r}, \omega) e^{j\omega t} d\omega + \frac{1}{2} \int_0^\infty \hat{U}^*(\boldsymbol{r}, \omega) e^{-j\omega t} d\omega \\ &= \frac{1}{2} \int_0^\infty \hat{U}(\boldsymbol{r}, \omega) e^{j\omega t} d\omega + \frac{1}{2} \int_{-\infty}^0 \hat{U}^*(\boldsymbol{r}, -\omega) e^{j\omega t} d\omega. \end{aligned}$$

Since $u(\boldsymbol{r}, t)$ is real, $\hat{U}$ must be conjugate symmetric; that is, $\hat{U}(\boldsymbol{r}, -\omega) = \hat{U}^*(\boldsymbol{r}, \omega)$. Therefore,

$$u(\boldsymbol{r}, t) = \frac{1}{2} \int_{-\infty}^\infty \hat{U}(\boldsymbol{r}, \omega) e^{j\omega t} d\omega,$$

which means that the spectrum is nothing but (twice) the Fourier transform of the wave $u(\boldsymbol{r}, t)$:

$$\hat{U}(\boldsymbol{r}, \omega) = 2 \int_{-\infty}^\infty u(\boldsymbol{r}, t) e^{-j\omega t} dt.$$

The pulsed plane wave. The analogue to the monochromatic plane wave for polychromatic light is called the *pulsed plane wave*: it consists of the superposition of plane waves traveling in the same direction and is given by

$$\begin{aligned} U(\boldsymbol{r}, t) &= \int_0^\infty \hat{U}(\boldsymbol{r}, \omega) e^{j\omega t} d\omega \\ &= \int_0^\infty A(\omega) e^{-j\boldsymbol{k}(\omega)^\top \boldsymbol{r}} e^{j\omega t} d\omega, \end{aligned}$$

where $A(\omega)$ is the complex amplitude of the component with frequency ω.

Intensity of optical waves

It is not possible to physically measure or observe the quantity $u(\boldsymbol{r}, t)$ directly. What we can observe however is the *optical intensity*, which is proportional to the time-averaged squared

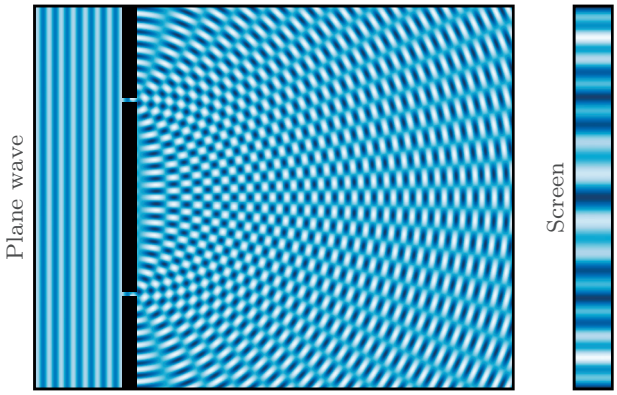


Figure 1.6: Double slit experiment: a plane wave passes through two slits in a wall. This generates cylindrical waves that interfere with each other, creating interference patterns on the screen.

amplitude. For a monochromatic wave, in phasor notation, this can be expressed as the squared magnitude of $U(\boldsymbol{r}, t)$:

$$I(\boldsymbol{r}) = |U(\boldsymbol{r})|^2 = U(\boldsymbol{r}, t)U^*(\boldsymbol{r}, t).$$

Using Parseval's theorem [143], we can also compute the intensity of a polychromatic wave with

$$\begin{aligned} I(\boldsymbol{r}) &= \int_{-\infty}^{\infty} U(\boldsymbol{r}, t)U^*(\boldsymbol{r}, t)dt \\ &= 4\int_{0}^{\infty} P(\boldsymbol{r}, \omega)d\omega, \end{aligned}$$

where $P(\boldsymbol{r}, \omega) = U(\boldsymbol{r}, \omega)U^*(\boldsymbol{r}, \omega)$ is the *spectral density function* of polychromatic light.

Note that the problem of recovering U from its intensity is called *phase retrieval* and is the topic of Chapter 6.

Properties of light waves

Optical waves can interact with objects or with each other. Two concepts that are particularly interesting in the context of this thesis are reflections and interference.

Reflection When the light changes the medium in which it travels, a (partial) reflection occurs. In the process, the wave can undergo a phase shift and be attenuated; this is expressed by the *reflection coefficient* $r = \rho e^{j\theta}$, where ρ is the attenuation factor and θ the phase shift[3]. With this notation, we can represent the reflection of planar waves with the following operator:

$$\mathcal{R}\{U\}(\boldsymbol{r}, t) = rAe^{-j(\boldsymbol{k}-2(\boldsymbol{k}^\top \boldsymbol{n})\boldsymbol{n})^\top \boldsymbol{r}} e^{j\omega t}.$$

While this expression represents a general reflection, it can be simplified in some specific cases. For instance, when we assume that we have a perfect metallic mirror ($r = -1$) and that the

[3]Note that we use $\boldsymbol{r} = [x, y, z]^\top$ for the position and $r = \rho e^{j\theta}$ for the reflection coefficient; these two quantities should not be confused.

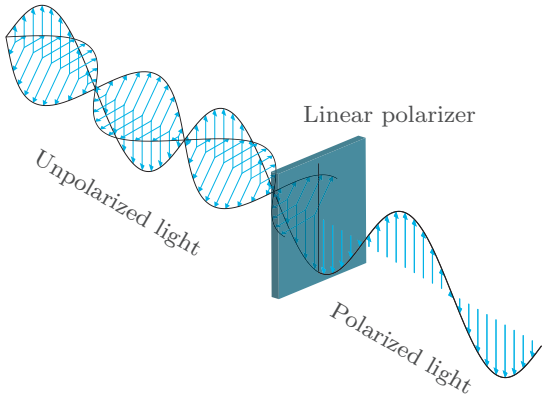


Figure 1.7: A linear polarizing filter: unpolarized light is transformed into polarized light that oscillates in the direction of the polarizing filter.

wave travels in the opposite direction with respect to the surface normal $\boldsymbol{n}$ of the mirror, that is, when $\boldsymbol{k} = -\boldsymbol{n}$, we have

$$\mathcal{R}\{U\}(\boldsymbol{r}, t) = -U^*(\boldsymbol{r})e^{j\omega t}.$$

Interference Let U_1 and U_2 be two plane waves. Since the wave equation is linear, the principle of superposition applies: $U_1 + U_2$ is also a solution to the wave equation, and hence, a valid optical wave. Given $U_1 = I_1^{1/2} e^{j\varphi_1}$ and $U_2 = I_2^{1/2} e^{j\varphi_2}$, the result is a wave with complex amplitude

$$U = U_1 + U_2$$

and intensity given by the *interference relation*

$$\begin{aligned} I &= |U_1 + U_2|^2 \\ &= |U_1|^2 + |U_2|^2 + U_1^* U_2 + U_1 U_2^* \\ &= I_1 + I_2 + 2\sqrt{I_1 I_2}\cos(\varphi). \end{aligned}$$

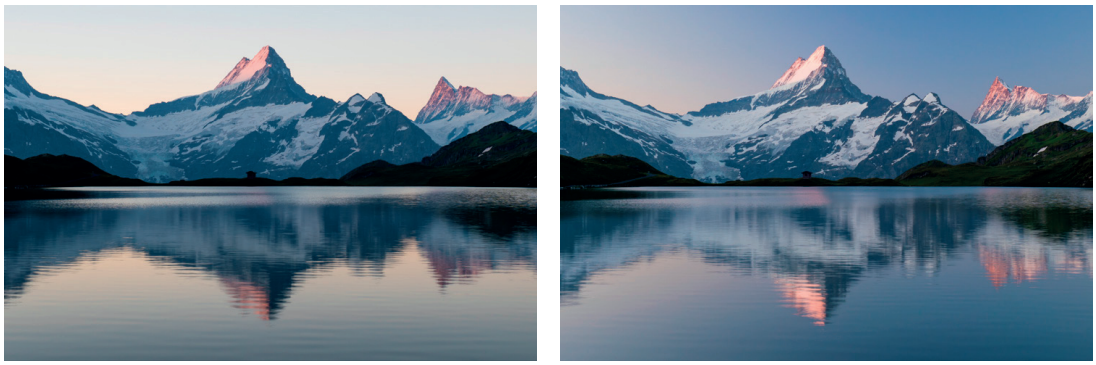
where $\varphi = \varphi_2 - \varphi_1$. When $I_1 = I_2 = I_0$, the interference relation reduces to

$$\begin{aligned} I &= 2I_0(1 + \cos(\varphi)) \\ &= 4I_0 \cos^2\left(\frac{\varphi}{2}\right). \end{aligned}$$

The notion of interference is often leveraged to demonstrate the wave-nature of light, as for instance in the famous *double slit experiment*. In this experiment, a plane waves goes through two narrows slits. This results in two circular waves that interfere with each other. By placing a screen, it is possible to visualize this interference, as illustrated in Figure 1.6.

1.2.3 Electromagnetic theory

We now turn to our third model for light. Again, we refer to the sun to underline where the wave theory of light as described above is limited. If you have ever worn polarized sunglasses, you



(a) Without filter (b) With filter

Figure 1.8: Effect of a polarizing filter on a natural scene: the sky is darker and bluer, and some of the bright reflections on the surface of the lake are diminished, especially near the shore. As a result, contrasts are increased. The effect is stronger in the directions that are perpendicular to the sun, which is rising on the left outside of the frame.

probably have noticed that they are extremely efficient at cutting the glare due to large horizontal surfaces such as lakes or roads. When you rotate them by 90° however, their effect is null. The same phenomenon can be observed in Figure 1.1 which was photographed with a polarizing filter. On the left side of the photograph, the filter was rotated such that its effect is maximum; as a result, the surface reflection is strongly reduced. On the right side, its effect is minimum, and the scene exhibits a strong reflection on the surface of the water. This short experiment suggests that an ingredient is missing from the wave theory: light seems to have an orientation in addition to its direction of propagation. Therefore, instead of describing it with a scalar field as in the wave theory, we need a vector field to represent it, as illustrated in Figure 1.2c. The orientation of the light is called *polarization*. Large horizontal surfaces generate horizontally polarized light and the sunglasses can selectively remove this light, while letting through the light with other polarizations.

Historically, James Clerk Maxwell is the father of electromagnetism. Not only did he unify electricity and magnetism, he also considered light as an electromagnetic field, governed by what is now known as *Maxwell's equations* [116]. In this model, light is composed of electric and magnetic fields that travel perpendicularly to each other. Additionally, Maxwell plays a special role in the history of color photography: he was the first person to realize color photographs with the so-called *three-color method*.

Light polarization

Formally, polarization is a physical property of light that characterizes the orientation of electromagnetic transverse waves in space. A wave is said to be *polarized* when its orientation follows a regular pattern in space: for example, when the wave oscillates in a fixed direction it is called *linearly* polarized, and when the polarization rotates at a constant rate along the wave propagation axis it is called *circularly* polarized. In contrast, *unpolarized* light is composed of a mixture of electromagnetic waves with different orientations.

As illustrated in Figure 1.7, a linear polarizing filter (or polarizer) is a physical device that

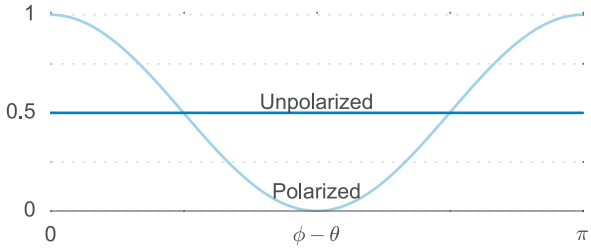


Figure 1.9: Illustration of Malus' law: while the intensity of unpolarized light is equally affected for all polarization orientations, polarized light is attenuated according to a $\cos^2(\phi - \theta)$ factor.

only lets through light with a given polarization. Polarizing filters diminish reflections from shiny surfaces such as wet objects or vegetation. Hence, they are widely used in photography[4], since they allow to reduce undesirable glare and tend to make skies bluer in natural scenes, as illustrated in Figure 1.8. Another typical use case of polarizers is to reduce reflections from a body of water; this is what we observed in Figure 1.1. In fact, it is even possible to completely cut the surface reflection. To achieve this, we would need to position the camera at *Brewster's angle*, which is defined as the angle for which the reflected light is perfectly polarized. For an interface between air and water, Brewster's angle is approximately 53° with respect to the surface normal.

Given a light wave polarized in direction ϕ with initial intensity I_0, the polarizing filter projects it onto its orientation θ. This is summarized by *Malus' law*, which quantifies the light intensity I after it goes through the filter:

$$I = I_0 \cos^2(\phi - \theta). \tag{1.4}$$

Unpolarized light can be thought of as containing all orientations equally, hence the resulting intensity is simply $\frac{I_0}{2}$, which is the mean value of $I_0 \cos^2(\phi - \theta)$ over all orientations. Figure 1.9 illustrates the effect of a linear polarizer on both polarized and unpolarized light, following Malus' law.

1.2.4 Wave-particle duality

In the twentieth century, physicists revisited the corpuscular theory, describing light as packets of energy, or *photons*. They also realized that some behaviors of light could only be explained from a photonic point of view, while others were only justified by the wave model, essentially leading to the *wave-particle duality* theory. The wave-particle duality theory is outside the scope of this thesis, but research on the topic has ead to significant advancements in the field of photography. For example, Einstein's photoelectric effect [58], whereby some metals emit electrons when light is shone on them, is crucial to build sensors for digital cameras.

[4]In general, only *circular* polarizing filters are used in photography; the reason is that linear polarizers interfere with the auto-focus and light-metering sensors of digital cameras. A circular polarizer can be obtained by adding a *quarter-wave plate* to a linear polarizer.

1.3 Thesis outline and main contributions

In this thesis, we revisit the aforementioned different manifestations of light with a signal processing perspective and introduce tools, sampling schemes, algorithms and analyses based on the various facets of light. An important signal processing concept we regularly rely on is *sampling*. Notwithstanding quantum theory and microscopic effects, our world is described by continuous quantities. However, computers think and work with values that are discrete in time and amplitude. Therefore, we need a way to convert a continuous signal to discrete values. Ideally, it is desirable not to lose information or distort the signals in the process: this is precisely the role of sampling theorems, which provide conditions and theoretical guarantees regarding the perfect reconstruction of sampled signals. The most famous sampling theorem is to due to Shannon [165, 166], Nyquist [136], Edmund and John Whittaker [195, 196], and Kotelnikov [96]: it describes the sufficient sampling rate such that a bandlimited continuous signal can be sampled and perfectly reconstructed. In this dissertation, we explore and propose sampling theorems for the case when the signals of interest are not bandlimited.

Another recurring theme is the notion of *sparsity*. Loosely speaking, a signal is said to be sparse if it can be represented by a relatively small number of parameters. The particular structure of sparsity is often used to regularize ill-posed problems. Sparse regularization is for instance the foundation of compressed sensing [53], which enables the recovery of a sparse signal from a small number of measurements. In the context of this thesis, we take advantage of sparsity in several chapters.

Chapter 2—Sampling Diracs and Non-Bandlimited Shapes Before we get into the details of light, we first develop new sampling results for non-bandlimited shapes. We build on the theory of finite rate of innovation (FRI) and provide theoretical guarantees for sampling and perfect reconstruction for signals that can be expressed as weighted sums of Lorentzian functions; such signals can be considered as sparse since they can be represented with few parameters. We call this new sampling scheme *variable pulse width* (VPW)-FRI. The VPW-FRI framework is first developed for 1D signals, and then generalized to higher dimensions and spherical signals.

In the noiseless case, perfect recovery is guaranteed by a set of theorems. In addition, we verify that our algorithm is robust to model mismatch and noise.

Summary of Contributions in Chapter 2

- We introduce VPW-FRI, a new model for the FRI sampling framework that enables the use of pulses with variable width to be sampled and perfectly reconstructed.
- We provide a noise analysis of the method.
- We generalize VPW-FRI to higher dimensions and spherical signals.

Chapter 3—Corpuscular Theory: Acquisition of Reflectance Functions The FRI sampling theory is connected to the topic of light in Chapter 3, which has a more practical flavor. In this chapter, we start by exposing the background regarding the acquisition of the reflectance function of materials. We then describe our *light domes*, which are devices that we built to perform such acquisition. We then show how the FRI theory is leveraged to acquire the reflectance function of materials. First, we propose a technique to sample and represent specularities, which are

narrow pulses that can be challenging to acquire. This technique does not assume the sampling kernel—in other words the shape of the specularity—to be known a priori. Finally, we apply the VPW-FRI theory from Chapter 2 to the problem of representating the reflectance function of materials. Within this framework, we demonstrate applications in 1 and 2 dimensions, as well as in the spherical domain, which is the natural representation of reflectance functions.

Summary of Contributions in Chapter 3

- We build *light domes* that we exploit to sample and acquire BRDFs.
- We apply VPW-FRI to the problem of sampling the BRDF.
- We propose another variation of FRI when the sampling kernel is unknown and apply it to the problem of specularity estimation.

Chapter 4—Wave Theory and Interferences: Lippmann Photography The wave nature of light is the foundation of Chapter 4, where the theory of wave optics is combined with a signal processing framework to analyze Lippmann photography. Lippmann (or interferential) photography is the first and only analog photography method that can capture the full spectrum of a scene in a single take. This century old technique records the colors of a scene by creating interference patterns inside the depth of a photosensitive plate.

Lippmann photography provides a great opportunity to demonstrate several fundamental concepts in signal processing. Conversely, a signal processing perspective enables us to shed new light on the technique. For instance, we describe new behaviors whose explanations were ignored by physicists to date. Furthermore, we show that the spectra generated by Lippmann plates are in fact distorted versions of the original spectra. More interestingly, we propose an algorithm to eliminate these artifacts and recover the original spectra. Since most of the Lippmann plates have been captured in the early 1900s, the combination of these artworks and our reconstruction algorithm provides a tool to look into the past with hyperspectral glasses. Finally, we verify our model and algorithms through extensive experiments by creating our own Lippmann photographs.

Summary of Contributions in Chapter 4

- We propose a rigorous study of the Lippmann process and an analytical model to explain the reproduction of colors.
- We introduce physically accurate simulations that allow us to study the assumptions made in the analytical model.
- We precisely characterize the color reproduction of Lippmann plates and we report new undocumented effects based on the influence of the reflector on the reflected spectrum.
- We study conditions for the inversion of the recording and rendering procedures.
- We create our own Lippmann plates to verify the theory.

Chapter 5—Electromagnetic Theory: Micro-Polarizer Arrays We then move one step further and consider the electromagnetic nature of light and study the polarization phenomenon. Depending on their nature, reflections of polarized light on surfaces can preserve or not its polarization. We take advantage of this observation and study the acquisition of images using polarizing filters, which provide a powerful way to separate diffuse and specular reflection. Traditional methods rely on several captures and require proper alignment of the filters. Recently, camera manufacturers have proposed to embed polarizing micro-filters in front of the sensor, creating a mosaic of pixels with different polarizations.

In this chapter, we investigate the advantages of such camera designs. In particular, we consider different design patterns for the filter arrays and propose a suite of algorithms to demosaic an image generated by such camera. This essentially allows us to separate the diffuse and specular components using a single image. The performance of our algorithm is then compared with a color-based method using synthetic and real data. Finally, we demonstrate how to retrieve the normals of a scene using the diffuse images estimated by our method.

Summary of Contributions in Chapter 5

- We investigate different designs for cameras micro-polarizing filter arrays.
- We propose algorithms to separate the diffuse and specular components from a mosaic of polarized pixels.
- We demonstrate our separation technique on synthetic and real data.
- We show that our separation technique improves the estimation of the surface normals of a scene.

Chapter 6—Super-Resolution Phase Retrieval We close this thesis as we started it: with a more signal processing-oriented chapter, where we leverage the FRI sampling theory to solve the problem of *phase retrieval*. More specifically, we consider the problem of phase retrieval for sparse signals, which are defined as a small number of atoms with given locations and amplitudes. This problem is also relevant in the context of light as the phase information is always lost in imaging devices. Therefore, phase retrieval algorithms are of great interest in fields such as X-ray crystallography [91, 119], astronomy [73] or optical imaging [169].

Solving the phase retrieval problem is equivalent to recovering a signal from its auto-correlation function. We describe an algorithm that tackles the problem in three steps: i) we super-resolve the auto-correlation function of the signal from a limited number of samples by invoking the finite rate of innovation sampling theory ii) we design a greedy algorithm that estimates the locations of a sparse solution given the super-resolved auto-correlation function, iii) we retrieve the amplitudes of the atoms given their locations and the measured auto-correlation function.

Along with the algorithm, we derive its performance bound with a theoretical analysis and propose a set of enhancements to improve its computational complexity and noise resilience. Finally, we demonstrate the benefits of the proposed method via a comparison against Charge Flipping, a notable algorithm in crystallography.

Summary of Contributions in Chapter 6

- We apply FRI to super-resolve the autocorrelation function of a signal.
- We propose a greedy algorithm to recover a sparse signal from its autocorrelation function.
- We analyze our algorithm from a theoretical point of view and through simulations.
- We demonstrate that our solution outperforms the Charge Flipping algorithm, which is the state-of-the-art in crystallography.

Chapter 2

Sampling Diracs and Non-Bandlimited Shapes*

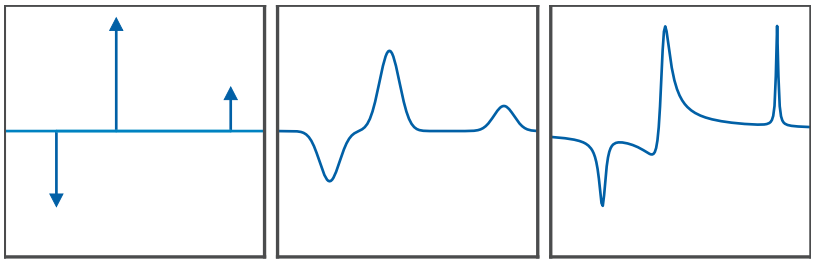
If we want to solve a problem that we have never solved before, we must leave the door to the unknown ajar.

What Do You Care What Other People Think?
Richard Feynman

2.1 Introduction

Sampling theorems provide a bridge between analog signals of the real world and their discrete representations. In the ideal case, they allow perfect reconstruction of a continuous signal from only a limited number of discrete samples. The most notable embodiment of this concept is the well-known Nyquist-Shannon theorem [166]. It states that any signal that belongs to the shift-invariant subspace of bandlimited functions can be perfectly reconstructed from a set of uniformly spaced discrete samples, provided that the sampling frequency is at least twice the bandwidth of the signal. While very powerful, this result falls short for many classes of signals.

*The material in this chapter is the result of joint work with Adam Scholefield, Loïc Baboulaz and Martin Vetterli. Author contributions: GB, LB and MV designed the initial research on the 1D case; GB, AS and MV designed the extension to higher dimensions; GB performed research, ran simulations and experiments, and wrote the chapter based on [10], which was written by himself.



(a) Stream of Diracs (b) Fixed-width (c) Variable-width

Figure 2.1: Signals with finite rate of innovation: (a) stream of Diracs recoverable with the original FRI theory; (b) fixed-shaped pulses recoverable using Strang-Fix theory; (c) variable-width pulses, recoverable by the proposed technique.

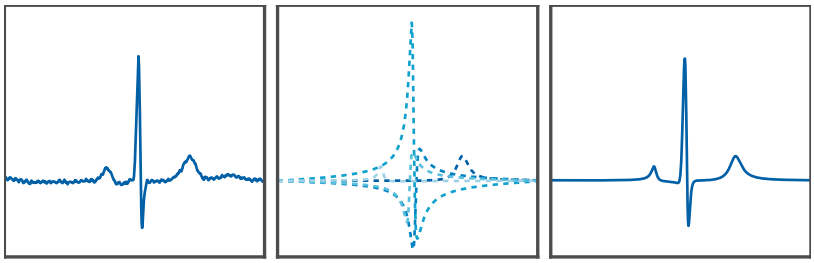
Over the years, sampling results have emerged for signals belonging to more general classes, including signals that lie in general shift-invariant subspaces [187] and *some* other signals with a finite rate of innovation (FRI) [191], i.e., signals that have a finite number of degrees of freedom per unit time. This latter class extends the former to many other types of signals such as streams of Dirac deltas as depicted in Figure 2.1a. Other examples include piecewise polynomials or piecewise sinusoidal signals [23].

All these schemes assume that we only have access to samples of a filtered version of the signal; that is, the original continuous-time signal $x(t)$ is filtered with a kernel $\varphi(t)$ before being uniformly sampled at time instants $t = nT$. Typically, $\varphi(t)$ is a lowpass filter that is either chosen by design, or, as is more common, is a characteristic of the acquisition system.

Like traditional Nyquist-Shannon sampling, the original FRI theory was developed for infinite support kernels, such as the sinc ideal low-pass filter; however, the results have been recently extended to physically realizable kernels of compact support that satisfy Strang-Fix conditions [55, 178]. Another benefit of these more general kernels is that, when applied to streams of Diracs, they can be used to add more variety to the signals admitted by FRI. For example, it has enabled FRI theory to be utilised to retrieve streams of short pulses from ultrasound imaging [184] and action potentials of neurological data from calcium imaging [137]. Note that, although this allows more general pulses, as illustrated in Figure 2.1b, each pulse is still constrained to have the same shape (up to a multiplicative factor).

The question we address in this chapter is the following: Can we find a signal model that allows a different shape for each pulse? This is motivated by the fact that in nature, several signals can be parametrized as pulses with various shapes. In what follows, we show that it is indeed possible by considering signals that consist of sums of Lorentzian[1] pulses with variable width and different degrees of asymmetry. While there are numerous parametric functions—Gaussians, B-splines, or wavelet basis functions to name a few—that are suitable to model signals with variable pulse shapes, the choice of Lorentzian functions is not innocuous; indeed, as we see in the next sections, the auto-regressive nature of their spectral coefficients enables us to estimate the parameters of the model using spectral estimation techniques such as *Prony's method*. Other algorithms estimating the parameters of sums of Lorentzians are described and analyzed in [3, 4, 24]; in this chapter, we connect these results to sampling theory.

[1] A Lorentzian function—also known as Cauchy distribution—is a bell-shaped curve that looks similar to a Gaussian but with heavier tails.



(a) Original signal (b) VPW decomposition (c) Estimated signal

Figure 2.2: Example of decomposition of an ECG signal into VPW pulses: (a) one period of a pulse from the MIT-BIH arrhythmia database [71]; (b) its decomposition into five asymmetric VPW pulses; (c) the VPW approximation (sum of the five pulses).

In addition to FRI techniques, recent extensions of compressive sensing also enable continuous parameter estimation but attack the problem from a different angle: while FRI theory relies on non-linear methods, these approaches enforce sparsity with the atomic norm and convex relaxations [25, 39, 40, 181].

In this chapter, we demonstrate, via a set of theorems, that sums of Lorentzians can be sampled and perfectly reconstructed, by adapting standard FRI machinery. Whereas the core of the original FRI theory is composed of Diracs, the atoms of our signal model are called *variable pulse width* (VPW) pulses.[2] As highlighted in Figure 2.1c, the VPW model brings more versatility and flexibility to the set of FRI signals that can be acquired. We present here the theory of sampling and reconstructing VPW signals and generalize it to higher dimensions and spherical signals. The VPW model can be used in many applications: as depicted in Figure 2.2, it provides a faithful and compact representation of ECG signals. In the next chapter, we move back to the main theme of this thesis—light—and show how the VPW-FRI framework allows us to efficiently sample and reconstruct reflectance functions.

2.2 Background

Before considering VPW signals, we briefly review the standard FRI theory.

2.2.1 Signals with finite rate of innovation

The *rate of innovation* ρ of a signal is defined as its number of degrees of freedom per unit time [191]. Ideally, we seek an algorithm to reconstruct an FRI signal using *only* ρ parameters per unit time, in which case sampling at the rate of innovation allows perfect reconstruction. This is precisely what is achieved by the Nyquist-Shannon theorem, where the degrees of freedom are embodied by the expansion coefficients in the orthonormal basis formed by shifted sinc functions.

Another archetypal FRI signal is a τ-periodic stream of K weighted Dirac pulses:

$$f(t) = \sum_{n\in\mathbb{Z}} \sum_{k=0}^{K-1} c_k \delta(t - t_k - n\tau).$$

[2] Variable pulse width (VPW)-FRI was coined in [11, 152] in the context of electrocardiogram (ECG) representation. Other notable applications of VPW on multiple lead ECG signals are reported in [122, 123].

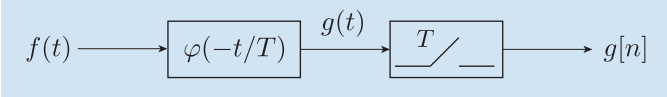


Figure 2.3: Sampling operation. The continuous-time signal $f(t)$ is filtered with the kernel $\varphi(t)$ and then recorded at time instants $t = nT$ ($n \in \mathbb{Z}$) to obtain the discrete-time signal $g[n]$.

Since a Dirac can be completely described by its position and amplitude, one period of the signal $f(t)$ has precisely $2K$ free parameters (K positions and K amplitudes) and a rate of innovation of $\rho = 2K/\tau$.

In what follows, we show how it is possible to reconstruct $f(t)$ from $N \geq 2K + 1$ samples obtained after an ideal low pass filter. To do this, we first demonstrate how we can calculate some of the Fourier series coefficients of f from its samples and then show how these Fourier series coefficients can be used to calculate the $2K$ parameters of the signal.

2.2.2 From samples to Fourier series coefficients

Although we do not have direct access to the Fourier series coefficients of $f(t)$, we can obtain a subset of them by taking the DFT of the samples obtained from the sampling scheme depicted in Figure 2.3, with $\varphi(t)$ equal to an ideal low-pass filter. This result was first shown as part of [191, Theorem 1], and we reproduce it here in the following Lemma.

Lemma 2.1 (Vetterli, Blu and Marziliano 2002) Assume we take N samples of a τ-periodic function f, according to Figure 2.3; that is, we have

$$g[n] = \langle f(t), \varphi(t - nT) \rangle, \quad n \in 0, 1, \ldots, N-1.$$

Furthermore, assume

$$\varphi(t) = B\,\mathrm{sinc}(Bt) = B\sin(\pi Bt)/(\pi Bt),$$

with $T = \tau/N$ and B satisfying $N = 2\lfloor B\tau/2 \rfloor + 1$. Define the DFT of $g[n]$ as

$$G[m] = \sum_{n=0}^{N-1} g[n] e^{-2\pi jnm/N}$$

and the Fourier series coefficients of $f(t)$ as

$$F[m] = \frac{1}{\tau} \int_{-\tau/2}^{\tau/2} f(t) e^{-2\pi jmt/\tau} dt.$$

Then, $F[m] = NG[m]$ for $|m| \leq \lfloor B\tau/2 \rfloor$.

Proof. See Appendix 2.A. ■

2.2.3 From Fourier series coefficients to signal parameters

When $f(t)$ is a τ-periodic stream of K weighted Dirac pulses, its Fourier series coefficients are given by

$$F[m] = \frac{1}{\tau} \sum_{k=0}^{K-1} c_k e^{-2\pi j t_k m/\tau}, \tag{2.1}$$

where $m \in \mathbb{Z}$. Thus, our problem is to calculate the parameters c_k and t_k from a finite number of Fourier series coefficients, $F[m]$. This is a common problem in spectral analysis, which can be solved using, for example, *Prony's method*. The idea is to find a filter $A[m]$ that, when convolved with the signal, produces a null signal; in other words, $A[m]$ annihilates the signal $F[m]$:

$$(F * A)[m] = 0, \qquad \forall m \in \mathbb{N}. \tag{2.2}$$

For this reason, Prony's method is also referred to as the *annihilating filter method*. Since $F[m]$ has the form given in (2.1), it is annihilated by the filter whose z-transform is

$$A(z) = \sum_{k=0}^{K} A[k] z^{-k} = \prod_{k=1}^{K} (1 - u_k z^{-1}),$$

where $u_k = e^{-2\pi j t_k/\tau}$.

Therefore, given the roots of the annihilating filter, we can compute the locations of the Diracs as $t_k = -\frac{\tau \angle u_k}{2\pi}$, where $\angle u_k$ is the phase of u_k. It remains to show how to find the annihilating filter coefficients, and thus the roots. To that end, (2.2) can be rewritten in matrix form as

$$\boldsymbol{S}\boldsymbol{h} = \boldsymbol{0}, \tag{2.3}$$

where $\boldsymbol{S} \in \mathbb{C}^{K \times (K+1)}$ is a Toeplitz matrix of rank K formed by consecutive values of $F[m]$ and $\boldsymbol{h}$ is a vector containing the $K+1$ annihilating filter coefficients $A[k]$. Observe that we need at least $2K+1$ consecutive coefficients to solve the system of equations given in (2.3) and fully recover the annihilating filter; the set of coefficients we choose does not necessarily need to be located around the DC component, but the most efficient strategy is to pick $F[m]$ for $m = -K-1, \ldots, K$, as it minimizes the frequency at which the signal is sampled. In this case, the bandwidth B of the low-pass filter $\varphi(t)$ should be larger or equal to $(2K+1)/\rho$, which represents one more sample per unit of time than the rate of innovation of the signal.[3]

Finally, we need to show how to retrieve the coefficients $\{c_k\}_{k=0}^{K-1}$, which, after calculating $\{t_k\}_{k=0}^{K-1}$, is a simple linear problem. Taking K consecutive Fourier coefficients (or more) and letting

$$\boldsymbol{V} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ u_0 & u_1 & \cdots & u_{K-1} \\ \vdots & \vdots & \ddots & \vdots \\ u_0^{K-1} & u_1^{K-1} & \cdots & u_{K-1}^{K-1} \end{bmatrix},$$

$$\boldsymbol{f} = \begin{bmatrix} F[0] & F[1] & \cdots & F[K-1] \end{bmatrix}^\top,$$

$$\text{and } \boldsymbol{c} = \begin{bmatrix} c_0 & c_1 & \cdots & c_{K-1} \end{bmatrix}^\top,$$

[3] The "+1" term is due to a technicality, namely that the DC component of the Fourier series has no phase information, and in that sense, contributes little to the reconstruction.

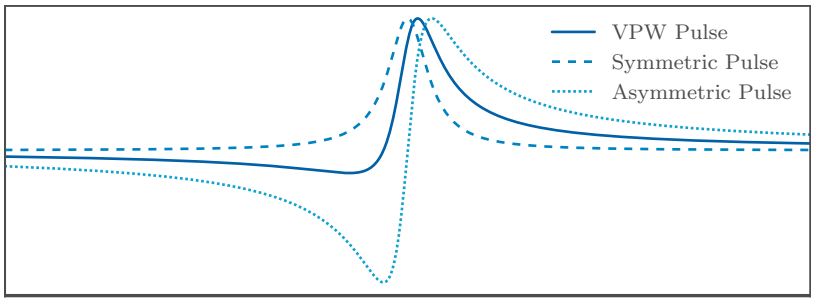


Figure 2.4: Example of a VPW pulse and its decomposition. Continuous line: VPW pulse; dashed line: symmetric pulse; dotted line: asymmetric pulse. The parameters used are $t_0 = 0.5, r_0 = 0.03$ and $c_0 = d_0 = 0.1$. Only one period is shown.

The expression (2.1) can be written as

$$\frac{1}{\tau}\boldsymbol{V}\boldsymbol{c} = \boldsymbol{f}. \tag{2.4}$$

The Vandermonde matrix $\boldsymbol{V}$ is non-singular since $u_k \neq u_\ell$ for any $k \neq \ell$. Therefore, the amplitudes $\boldsymbol{c}$ can be recovered by $\boldsymbol{c} = \tau \boldsymbol{V}^{-1}\boldsymbol{f}$.

After this brisk review of FRI sampling and reconstruction, we are ready to extend the model to more general pulses.

2.3 Pulses with variable width: The 1D case

This section demonstrates how to generalize the Dirac pulse model to account for signals with variable width and asymmetry. More precisely, we show how to sample and reconstruct variable-width pulses of the form

$$f_k(t) = f_k^s(t) + f_k^a(t), \tag{2.5}$$

where

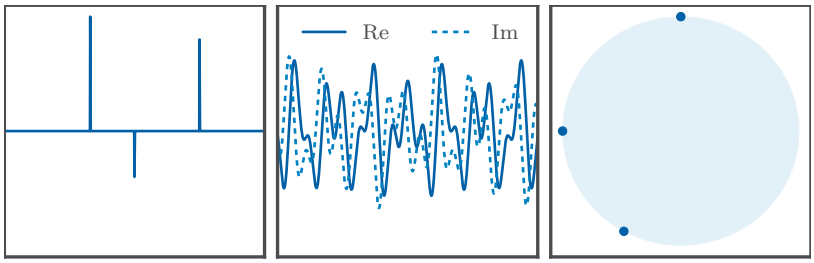
$$f_k^s(t) = c_k \sum_{n\in\mathbb{Z}} \frac{r_k}{\pi\left(r_k^2 + (t - t_k - n\tau)^2\right)}, \tag{2.6}$$

and

$$f_k^a(t) = d_k \sum_{n\in\mathbb{Z}} \frac{t - t_k - n\tau}{\pi\left(r_k^2 + (t - t_k - n\tau)^2\right)}. \tag{2.7}$$

Here, $f_k^s(t)$ and $f_k^a(t)$ are the symmetric and anti-symmetric components of the pulse, respectively. The combination of the symmetric and asymmetric pulses (2.5) is the main building block of our signal model and we call it a *variable pulse width* (VPW) pulse. An example of a VPW pulse and its decomposition into symmetric and asymmetric components is shown in Figure 2.4.

Expressions (2.6) and (2.7) follow naturally from a simple observation about the roots of the annihilating filter. In particular, in the classical FRI theory for streams of Diracs, the roots of the z-transform of the annihilating filter all have a magnitude of 1 and thus lie on the unit circle, as shown in Figure 2.5. Since these roots are complex numbers, they can potentially carry more information than just the location of the pulses; indeed, in the classical case, only the phase is extracted and the information regarding the magnitude is simply disregarded.



(a) Time domain (b) Frequency domain (c) Roots of the filter

Figure 2.5: Classical FRI for streams of Diracs: (a) time-domain signal; (b) real and imaginary parts of the spectrum, which is made of a sum of complex exponentials; (c) complex roots of the z-transform of the filter annihilating the spectrum.

2.3.1 Variable width

To leverage the information about the magnitude of the root and control the width of the pulses, we introduce a new parameter r_k called the *width* and we redefine the roots as

$$u_k = e^{-2\pi(r_k + jt_k)/\tau}.$$

With this adjustment and for a single root, we can build the following Fourier series coefficients:

$$F_k^s[m] = \frac{c_k}{\tau} e^{-2\pi(r_k|m| + jt_k m)/\tau}, \qquad m \in \mathbb{Z}. \tag{2.8}$$

The absolute value of m is used to ensure conjugate symmetry of the spectrum. The corresponding τ-periodic time-domain signal is precisely the Lorentzian function described in (2.6).

2.3.2 Asymmetry

Pulses such as (2.8) are symmetric; in many practical applications however, asymmetric pulses are desirable. The symmetric formulation can be generalized by considering the following Fourier series coefficients:

$$F_k^a[m] = -\frac{jd_k}{\tau} \operatorname{sgn}(m) e^{-2\pi(r_k|m| + jt_k m)/\tau}, \quad m \in \mathbb{Z},$$

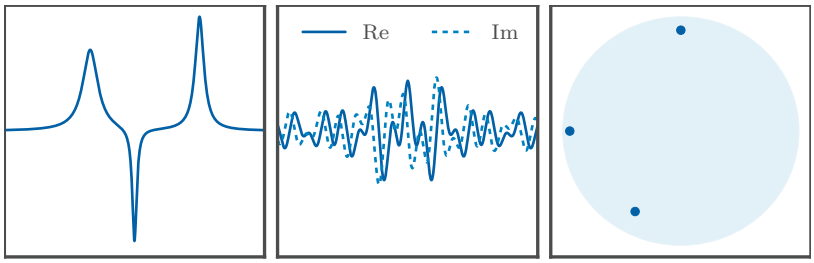
where $d_k \in \mathbb{R}$. The signal F_k^a is the Hilbert transform of F_k^s and its (τ-periodic) time-domain representation is given in (2.7).

From its Fourier series expression, we can also derive an alternative formula for $f_k(t)$, which avoids the infinite sum and is given by[4]

$$f_k(t) = \frac{c_k}{\tau} \frac{1 - |z_k(t)|^2}{(1 - z_k(t))(1 - z_k^*(t))} + \frac{d_k}{\tau} \frac{2\operatorname{Im}\{z_k(t)\}}{(1 - z_k(t))(1 - z_k^*(t))},$$

where $z_k(t) = e^{2\pi(-r_k + j(t - t_k))/\tau}$.

[4]The details of the derivations are provided in Appendix 2.B.



(a) Time domain (b) Frequency domain (c) Roots of the filter

Figure 2.6: VPW pulses: (a) time-domain signal; (b) real and imaginary parts of the spectrum, which is made of a damped sum of complex exponentials; (c) complex roots of the z-transform of the filter annihilating the spectrum.

VPW pulses are in essence a generalization of Diracs; by setting $d_k = 0$ and taking the limit of $f_k(t)$ when r_k goes to zero, we obtain a Dirac delta with amplitude c_k located at time t_k.

As illustrated in Figure 2.6, several VPW pulses can be combined to form a *VPW signal*:

$$f(t) = \sum_{k=0}^{K-1} f_k(t) \tag{2.9}$$
$$F[m] = \sum_{k=0}^{K-1} F_k[m], \qquad m \in \mathbb{Z}.$$

2.3.3 Recovery of VPW pulse parameters

We now show how to sample and reconstruct a sum of Lorentzian pulses.

THEOREM 2.2 Let f be the sum of K continuous τ-periodic asymmetric Lorentzians as described in (2.9). Assume we take N samples of f, according to Figure 2.3; i.e., we obtain

$$g[n] = \langle f(t), \varphi(t - nT) \rangle, \quad n \in 0, 1, \ldots, N-1.$$

Furthermore, assume $\varphi(t)$ is such that the Fourier series coefficients of f, denoted $F[m]$, can be obtained from the samples, for $0 \leq m \leq \lfloor (N-1)/2 \rfloor$. For example, if $\varphi(t) = B\,\mathrm{sinc}(Bt)$, with $T = \tau/N$ and B satisfying $N = 2\lfloor B\tau/2 \rfloor + 1$, this is achieved by Lemma 2.1.

Then, if $N \geq 4K + 1$, the samples, $\{g[n] : n \in 0, 1, \ldots, N-1\}$, are sufficient to recover f.

PROOF. To prove the theorem, we first show how to construct the unique annihilating filter for the Fourier series coefficients of f. We then show how the $4K$ parameters of f can be retrieved from this filter. Let

$$A(z) = \sum_{k=0}^{K} A[k] z^{-k} = \prod_{k=0}^{K-1} (1 - u_k z^{-1}),$$

Algorithm 2.1 VPW-FRI parameter estimation

Input: $N \geq 4K+1$ discrete samples $x[n]$ $(n = 0, 1, \dots N-1)$ of a VPW-FRI continuous-time signal $x(t)$ and K its number of pulses.

Output: The locations, widths and amplitudes $\{t_k, r_k, c_k\}_{k=0}^{K-1}$ of $x(t)$.

Compute $X[m]$ the FFT of $x[n]$
Build the Toeplitz matrix $\boldsymbol{S}$ from the positive $X[m]$ of the spectrum
Denoise $\boldsymbol{S}$ or $x[n]$ according to one of the strategies of Section 2.3.5
Find the annihilating filter coefficients by solving $\boldsymbol{Sh} = \boldsymbol{0}$
Compute u_k, the roots of $\boldsymbol{h}$
$t_k = -\frac{\tau \angle u_k}{2\pi}$
$r_k = \frac{\tau \log |u_k|}{2\pi}$
Build the Vandermonde matrix $\boldsymbol{V}$ from the roots u_k
Solve $\frac{1}{\tau}\boldsymbol{Vc} = \boldsymbol{x}$, where $\boldsymbol{x}$ contains the FFT coefficients $X[m]$ and $\boldsymbol{c} \in \mathbb{C}$ the amplitudes c_k
return $\{t_k, r_k, c_k\}_{k=0}^{K-1}$

and $F[m]$ be the Fourier series coefficients of f. Then,

$$\begin{aligned}(A * F)[m] &= \sum_{\ell=0}^{K} A[\ell] F[m-\ell] \\ &= \sum_{\ell=0}^{K} \sum_{k=0}^{K-1} (c_k - j d_k) A[\ell] u_k^{m-\ell} \\ &= \sum_{k=0}^{K-1} (c_k - j d_k) \left(\sum_{\ell=0}^{K} A[\ell] u_k^{-\ell} \right) u_k^m = 0.\end{aligned}$$

Note that we have restricted ourselves to positive indices ($m \geq 0$). This is because, due to its decaying nature, the VPW spectrum has a *cusp* at $m = 0$, which prevents us from performing the annihilation on both positive and negative values of the spectrum. Finding the unique filter A requires at least $2K+1$ Fourier series coefficients corresponding to non-negative frequencies. By Lemma 2.1, this can be achieved with $N \geq 4K+1$ samples of f.

As in the case of streams of Diracs, the locations $\{t_k\}_{k=0}^{K-1}$ can be retrieved from the K roots of $A(z)$: $u_k = e^{-2\pi j t_k/\tau}$. Moreover, the widths $\{r_k\}_{k=0}^{K-1}$ are given by $r_k = \tau \log|u_k|/(2\pi)$. Finally, the parameters $\{c_k\}_{k=0}^{K-1}$ and $\{d_k\}_{k=0}^{K-1}$ are retrieved by solving (2.4) with

$$\boldsymbol{c} = \begin{bmatrix} c_0 + j d_0 & c_1 + j d_1 & \cdots & c_{K-1} + j d_{K-1} \end{bmatrix}^\top . \qquad \blacksquare$$

The estimation procedure is summarized in Algorithm 2.1. Since the algorithm requires $4K+1$ samples to retrieve $4K$ parameters, it is only one sample away from the rate of innovation.

2.3.4 Sampling kernels

Although we concentrate on the ideal low-pass filter or Dirichlet sampling kernel, VPW signals are fully compatible with the more general kernels presented in the FRI literature. In particular, as outlined in Theorem 2.2, we can use any kernel that allows us to retrieve enough Fourier series coefficients of the continuous-time signal. In [55], it is shown that this can be achieved

with kernels that satisfy the generalized Strang-Fix conditions, such as *cardinal exponential splines* [188] or sum of sinc kernels [184].

Moreover, even if the sampling kernel does not meet these conditions, we can relax the requirement of perfect reconstruction and achieve accurate approximations using either the correction filter proposed in [34] or the framework set out in [189].

2.3.5 Denoising techniques

In practice, it is unlikely that a signal perfectly matches the VPW-FRI model. Even if this is the case, the acquisition device might introduce some noise in the process. This justifies the need for a denoising block before the estimation of the parameters. Here, we review several methods borrowed from spectral estimation theory [177] and evaluate their performance on our estimation procedure.

We model the noise $\epsilon[n]$ as i.i.d. additive Gaussian with zero mean and covariance matrix $\sigma^2 \boldsymbol{I}$. Furthermore, we assume that it is introduced after sampling; that is, the noisy measurements are

$$\widetilde{f}[n] = f[n] + \epsilon[n] \qquad n = 0, 1, \dots N-1.$$

We denote the noisy Toeplitz matrix built from the samples $\widetilde{f}[n]$ as $\widetilde{\boldsymbol{S}}$. Several denoising strategies have been introduced in the classical FRI framework and they are directly applicable to the VPW-FRI model. We summarize and compare the main techniques below.

Cadzow

Cadzow's algorithm [37] is an iterative method that alternatively enforces rank K and a Toeplitz structure for $\widetilde{\boldsymbol{S}}$. The low rank approximation is performed by computing the SVD of $\widetilde{\boldsymbol{S}}$ and keeping only its K largest singular values. To ensure that $\widetilde{\boldsymbol{S}}$ is Toeplitz, we simply average across diagonals. In practice, the method seems to converge to a solution that satisfies both properties, but it should be noted that this has not been proven theoretically. Furthermore, there is no guarantee that the final solution is close, in any way, to $\boldsymbol{S}$.

The matrix pencil

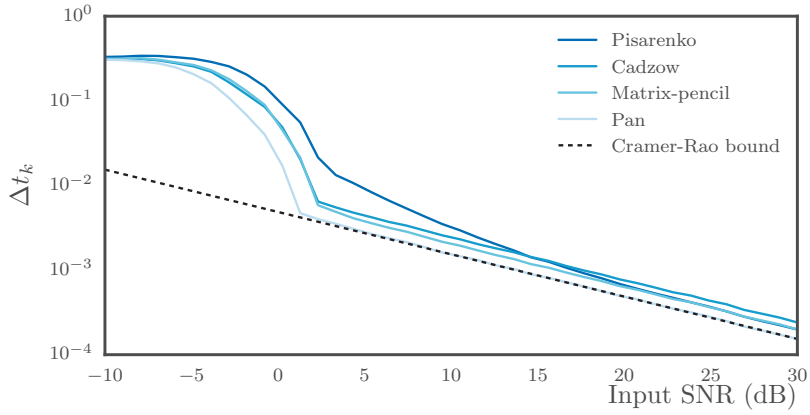
Another widely used technique is the matrix pencil method [78]—also known as ESPRIT [158]—which exploits the rotational invariance property of the underlying signal subspace. Unlike Cadzow, it is a non-iterative algorithm.

Pisarenko

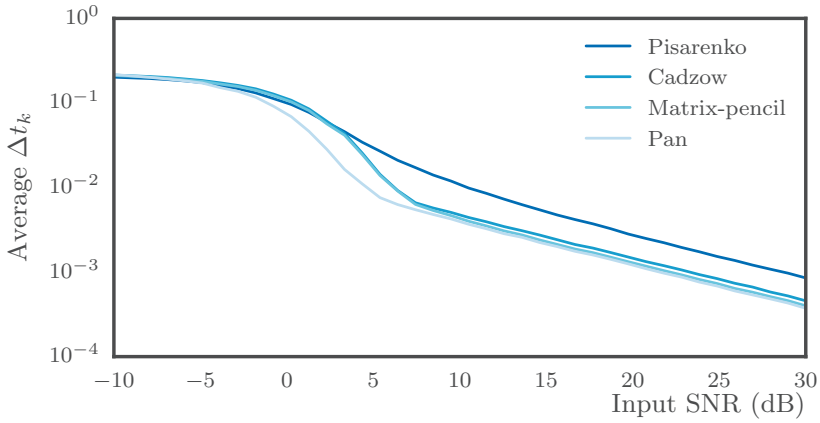
Pisarenko's method [150] estimates the annihilating filter by simply extracting the last column of the matrix $\boldsymbol{V}$ of the SVD of $\widetilde{\boldsymbol{S}} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^*$. Indeed, the last columns of $\boldsymbol{V}$ provide an orthogonal basis for the nullspace of $\widetilde{\boldsymbol{S}}$ and in our case, the nullspace of $\boldsymbol{S}$ is one-dimensional. Note that Pisarenko's method or the matrix pencil can actually be used sequentially after Cadzow.

Pan's method

The last method we survey is inspired by IQML [36] and has been recently introduced to the FRI framework by Pan in [142]. The annihilating filter $\boldsymbol{h}$ is found iteratively and its computation is



(a) Single VPW pulse



(b) Two VPW pulses

Figure 2.7: Comparison of the effect of different denoising approaches on the estimation of the location of (a) one and (b) two VPW pulses: 10 Fourier coefficients are used in the estimation procedure and the experiment has been repeated 5,000 times and averaged. For the single pulse case, the Cramér-Rao bound for the estimation of the location has been computed (dashed line).

formulated as

$$\begin{aligned} \min_{\widehat{\boldsymbol{f}},\boldsymbol{h}} \quad & \|\widetilde{\boldsymbol{f}} - \boldsymbol{B}\widehat{\boldsymbol{f}}\|_2^2 \\ \text{subject to} \quad & \boldsymbol{h} * \widehat{\boldsymbol{f}} = \boldsymbol{0}, \end{aligned}$$

where $\boldsymbol{B}$ is a linear transformation that maps the signal that can be annihilated to the measurements. Unlike the other techniques, this approach is not applied to $\widetilde{\boldsymbol{S}}$ but directly to the measured signal: this can result in significant performance improvements since the transformation operation from the measurements to the signal to be annihilated sometimes dramatically alters the structure of the noise. The main ingredient of the algorithm is the matrix $\boldsymbol{B}$, which is different for every FRI problem. In the case of VPW-FRI, $\boldsymbol{B}$ is the transform that maps the denoised DFT coefficients $\widehat{\boldsymbol{f}}$ to the discrete-time measured signal $\widetilde{\boldsymbol{f}}$.

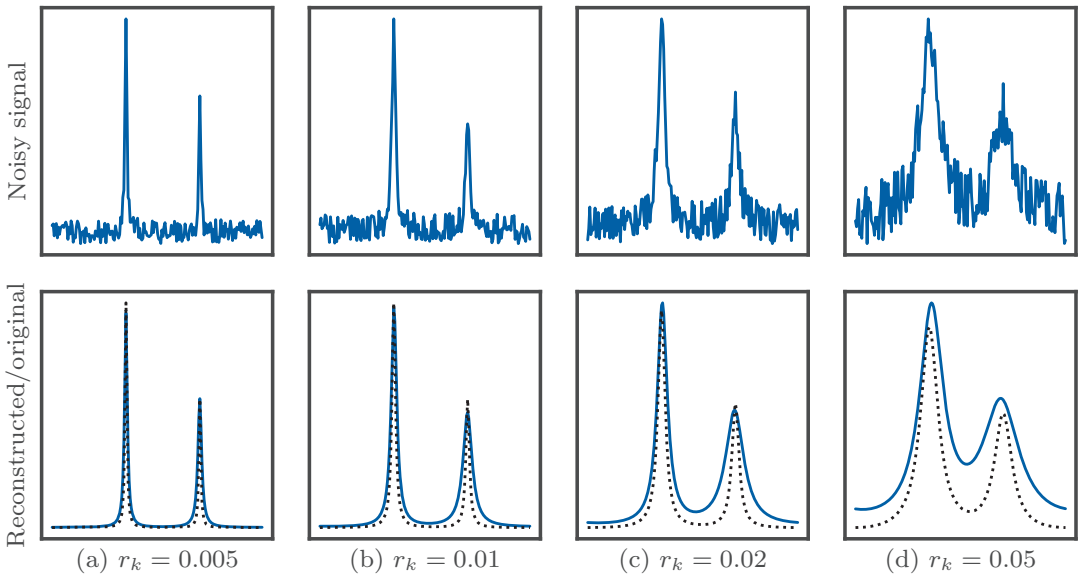


Figure 2.8: Reconstruction error for different pulse widths. The top row represents noisy signals with an SNR of 0 dB used as inputs to the VPW estimation and the bottom row shows the estimated signal in blue as well as the original signal with dotted black lines. The measured mean squared errors are (a) 0.002, (b) 0.006, (c) 0.012, (d) and 0.030.

In Figure 2.7, we compare the performance of the different denoising methods under various levels of noise for one and two VPW pulses. We observe that the matrix-pencil and Cadzow have a similar performance. Furthermore, Pan's method outperforms all other methods by a fairly large margin.

Cramér-Rao bound

For one pulse, we also derive and display the Cramér-Rao bound (CRB), which provides a theoretical lower bound on the best possible performance achievable by any unbiased estimator. The details of the computations of the CRB are provided in Appendix 2.C. We see that, for a pulse with width $r_0 = 0.02$, the VPW-FRI estimation procedure is fully efficient down to SNRs of about 0 dB. Past this threshold, we notice a gap between the estimator and the CRB. Further analysis on this gap can be found in [16] in the context of FRI for streams of Diracs.

We also notice that the quality of estimation depends on the width parameter: for an equivalent SNR, the narrower the pulse, the better the performance. This effect is showcased in Figure 2.8. Intuitively, a higher value of r_k results in a spectrum that decays faster; this makes the estimation more challenging and less resilient to noise.

2.4 Pulses with variable width: The 2D case

We now describe how the VPW model generalizes to 2D, by extending results on 2D streams of Diracs [113]. Later, we will also consider the spherical case, which is particularly relevant to

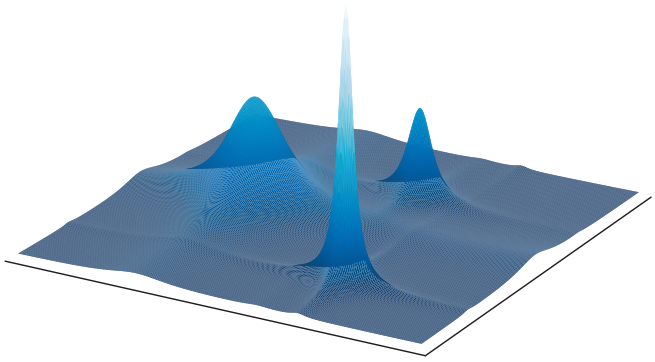


Figure 2.9: Example of a 2D VPW signal composed of three pulses. The parameters used are $(x_k, y_k) = \{(0.6, 0.7), (0.7, 0.2), (0.3, 0.5)\}$, $r_k = \{0.035, 0.02, 0.085\}$, $c_k = \{0.75, 0.7, 5\}$, $\tau = 1$.

reflectance functions.

The second dimension introduces more flexibility in the model, allowing us to reproduce several different shapes. In what follows, we discuss four different pulse shapes, each offering different properties such as asymmetry or radiality. We also demonstrate how to estimate the parameters using the FRI algorithm.

We use the pair of variables (x, y) to index 2D spatial measurements and (m, n) to index 2D Fourier series coefficients. We also assume that the 2D sampling kernel is a separable kernel given by the tensor product of two 1D functions: $\varphi(x, y) = \varphi_x(x)\varphi_y(y)$. Therefore, the natural 2D extension of our model is given by the tensor product of two 1D VPW pulses, say $f_{k,x}(x)$ and $f_{k,y}(y)$. The Fourier series coefficients are given by

$$\begin{aligned} F_k[m, n] &= \frac{1}{\tau} F_{k,x}[m] F_{k,y}[n] \\ &= \frac{1}{\tau} e^{-\frac{2\pi}{\tau} j(x_k m + y_k n)} e^{-\frac{2\pi}{\tau} r_k (|m| + |n|)}, \end{aligned}$$

Similarly, we define a VPW signal as a linear combination of K pulses; that is,

$$F[m, n] = \sum_{k=0}^{K-1} F_k[m, n], \qquad m, n \in \mathbb{Z}. \tag{2.10}$$

An example of a 2D VPW signal composed of three pulses is shown in Figure 2.9.

2.4.1 Estimation of the parameters

Observe that if we fix one of the frequency indices, say m, the expression (2.10) reduces to

$$F[n] = \frac{1}{\tau^2} \sum_{k=0}^{K-1} \widetilde{c}_k e^{-\frac{2\pi}{\tau} j y_k n} e^{-\frac{2\pi}{\tau} r_k |n|}, \tag{2.11}$$

where $\widetilde{c}_k = c_k e^{-j\frac{2\pi}{\tau} x_k m} e^{-\frac{2\pi}{\tau} r_k |m|}$. Now that we have reduced the problem to the 1D case, we can invoke Lemma 2.1 and Theorem 2.2 to connect the DFT coefficients with the Fourier series coefficients and guarantee the uniqueness of reconstruction of the continuous-space signal. As in

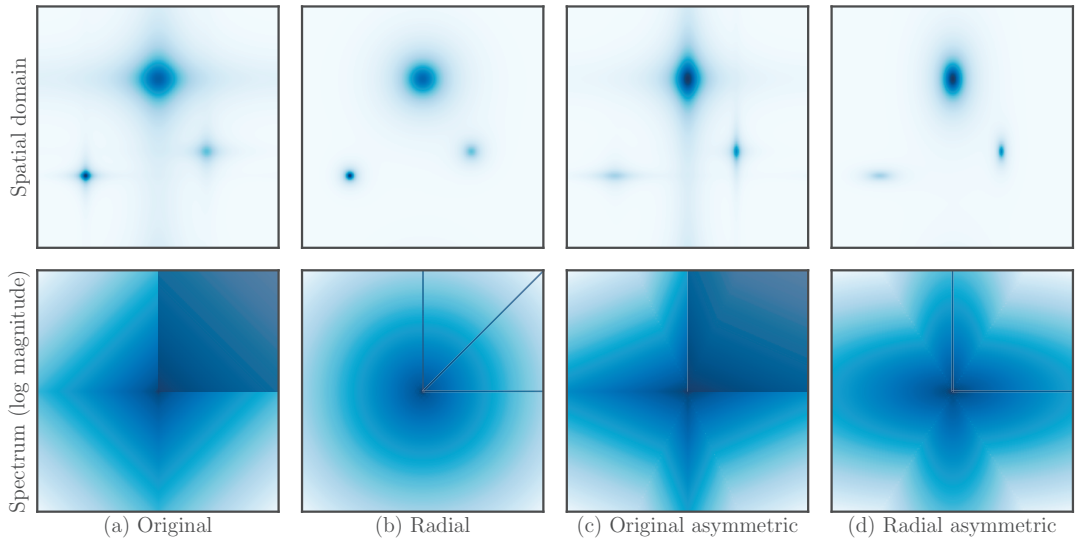


Figure 2.10: Family of 2D VPW pulses and their spectrum: (a) original formulation, (b) radial formulation, (c) original anisotropic formulation, and (d) radial anisotropic formulation. The area of annihilation is represented by a dark overlay.

the 1D case, we cannot exploit the full spectrum for annihilation; due to the singularity at the DC component, we restrict ourselves to the positive indices m and n. We refer to the sets of positive n for a fixed m as *half-columns* and to the sets of positive m for a fixed n as *half-rows*. The coefficients used are part of what we call the *annihilation region*. The spectrum and the annihilation region of the signal described in (2.10) are shown in Figure 2.10a.

Following (2.11), we see that each half-column of the spectrum can be annihilated independently by the same filter. More precisely, we can build Toeplitz matrices with each of these half-columns and stack them on top of each other to obtain the matrix $\boldsymbol{S}^c$:

$$\boldsymbol{S}^c = \begin{bmatrix} \boldsymbol{S}^c_0 \\ \boldsymbol{S}^c_1 \\ \vdots \\ \boldsymbol{S}^c_{\lfloor M/2 \rfloor - 1} \end{bmatrix},$$

where $\boldsymbol{S}^c_m$ is the Toeplitz matrix corresponding to the mth half-column of the spectrum. Similarly, by fixing n, the half-rows of the spectrum can be annihilated and we can create the matrix $\boldsymbol{S}^r$. Both $\boldsymbol{S}^c$ and $\boldsymbol{S}^r$ are block Toeplitz and can be used independently to recover the locations along the x and y axes, respectively, following the procedure presented in Section 2.3.3.

To solve the annihilation equation, each half-column and half-row of the spectrum must have at least $K+1$ elements. Therefore, we need at least $(2K+1)^2$ samples to recover $4K$ parameters; unlike in the 1D case, we need to sample the signal at a significantly higher rate than the rate

of innovation to achieve perfect recovery.[5]

Another detail needs to be taken care of in 2D. Since we reformulate the estimation of the parameters as two 1D problems, we obtain two sets of locations $\{x_k\}_{k=0}^{K-1}$ and $\{y_k\}_{k=0}^{K-1}$ but we do not know the *corresponding pairs* of these recovered locations. If K is small, we can solve this issue by combinatorial search over all the possible pairs; for each combination of locations, we compute the amplitude of the corresponding pulses and compare the reproduced signal with the original. We then select the pair whose reproduction corresponds (or, in the noisy case, is closest) to the measured signal.

Unfortunately, this brute-force approach quickly becomes impractical as K gets larger. To cope with this issue, we use the following heuristic. Instead of searching through all pairs of K pulses, we suppose that all combinations are valid, that is, we have K^2 potential pulses. We then compute the amplitude of each of these pulses by solving (2.4). Finally, the correct pairs of locations are revealed by the K largest pulse amplitudes. This simple approach has no proven correctness, but experiments have shown that it performs well in practice.

2.4.2 Radial formulation

Observe that the 2D Lorentzian pulse described in (2.10) has heavy tails along the two main axes. Perhaps a more suitable model is given by

$$F[m,n] = \frac{1}{\tau^2} \sum_{k=0}^{K-1} c_k e^{-\frac{2\pi}{\tau} j(x_k m + y_k n)} e^{-\frac{2\pi}{\tau} r_k \sqrt{m^2+n^2}}, \tag{2.12}$$

for $m, n \in \mathbb{Z}$. The difference with (2.10) resides in the damping factor: in (2.10), the exponent is the ℓ^1-norm of the Fourier indices, while in (2.12), the ℓ^2-norm is used. As a result, the spectrum of (2.12) does not exhibit the sharp transitions of the original representation; this corresponds to radial pulses in the primal domain, as shown in Figure 2.10b.

The disadvantage of this representation is that its region of annihilation is not as large as in the previous model. Such a signal can be annihilated along the DC components, that is, when we set either m or n equal to zero, but also along the diagonal line defined as $m = n$. More generally, the spectrum can be annihilated along the line defined as $m = an$ or $n = bm$ for $a, b \in \mathbb{N}$, as long as we have a sufficient number (at least $K + 1$) of coefficients on that line. The annihilation region of this signal can be seen in Figure 2.10b.

2.4.3 Anisotropic pulses

We can further generalize the 2D pulse shape and introduce anisotropy in the model by using a different width parameter for each axis. For instance, the radial formulation can be rewritten as

$$F[m,n] = \frac{1}{\tau^2} \sum_{k=0}^{K-1} c_k e^{-\frac{2\pi}{\tau} j(x_k m + y_k n)} e^{-\frac{2\pi}{\tau} \sqrt{r_{m,k}^2 m^2 + r_{n,k}^2 n^2}},$$

where $r_{m,k}$ and $r_{n,k}$ are the respective widths along the main axes. Examples of anisotropic pulses for both the original and radial formulations are illustrated in Figure 2.10c and 2.10d.

[5] Recent results [141] have shown that it is in fact possible to sample and perfectly reconstruct streams of Diracs when sampling at the rate of innovation in 2 or more dimensions; this approach can also be generalized to the VPW pulse shape.

2.5 Pulses with variable width: On the sphere

Recently, FRI theory has been developed for spherical signals by Deslauriers-Gauthier and Marziliano [51]. Building on their work, we show that the VPW framework can also be generalized to the spherical domain.

2.5.1 Preliminaries

We start by reviewing a few definitions used in spherical signal processing. We refer the interested reader to [52, 153] for a more thorough introduction.

Functions on the sphere $\mathbf{S}^2 = \{\boldsymbol{x} \in \mathbb{R}^3 : \|\boldsymbol{x}\| = 1\}$ are usually characterized by elevation and azimuth angles, denoted θ and ϕ respectively. The spherical Dirac delta is defined as

$$\delta(\theta, \phi; \theta_0, \phi_0) = \delta(\cos\theta - \cos\theta_0)\delta(\phi - \phi_0),$$

where (θ_0, ϕ_0) denotes its location. A sum of weighted Diracs is then defined as

$$f(\theta, \phi) = \sum_{k=0}^{K-1} c_k f_k(\theta, \phi) = \sum_{k=0}^{K-1} c_k \delta(\theta, \phi; \theta_k, \phi_k).$$

On the sphere, the spherical harmonic (SH) decomposition provides the analog to the Fourier series expansion. As shown in [52], for the signal $f(\theta, \phi)$, it is given by[6]

$$\widehat{f}_{\ell,m} = N_{\ell,m} \sum_{k=0}^{K-1} c_k P_{\ell,|m|}(\cos\theta_k) e^{-jm\phi_k},$$

with $\ell = 0, 1, \ldots$ and $|m| < \ell$. Here, $N_{\ell,m}$ is a normalization constant, and $P_{\ell,m}$ denotes the *associated Legendre polynomial*, defined as

$$P_{\ell,m}(\cos\theta_k) = (-1)^m (\sin\theta_k)^{|m|} \frac{d^m}{d(\cos\theta_k)^m} P_\ell(\cos\theta_k),$$

where $P_\ell(\cdot)$ is the Legendre polynomial. Since P_ℓ is of degree ℓ, the degree of $P_{\ell,m}$ is at most $\ell - m$.

2.5.2 Sampling the hemisphere

Several strategies can be employed to sample a hemisphere; the simplest one is *equiangular sampling*, i.e. uniform sampling along the azimuth and elevation angles. The drawback of this approach is that it results in non-uniform sampling of the hemisphere, as illustrated in Figure 2.11a: the area around the pole has a higher density than the area near the equator.

Another simple technique that leads to almost uniform sampling is *spherical Fibonacci sampling* and can be seen in Figure 2.11b. Other strategies such as quasi-random sampling [74, 174] can also be used in this context. In Chapter 3, we nevertheless use equiangular sampling since in practice it allows for a faster and more natural scanning of the hemisphere.

[6] In some of the spherical harmonics related literature, indices are expressed with a superscript/subscript notation, e.g. f_ℓ^m. We choose to include both indices in the subscript, to avoid confusion with powers.

Figure 2.11: Two possible sampling schemes on the hemisphere: (a) equiangular and (b) spherical Fibonacci sampling. Each of these hemispheres contains 300 samples.

2.5.3 From samples to spherical harmonic coefficients

We define the *bandwidth* L of a spherical signal as the smallest L such that $\widehat{f}_{\ell,m} = 0$ for all $\ell \geq L$. In [56, Theorem 3], it is proved that the SH coefficients of a spherical bandlimited function can be exactly computed given a minimum number of equiangular samples. For completeness, we restate this result below.

THEOREM 2.3 (DRISCOLL AND HEALY 1994) Let g be the result of convolving the spherical signal f with an ideal lowpass filter φ with bandwidth L. Assume we further sample g at points $(\theta_n, \phi_{n'})$, where $\theta_n = \frac{\pi n}{2L}$ for $n = 0, 1, \ldots, 2L-1$ and $\phi_{n'} = \frac{\pi n'}{L}$ for $n' = 0, 1, \ldots, 2L-1$. Then the SH coefficients of g are given by

$$\widehat{g}_{\ell,m} = \frac{\sqrt{2\pi}}{2L} \sum_{n=0}^{2L-1} \sum_{n'=0}^{2L-1} \alpha_n g(\theta_n, \phi_{n'}) Y^*_{\ell,m}(\theta_n, \phi_{n'}), \tag{2.13}$$

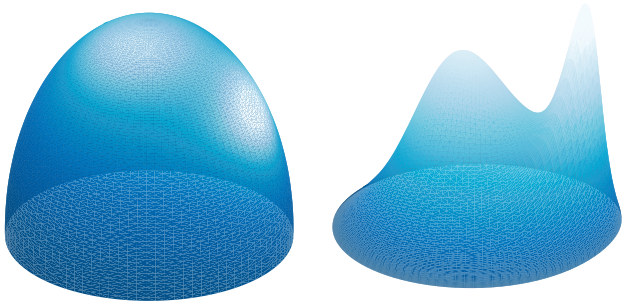
where $Y^*_{\ell,m}(\theta_n, \phi_{n'})$ are the SH basis functions evaluated at $(\theta_n, \phi_{n'})$ and $\{\alpha_n\}_{n=0}^{2L-1}$ is the unique solution to

$$\sum_{n=0}^{2L-1} \alpha_n P_\ell(\cos\theta_n) = \sqrt{2}\delta_{\ell 0},$$

for $\ell = 0, 1, \ldots, 2L-1$.

PROOF. See [56]. ■

Other sampling theorems exists for different sampling strategies: for example, if we choose N samples uniformly at random on the sphere, we can retrieve the SH coefficients of g when $N \geq L^2$, with probability one [18]. Finally, remark that if we have fewer samples, the spectrum often decays fast enough for (2.13) to yield a good approximation.



(a) Color plot (b) Planar projection

Figure 2.12: Example a spherical VPW signal composed of two pulses. There are different ways to display (hemi)spherical signals: (a) a plot of a unit hemisphere where the brightness of colors represents the intensity, and (b) a projection of the same signal onto the xy plane with the intensity represented on the z axis.

2.5.4 Retrieving Diracs from sectoral SH coefficients

In [51], the locations and amplitudes of the Diracs are retrieved from the *sectoral* spherical harmonics coefficients, which are the coefficients where $m = \ell$:

$$\widehat{f}_{\ell,\ell} = \widetilde{N}_{\ell,\ell} \sum_{k=0}^{K-1} c_k (\sin\theta_k)^\ell e^{-j\ell\phi_k}. \tag{2.14}$$

Here, $\widetilde{N}_{\ell,\ell}$ is a constant that encompasses $N_{\ell,\ell}$ as well as the appropriate associated Legendre coefficient. Since $\widetilde{N}_{\ell,\ell}$ is known and nonzero, we can divide the expression in (2.14) by it to obtain

$$g_\ell = \frac{\widehat{f}_{\ell,\ell}}{\widetilde{N}_{\ell,\ell}} = \sum_{k=0}^{K-1} c_k u_k^\ell, \tag{2.15}$$

where $u_k = \sin\theta_k e^{-j\phi_k}$. We recognize in (2.15) the familiar expression of a sum of complex exponentials, which can be annihilated using Prony's method to recover the roots $\{u_k\}_{k=0}^{K-1}$. To find the annihilating filter, we need the bandwidth L of the signal to be at least $2K$.

2.5.5 Retrieving VPW pulses from sectoral SH coefficients

We now extend the result of [51], adding width to the Diracs. Like the Cartesian case, we can introduce a damping factor $0 \leq r_k \leq 1$, which reduces the magnitude of the signal at higher frequencies. The spherical VPW pulses are thus defined as

$$\widehat{f}_{\ell,m} = N_{\ell,m} \sum_{k=0}^{K-1} c_k P_{\ell,|m|}(\cos\theta_k) e^{-jm\phi_k} e^{-r_k(\ell+|m|)}. \tag{2.16}$$

An example of a spherical VPW signal composed of two pulses is shown in Figure 2.12. The following theorem shows how to retrieve the parameters of this new signal model.

THEOREM 2.4 Let f be the sum of K continuous spherical VPW pulses as in (2.16). The spherical harmonics coefficients $\widehat{f}_\ell^m$ with $m = 0, 1, \ldots, 2K-1$ and $|\ell| < m$ are sufficient to recover f.

PROOF. As in [51], we can compute

$$g_\ell = \frac{\widehat{f}_{\ell,\ell}}{\widetilde{N}_{\ell,\ell}} = \sum_{k=0}^{K-1} c_k u_k^\ell, \tag{2.17}$$

where we now have

$$\widehat{f}_{\ell,\ell} = \widetilde{N}_{\ell,\ell} \sum_{k=0}^{K-1} c_k (\sin\theta_k)^\ell e^{-j\ell\phi_k} e^{-2r_k\ell}.$$

The roots, which can be found via the annihilating filter method, become $u_k = \sin\theta_k e^{-i\phi_k} e^{-2r_k}$. We can extract the azimuth ϕ_k from the phase of u_k, but the magnitude of u_k depends on both θ_k and r_k and at this point we cannot resolve the ambiguity. What we can do however is find the amplitude c_k by solving the linear system in (2.17). Furthermore, for $m = \ell - 1$ and $\ell \geq 1$, we have

$$\widehat{f}_{\ell,\ell-1} = \widetilde{N}_{\ell,\ell-1} \sum_{k=0}^{K-1} c_k e^{-r_k} \cos\theta_k u_k^{\ell-1},$$

$$g_{\ell-1} = \frac{\widehat{f}_{\ell,\ell-1}}{\widetilde{N}_{\ell,\ell-1}} = \sum_{k=0}^{K-1} c_k b_k u_k^{\ell-1},$$

where $b_k = e^{-r_k}\cos\theta_k$. Note that, since we know $\widehat{f}_\ell^{\ell-1}$ and $u_k^{\ell-1}$, we can compute b_k. Moreover, the knowledge of the azimuth ϕ_k enables us to calculate

$$a_k = \frac{u_k}{e^{-j\phi_k}} = \sin\theta_k e^{-2r_k}.$$

Finally, combining a_k and b_k and assuming a_k is nonzero, we compute the width

$$r_k = \frac{1}{2}\log\left(-b_k^2 + \frac{1}{2a_k^2}\sqrt{4a_k^2 + b_k^4}\right) \tag{2.18}$$

by keeping only the positive solution to the square root. Observe that b_k is raised to the 4th power and a_k is squared. Furthermore, in the noiseless case[7], both terms are real, so the expression under the square root is always positive. Using (2.18), we obtain $\theta_k = \arccos\left(\frac{b_k}{e^{r_k}}\right)$. ■

Like [51], our method suffers from a relatively poor spectrum usage. Indeed, only a small number of the spherical harmonics coefficients are used for the parameters' estimation.[8] In Chapter 3, we address this shortcoming by adding a nonlinear optimization routine that uses the entire spectrum to refine the VPW estimation. VPW-FRI can be used to model a wide range of signals. In the next chapter, we see how to use it to sample and represent reflectance distribution functions. Another interesting example is the compression of electrocardiogram signals, as explained below.

[7] In the noisy case, it is possible that b_k and a_k are complex. To prevent negative terms in the square root, we project them onto the real axis.

[8] A better spectrum usage was proposed in [52], but we have been unable to generalize their method to VPW.

2.6 Example: ECG compression

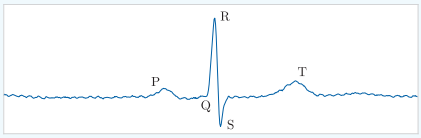


Figure 2.13: Example of an ECG pulse.

Several signals can be approximated using VPW pulses; a good example consists of the problem of sampling and representing electrocardiogram (ECG) signals. ECG signals are composed of five types of pulses, labeled P, Q, R, S, T in the medical literature (see also Figure 2.13). Each of these pulses represents a different stage in the activity of the heart and each has a different width and amplitude; while the P and T pulses are fairly wide, the Q,R,S pulses are narrow and close to each other. Having a signal model with different pulse widths is paramount here.

We briefly show how to to use VPW-FRI to compress ECGs. For our experiments, we use ECG records from the MIT-BIH database [71]. We compare the VPW-FRI pulse model with a Gaussian model, as is for example described in [26]. Each Gaussian pulse is expressed with 3 parameters: location, scale and amplitude. We fit them to the original signal using nonlinear optimization techniques. Additionally, we compare the results with ideal lowpass filtering followed by downsampling. Figure 2.14 shows an example of a compressed reconstruction. On the single record of Figure 2.14, VPW-FRI achieved a signal-to-residual ratio (SRR) of 22.3 dB compared to 14.0 dB for the Gaussian pulse model. We also observe that the Gaussian model sometimes fails to identify narrow pulses. Another disadvantage of this technique is that it is difficult to initialize the nonlinear optimization. In fact, in the results presented here, we initialized it with the VPW-FRI pulse locations to make it converge faster.

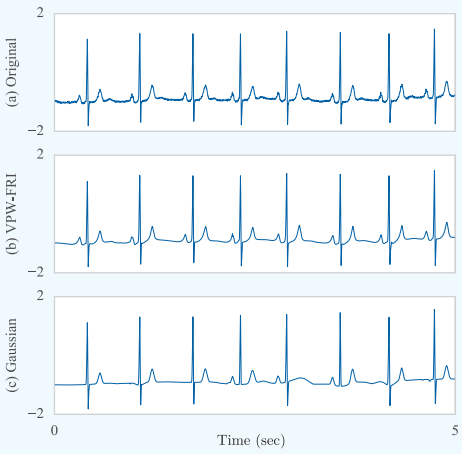


Figure 2.14: ECG compression on record 123 from the MIT-BIH database: (a) original ECG signal; (b) VPW estimation with 5 pulses for each heart beat; (c) estimation with 7 Gaussians.

2.7 Conclusion

We have presented variable pulse width, a new parametric model that fits within the FRI framework and allows for a wider range of pulse shapes. We proposed a method for estimating the parameters and quantified its performance on noisy signals. Moreover, we showed that VPW can be extended to 2D and spherical signals. In 2 dimensions, we proposed several pulse shapes and discussed their characteristics. Unlike in 1D, the spectrum usage is not optimal in both 2D and on the sphere; we need to sample at a rate quadratically higher than the rate of innova-

tion to ensure perfect recovery. Another limitation is that the spherical case allows only radial pulses. Nevertheless, this model is applicable to a broad range of signals, such as for example electrocardiograms. The next chapter showcases applications of the 2D and spherical versions of VPW-FRI to sample and represent reflectance functions.

2.A Proof of Lemma 2.1

Since $f(t)$ is periodic, with period τ, it has the following Fourier series expansion:

$$f(t) = \sum_{m\in\mathbb{Z}} F[m]e^{2\pi jmt/\tau}.$$

From the sampling scheme in Figure 2.3,

$$\begin{aligned} g[n] &= \langle f(t), \varphi\left(t - nT\right)\rangle \\ &= \sum_{m\in\mathbb{Z}} F[m]\left\langle e^{2\pi jmt/\tau}, B\,\mathrm{sinc}\left(B(t-nT)\right)\right\rangle \\ &= \sum_{m\in\mathbb{Z}} F[m]e^{2\pi jmnT/\tau}\int_{-\infty}^{\infty}\mathrm{sinc}\left(t\right)e^{2\pi jmt/(B\tau)}dt \\ &= \sum_{m=-\lfloor B\tau/2\rfloor}^{\lfloor B\tau/2\rfloor} F[m]e^{2\pi jmn/N}. \end{aligned}$$

This is a system of N equations in $2\lfloor B\tau/2\rfloor + 1$ unknown Fourier series coefficients $F[m]$, which can be inverted when $N \geq 2\lfloor B\tau/2\rfloor + 1$. When $N = 2\lfloor B\tau/2\rfloor + 1$, the system is, up to a $1/N$ factor, the $N \times N$ inverse-DFT of $F[m]$. Therefore,

$$F[m] = NG[m],$$

for $|m| \leq \lfloor B\tau/2\rfloor$.

2.B Time-domain formulae for VPW pulse

We show here the details of the derivation of the time-domain formulae for the VPW pulse. The time-domain representation of F_k^s (cf. (2.8)) is computed as

$$\begin{aligned} f_k^s(t) &= \sum_{m=-\infty}^{\infty} F_k^s[m]e^{j2\pi mt/\tau} \\ &= \frac{1}{\tau}\left(\sum_{m=0}^{\infty} z_k^m(t) + \sum_{m=0}^{\infty}(z_k^*(t))^m - 1\right) \\ &\overset{(a)}{=} \frac{1}{\tau}\left(\frac{1}{1-z_k(t)} + \frac{1}{1-z_k^*(t)} - 1\right) \\ &= \frac{1}{\tau}\frac{1-|z_k(t)|^2}{(1-z_k(t))(1-z_k^*(t))}, \end{aligned}$$

where $z_k(t) = e^{2\pi/\tau(-r_k+j(t-t_k))}$. Note that (a) assumes that $r_k > 0$, which only holds when the annihilating filter is *stable* (i.e., when its roots are inside the unit circle).

We can do the same reasoning for the asymmetric component F_k^a of the signal, whose time-domain representation is

$$\begin{aligned} f_k^a(t) &= \sum_{m=-\infty}^{\infty} F_k^a[m] e^{j2\pi mt/\tau} \\ &= \frac{i}{\tau}\left(\sum_{m=0}^{\infty} z_k^m(t) - \sum_{m=0}^{\infty} (z_k^*(t))^m + 1\right) \\ &\overset{(d)}{=} \frac{i}{\tau}\left(\frac{1}{1-z_k(t)} - \frac{1}{1-z_k^*(t)}\right) \\ &= \frac{1}{\tau}\frac{2\,\mathfrak{Im}\{z_k(t)\}}{(1-z_k(t))(1-z_k^*(t)}. \end{aligned}$$

Again, (d) only holds for $r_k > 0$.

2.C Computation of the Cramer-Rao bound for VPW signals

Let $\boldsymbol{\Theta} = \left[t_0, r_0, c_0\right]^\top$ be the vector of unknown parameters. The CRB relies on the Fisher information matrix, which is given by (cf. [54] for more details)

$$\begin{aligned} \boldsymbol{I}(\boldsymbol{\theta}) &= \mathbb{E}\left[\left(\frac{\partial \log \widetilde{f}[n])}{\partial \boldsymbol{\theta}}\right)\left(\frac{\partial \log \widetilde{f}[n])}{\partial \boldsymbol{\theta}}\right)^\top\right] \\ &= \frac{1}{\sigma_\epsilon^2}\sum_{n=0}^{N-1} \nabla f[n] \nabla f[n]^\top. \end{aligned}$$

Note that $f[n]$ and $\widetilde{f}[n]$ also depend on $\boldsymbol{\theta}$, but we have omitted it to simplify the notation.

Given any unbiased estimate $\widehat{\boldsymbol{\theta}}$ of the parameters, the CRB provides a lower bound on the covariance matrix of $\widehat{\boldsymbol{\theta}}$:

$$\operatorname{cov}(\widehat{\boldsymbol{\theta}}) \succeq \boldsymbol{I}^{-1}(\boldsymbol{\theta}).$$

Let our signal $\boldsymbol{f} = \left[f[0], f[1], \ldots, f[N-1]\right]^\top$ and its DFT $\widehat{\boldsymbol{f}} = \left[F[0], F[1], \ldots, F[N-1]\right]^\top$. Since $\boldsymbol{f} = \boldsymbol{W}^* \widehat{\boldsymbol{f}}$—where $\boldsymbol{W}$ is the DFT matrix of size N—and we have a closed-form formula for $F[m]$, we choose to first compute the derivatives for the gradient in the frequency domain and then perform the inverse DFT to get $\nabla f[n]$. If we restrict ourselves to positive m and a single VPW pulse, we have, given $F[m] = c_0 e^{-2\pi m(r_0+jt_0)/\tau}$,

$$\begin{aligned} \frac{\partial F[m]}{\partial t_0} &= -c_0 \frac{2\pi jm}{\tau} e^{-2\pi m(r_0+jt_0)/\tau} \\ \frac{\partial F[m]}{\partial r_0} &= -c_0 \frac{2\pi m}{\tau} e^{-2\pi m(r_0+jt_0)/\tau} \\ \frac{\partial F[m]}{\partial c_0} &= e^{-2\pi m(r_0+jt_0)/\tau}. \end{aligned}$$

Chapter 3

Corpuscular Theory: Acquisition of Reflectance Functions*

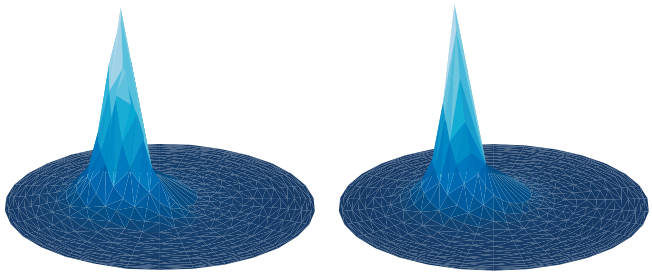
It is exciting to think that it costs nothing to create a new particle.

Lectures on Gravitation
Richard Feynman

3.1 Introduction

We demonstrate in this chapter via four short applications how the finite rate of innovation (FRI) theory and the variable pulse width (VPW) model can be exploited to acquire and estimate the bidirectional reflectance distribution function (BRDF), which is a high-dimensional function that characterizes how light is reflected by surfaces. Usually, BRDFs contain high frequency components—called *specularities*—that are difficult to catch with classical sampling theory. Moreover, their shape varies through space, which makes existing FRI models inappropriate. The acquisition of BRDFs in the spherical domain requires dense sampling of the hemisphere, as can be seen in [115]. More elaborate techniques propose to perform the acquisition in a *transform*

*The material in this chapter is the result of joint work with Adam Scholefield, Ivan Domanić and Martin Vetterli. Author contributions: GB, AS, ID and MV designed the extension from Chapter 2 to make it suitable for reflectance functions; GB performed research under the guidance of AS and MV; GB ran simulations and experiments, and wrote the chapter based on [7, 10], which were written by himself.



(a) Original (b) VPW model

Figure 3.1: Bidirectional reflectance distribution function (BRDF) representation in 2D: (a) example of a measured BRDF and (b) its representation using a single VPW pulse.

domain: for example in [183], the acquisition is done in the spherical harmonics domain, which substantially reduces the number of samples. The disadvantage is that it requires a specialized gantry, with a high number of light sources that can recreate the spherical harmonics functions. In this chapter, we leverage FRI theory to reduce the number of samples directly in the spherical domain. An example of a specularity approximated with a single spherical VPW pulse is shown in Figure 3.1; here, the direction of outgoing light is kept fixed.

By utilizing a simple parametric model, we are enforcing a strong prior that is particularly beneficial in certain scenarios. For example, in the BRDF example, it is typical to only have a few samples of sharp specularities; it is thus very challenging to obtain an accurate representation of the underlying reflectance function. Despite this, the proposed approach is able to recover a good approximation of these specularities, and the continuous representation enables us to later render the object from any illumination direction. This is in contrast to super-resolution approaches, where a denser sampling of the BRDF would only allow us to render the object under certain illumination directions or would force us to apply some other type of interpolation. In fact, graphics renderers deal with parametric forms of the reflectance function precisely for these reasons.

We start with a brief review of the BRDF definitions and existing models and then propose FRI to facilitate the acquisition of specularities. At first, we suggest to use diffusers to blur and widen specularities; this is equivalent to convolving the specularity with an *unknown* sampling kernel. We show that we can adapt the FRI framework to such situations.

Next, we describe three experiments to justify the use of VPW-FRI; each of them focuses on a different domain (1D, 2D and spherical) and makes use of several data sources: synthetic data, real data from a publicly-available dataset, as well as our own acquired data using a custom-made gantry.

3.2 Background

The illumination has a strong influence on the appearance of objects: for example, a shiny material might look brighter if light is directed at it from a particular angle and darker otherwise. The role of the BRDF [135] is precisely to describe the reflective characteristics of materials; it

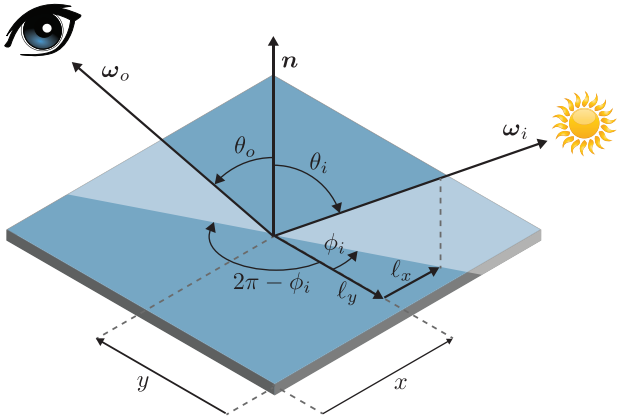


Figure 3.2: Vectors involved in the BRDF function: $\boldsymbol{\omega}_i$ is the incoming light direction, $\boldsymbol{\omega}_o$ is the outgoing light direction, and $\boldsymbol{n}$ is the surface normal. The vectors ℓ_x and ℓ_y represent the projection of $\boldsymbol{\omega}_i$ onto the xy-plane.

enables us to decouple the reflectance properties of objects from the illumination.

DEFINITION 3.1 (BRDF) The BRDF ρ is a four-dimensional function that measures the radiance L_o that is reflected from a specific point in a direction $\boldsymbol{\omega}_o = [\theta_o, \varphi_o]^\top$, given a light source with irradiance E_i coming from a direction $\boldsymbol{\omega}_i = [\theta_i, \varphi_i]^\top$.

$$\rho_\lambda(\boldsymbol{\omega}_i, \boldsymbol{\omega}_o) = \frac{dL_o(\boldsymbol{\omega}_o)}{dE_i(\boldsymbol{\omega}_i)} = \frac{dL_o(\boldsymbol{\omega}_o)}{L_i(\boldsymbol{\omega}_i)\cos\theta_i d\boldsymbol{\omega}_i}.$$

The directions $\boldsymbol{\omega}_o$ and $\boldsymbol{\omega}_i$ are measured with respect to the surface normal $\boldsymbol{n}$ of the material at the point of interest. The BRDF is specific to the wavelength, and in practice we often use a separate reflectance function for each spectral band; for the sake of clarity, we focus here on a single channel. The BRDF has the following properties and characteristics:

(1) **Positivitiy:** The BRDF is positive.

$$\rho(\boldsymbol{\omega}_i, \boldsymbol{\omega}_o) \geq 0, \qquad \forall \boldsymbol{\omega}_i, \boldsymbol{\omega}_o.$$

(2) **Helmholtz reciprocity:** Light traveling along the same path but in the opposite direction is reflected in the same way.

$$\rho(\boldsymbol{\omega}_i, \boldsymbol{\omega}_o) = \rho(\boldsymbol{\omega}_o, \boldsymbol{\omega}_i).$$

(3) **Conservation of energy:** The quantity of light reflected must be less than or equal to the quantity of incident light:

$$\int_\Omega \rho(\boldsymbol{\omega}_i, \boldsymbol{\omega}_o)\cos\theta_i d\boldsymbol{\omega}_i \leq 1, \qquad \forall \boldsymbol{\omega}_o.$$

Remark that some of the light hitting the surface might be absorbed by the material or refracted.

(a) Acquisition (b) Rendering

Figure 3.3: BRDF acquisition and rendering: (a) a close-up of a US penny under a specifc illumination and (b) a rendering of the same coin under a different artificial illumination.

(4) **Isotropy:** *Some* BRDFs are said to be *isotropic*, that is, they are unchanged if the incoming and outgoing vectors are rotated by the same amount about the surface normal.

There exist different generalizations of the BRDF: among others, the spatially varying bidirectional reflectance distribution function (SVBRDF) takes into account spatial variations and the bidirectional texture function (BTF) encompasses the SVBRDF parameters and adds occlusions and shadows. Both functions are six-dimensional.

3.2.1 The rendering equation

To understand how the BRDF transforms the incoming light, we need one more ingredient: the *rendering equation*. It quantifies the radiance L_o emitted from a point along the direction $\boldsymbol{\omega}_o$ by integrating the product of the BRDF ρ and the radiance of the incoming light L_i for all possible directions $\boldsymbol{\omega}$ on the hemisphere Ω:

$$L_o(\boldsymbol{\omega}_o) = \int_\Omega L_i(\boldsymbol{\omega})\rho(\boldsymbol{\omega}, \boldsymbol{\omega}_o)\langle \boldsymbol{n}, \boldsymbol{\omega}\rangle d\boldsymbol{\omega}. \tag{3.1}$$

The inner product $\langle \boldsymbol{n}, \boldsymbol{\omega}\rangle = \cos(\theta)$, which is often absorbed in ρ, is called the *normal attenuation* or *cosine term*. It accounts for the fact that objects appear darker when $\boldsymbol{\omega}$ is closer to grazing angle.

Despite its generality, the rendering equation does not take into consideration every aspect of the light's behavior. For instance, it does not capture inter-reflections, transmission, subsurface scattering or polarization. Nevertheless, (3.1) is a good approximation of the physics of light without being too complex. For example, the full knowledge of the BRDF or SVBRDF combined with the rendering equation allow the creation of realistic scenes under new illuminations and views. Figure 3.3 shows an example of a coin whose SVBRDF has been captured under several different illuminations; in Figure 3.3b, the same coin has been rendered under a new different illumination[1].

[1] A video of the re-illumination of the coin can be found at `http://go.epfl.ch/relight`.

3.2.2 BRDF acquisition

To acquire the BRDF of an object, we need to illuminate and observe it from several angles: for each pair $(\boldsymbol{\omega}_i, \boldsymbol{\omega}_o)$, we measure the reflected radiance. This process is often performed with a *goniophotometer*—a device that takes several photographs of a material under different illuminations; for instance in [115], almost 3 million BRDF samples have been acquired for 100 different materials that now compose the publicly available MERL BRDF database. The high dimensionality of BRDFs and SVBRDFs often results in huge datasets, and it motivates the need for lower-dimensional approximations and reparametrizations.

One of the approaches to reduce the number of samples needed in BRDF acquisition while still faithfully capturing the specularity is to extend the area illuminated by the light source. The earliest example of extended light sources shaped the radiance of the light as a Gaussian distribution and modeled the BRDF with a cosine and a Dirac pulse [129]. More recently, [65] suggested acquiring spatially-varying BRDFs of an almost flat surface using a *linear* light source and by only varying the inclination of the incoming light θ_i. Light sources can also take more elaborate patterns, and enable BRDF acquisition in a different domain, such as the spherical harmonic domain [183] or the Fourier domain [2].

It is also common to represent the incoming and outgoing light directions in the half-angle coordinate system [159]. With this parametrization, the BRDF is specified in terms of the half-angle $\boldsymbol{h}$ and the difference $\boldsymbol{d}$: $\boldsymbol{h} = (\theta_h, \varphi_h) = \frac{\boldsymbol{\omega}_i + \boldsymbol{\omega}_o}{\|\boldsymbol{\omega}_i + \boldsymbol{\omega}_o\|}$ is the angle that is half way between $\boldsymbol{\omega}_i$ and $\boldsymbol{\omega}_o$, while $\boldsymbol{d} = (\theta_d, \varphi_d)$ is the angle between $\boldsymbol{h}$ and $\boldsymbol{\omega}_i$. The advantage of this representation is that it is independent of φ_d for *isotropic materials*. This is the approach taken in [115] to reduce the dataset size to three dimensions. Other simplifications consist of fixing the viewing angle or projecting the vector $\boldsymbol{\omega}_i$ onto the xy-plane, as in [112]: in Figure 3.2, this projection is illustrated by the two components ℓ_x and ℓ_y.

3.2.3 Light domes

While powerful, most goniophotometers do not take into account spatial variations. To address this need, we constructed our own *light domes*. Light domes consist of a combination of light sources and an imaging device to capture slices of the SVBRDF. In our light domes, the light sources are placed on a hemisphere surrounding the object to be analyzed and a DSLR camera is located at the zenith. The digital camera enables us to capture the spatial variations of the SVBRDF. We developed two flavors of light domes.

1. The first design consists of 58 fixed point light sources spread on the hemisphere (see Figure 3.4a). Its conception is similar to other light domes that have been built by other research groups [48]. The position of the camera and the object to be analyzed is fixed, thus we give away outgoing angle information in exchange for spatial information. The advantage of such a design is that it is fast to turn on and off the lights, making SVBRDF acquisitions relatively quick. It also makes it possible to perform acquisitions with several light sources on at the same time. This light dome design is also the genesis of the startup *Artmyn*[2], which originated from our research group and specializes in the acquisition of the reflectance function of artworks.

2. Our second light dome consists of an extensible motorized arm with a light source at its extremity (see Figure 3.4b). The arm can be rotated around the center, where a camera

[2] https://artmyn.com/

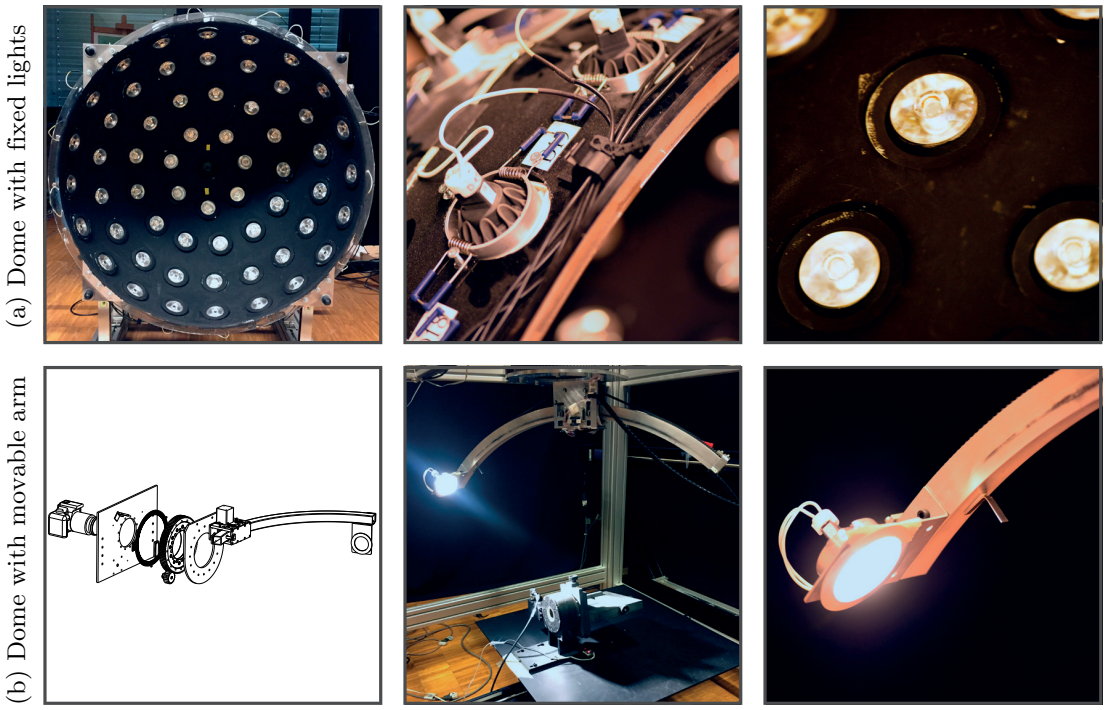


Figure 3.4: Design and photographs of our light domes: (a) dome with 58 fixed lights, and (b) dome with movable arm and arbitrary light positions.

is installed, which enables us to position the incoming light at almost any location on the hemisphere. This light dome started as a student project [6, 124, 175], and its design has been substantially upgraded over the years. It now has a sturdy chassis and can be mounted with several different cameras, including a hyperspectral line scanner. Additionally, a rotating platform that holds the objects of interest can be mounted on the chassis (see Figure 3.4b, middle): this allows us to orient the objects along different angles and to acquire outgoing angle information. The advantages of having a moving arm holding the light source are multiple. First, it makes it possible to sample the incoming light hemisphere with an arbitrary resolution, while with fixed lights, the sampling patterns are imposed by the position of the light. Second, it enables us to sample the BRDF along trajectories of light: this can be done by performed acquisitions with longer exposure times while moving the arm. In the same vein, we can also use the trajectories of the light to simulate some extended light sources. Finally, we can also imagine a dynamic sampling scheme that adapts based on the points of the SVBRDF that have already been measured.

3.2.4 BRDF shape and models

In its elementary form, a BRDF is decomposed into two parts: the *diffuse* (or *Lambertian*) part—a constant term that represents the amount of light that is equally reflected in all directions—and the *specular* component—a peak whose direction has the same angle as the incoming light with

respect to the surface normal. In mathematical terms,

$$\rho(\boldsymbol{w}_i, \boldsymbol{w}_o) = \underbrace{\frac{k_d}{\pi}}_{\text{diffuse}} + \underbrace{k_s \rho_s(\boldsymbol{w}_i, \boldsymbol{w}_o)}_{\text{specular}},$$

where k_d is the diffuse albedo, k_s the specular albedo, and ρ_s is a function of a few arguments that represents the shape of the specular peak. A wide range of parametric models have been proposed to represent ρ_s [115]; one of the most flexible and accurate [134] being the Cook-Torrance model [44], also known as the *microfacet model*. It is given by

$$\rho_s(\boldsymbol{w}_i, \boldsymbol{w}_o) = \frac{F(\boldsymbol{\omega}_o, \boldsymbol{h}) G(\boldsymbol{\omega}_i, \boldsymbol{\omega}_o, \boldsymbol{h}) D(\boldsymbol{h})}{4\langle \boldsymbol{\omega}_i, \boldsymbol{n} \rangle \langle \boldsymbol{\omega}_o, \boldsymbol{n} \rangle},$$

where the function F is called the *Fresnel factor*, G is the *geometry term*, and D is the *microfacet distribution*. While F and G are fairly constant throughout the range of values of θ_h (except near grazing angle, where they exhibit a particular behavior), D is the factor we are most interested in since it describes the global shape of the specularity. Several shapes have been proposed for D; two of the most widely used are the Beckmann [194] and the GGX [192] distributions:

$$D_{Beckmann}(\boldsymbol{h}) = \frac{1}{\pi \alpha^2 \langle \boldsymbol{n}, \boldsymbol{h} \rangle^4} \exp\left(-\frac{\tan(\theta_h)^2}{\alpha^2}\right),$$

$$D_{GGX}(\boldsymbol{h}) = \frac{\alpha^2}{\pi \langle \boldsymbol{n}, \boldsymbol{h} \rangle^4 (\alpha^2 + \tan(\theta_h)^2)^2},$$

where θ_h is the angle between $\boldsymbol{h}$ and $\boldsymbol{n}$, and α is a *roughness parameter*. Observe that the microfacet distribution is 1-dimensional in these models: it depends only on $\tan(\theta_h)$.

Another popular and relatively simple model to represent specularities is due to Ward [194]. It represents specularities with elliptical-Gaussian distribution function

$$\rho_{s,Ward}(\boldsymbol{\omega_i}, \boldsymbol{\omega_o}) = \frac{k_s}{4\pi\alpha^2 \sqrt{\langle \boldsymbol{n}, \boldsymbol{\omega_i} \rangle \langle \boldsymbol{n}, \boldsymbol{\omega_o} \rangle}} \exp\left(-\frac{\tan(\theta_h)^2}{\alpha^2}\right).$$

Shortly, we will investigate the use of FRI and the VPW model studied in Chapter 2 to model ρ_s.

3.3 FRI with unknown sampling kernel

Our first attempt at adapting the FRI sampling framework to the problem of BRDF acquisition is based on extended light sources. We model the specularity as a Dirac pulse $x(t)$ located at t_0. Since most BRDFs contain only a single specularity, we restrict ourself to a single pulse model. The Fourier series coefficients of $x(t)$ are simply given by

$$X[m] = \frac{1}{\tau} a_0 e^{-i2\pi t_0 m/\tau}, \qquad m \in \mathbb{Z}.$$

As seen in Chapter 2, one of the key elements in the FRI theory is the search for an annihilating filter A such that $(A * X)[m] = 0, \forall m \in \mathbb{Z}$. For a single pulse, the annihilating filter coefficients are found by solving the linear system of equations

$$A[0]X[m] + A[1]X[m-1] = 0,$$

for $m = 1, \ldots, M-1$ and $M \geq 2$. The single root u_0 of this filter satisfies

$$A(z) = A[0] + A[1]z^{-1} = 1 - u_0 z^{-1}, \tag{3.2}$$

where $A(z)$ is the z-transform of A.

Now assume that $x(t)$ is convolved with an *unknown* τ-periodic filter $\varphi(t)$ that represents the specularity shape combined with the shape of the light source illuminating the object. We impose the constraint that $\varphi(t)$ must be *symmetric* (the reason for this is made clear later). Let $y(t) = (x * \varphi)(t)$. By the convolution theorem, the Fourier series coefficients of $y(t)$ are given by

$$Y[m] = \frac{1}{\tau} a_0 u_0^m \Phi[m], \qquad m \in \mathbb{Z},$$

where $\Phi[m]$ are the Fourier series coefficients of $\varphi(t)$.

As in (3.2), we can find an annihilating filter for $Y[m]$ of length two: $A[0] = 1$ and $A[1] = -v_0$, where v_0 is the root of this new filter. Using $M \geq 2$ Fourier series coefficients, we can estimate v_0 by solving

$$\boldsymbol{y_1} = v_0 \boldsymbol{y_0}, \tag{3.3}$$

with

$$\boldsymbol{y_0} = \begin{bmatrix} Y[0] \\ Y[1] \\ \vdots \\ Y[M-2] \end{bmatrix} \text{ and } \boldsymbol{y_1} = \begin{bmatrix} Y[1] \\ Y[2] \\ \vdots \\ Y[M-1] \end{bmatrix}.$$

The closest approximation of v_0 in (3.3) is simply given by the orthogonal projection of $\boldsymbol{y_1}$ onto $\boldsymbol{y_0}$:

$$\begin{aligned} v_0 &= \frac{\boldsymbol{y_0}^* \boldsymbol{y_1}}{\boldsymbol{y_0}^* \boldsymbol{y_0}} \\ &= \frac{\sum_{m=1}^{M-1} \left(\alpha_0 \varphi[m-1] u_0^{m-1}\right)^* \left(\alpha_0 \varphi[m] u_0^m\right)}{\sum_{m=0}^{M-2} \left(\alpha_0 \varphi[m] u_0^m\right)^* \left(\alpha_0 \varphi[m] u_0^m\right)} \\ &= \underbrace{\frac{\sum_{m=1}^{M-1} \left(\varphi[m-1]^* \varphi[m] |u_0^{m-1}|^2\right)}{\sum_{m=0}^{M-2} \left(|\varphi[m]|^2 |u_0^m|^2\right)}}_{\text{purely real}} u_0. \end{aligned}$$

The key observation is that, since the sampling kernel is assumed to be symmetric, its Fourier series coefficients are real. Hence the expression multiplying u_0 is a real quantity, which implies that $\angle v_0 = \angle u_0$. In other words, the filter annihilating $Y[m]$ allows us to retrieve the location t_0 from its root as long as the unknown filter $\varphi(t)$ is symmetric.

As a final remark, note that even though the kernel shape is not necessary in the estimation of the location, it is needed for the recovery of the amplitude a_0: it is simply impossible to differentiate the contribution of the kernel $\varphi(t)$ from the contribution of the signal $x(t)$ by observing $y(t)$ only.

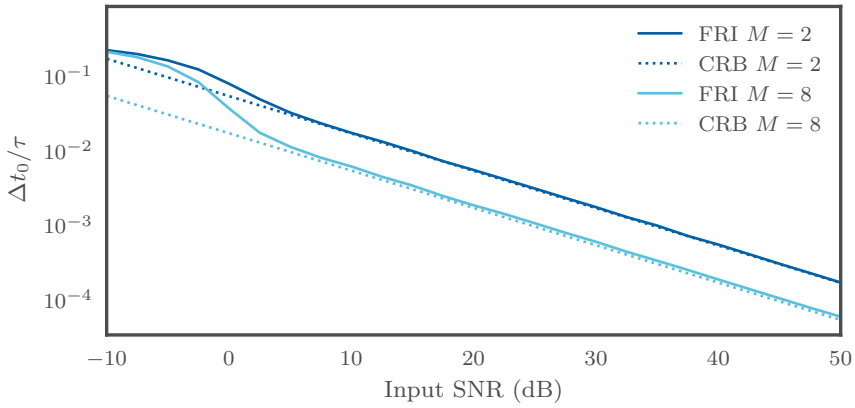


Figure 3.5: Sampling a single noisy τ-periodic Dirac convolved with a Gaussian kernel ($\tau = 1, \sigma = 0.05$). Using $2M = 2 \cdot 2$ and $2 \cdot 8$ samples, we measure the disparity $\Delta t_0/\tau$ between the FRI-estimated $\hat{t}_0$ and t_0 for a given input SNR (solid lines) and compare it with the CRB (dotted lines). Each data point is the average of $10,000$ repetitions.

Extension to 2D signals

As in Section 2.4, the generalization of the above result to 2-dimensional signals is straightforward if we assume that the sampling kernel is separable; i.e., $\varphi = \varphi_x * \varphi_y$, where φ_x is a horizontal and φ_y a vertical filter (see [113] for more details). The separability of the sampling kernel combined with the convolution theorem yields the following expression for the 2D Fourier series coefficients:

$$Y[m,n] = \frac{1}{\tau^2} a_0 e^{-i2\pi x_0 m/\tau} e^{-i2\pi y_0 n/\tau} \varphi_x[m]\varphi_y[n], \quad m,n \in \mathbb{Z},$$

where (x_0, y_0) is the location of the Dirac pulse and a_0 its amplitude. We can reduce the problem to the 1-dimensional case by selecting the coefficients $Y[m,0]$ for $m \in \mathbb{Z}$:

$$Y[m,0] = \frac{1}{\tau^2} A_0 e^{-i2\pi x_0 m/\tau} \varphi_x[m],$$

with $A_0 = a_0\varphi_y[0]$. In the same manner, we can estimate y_0 using the coefficients $Y[0,n]$ for $n \in \mathbb{Z}$. We need $M \geq 2$ Fourier coefficients in each dimension to find the coordinates (x_0, y_0).

3.3.1 Noisy case

We now assess the performance of the FRI estimation procedure in the presence of noise with the following simulation. We generate a single Dirac with a random location $t_0 \in [0,\tau)$, convolved with a Gaussian kernel of variance $\sigma^2 = 0.05$. We sample $y(t)$ as

$$y_s[n] = \langle y(t), h\left(t/T - n\right)\rangle \qquad n = 0, 1, \ldots, 2M-1,$$

where $T = \frac{\tau}{2M}$ and $h(t)$ is an ideal lowpass filter whose role is to avoid aliasing in the low-frequency Fourier coefficients. We then corrupt the samples $y_s[n]$ with additive white Gaussian noise (AWGN) of variance σ_ϵ^2. We compare the efficiency of the FRI estimation procedure with the Cramér-Rao bound; given $2M$ discrete uniform samples, the variance $\mathrm{var}(\hat{t}_0)$ of any unbiased

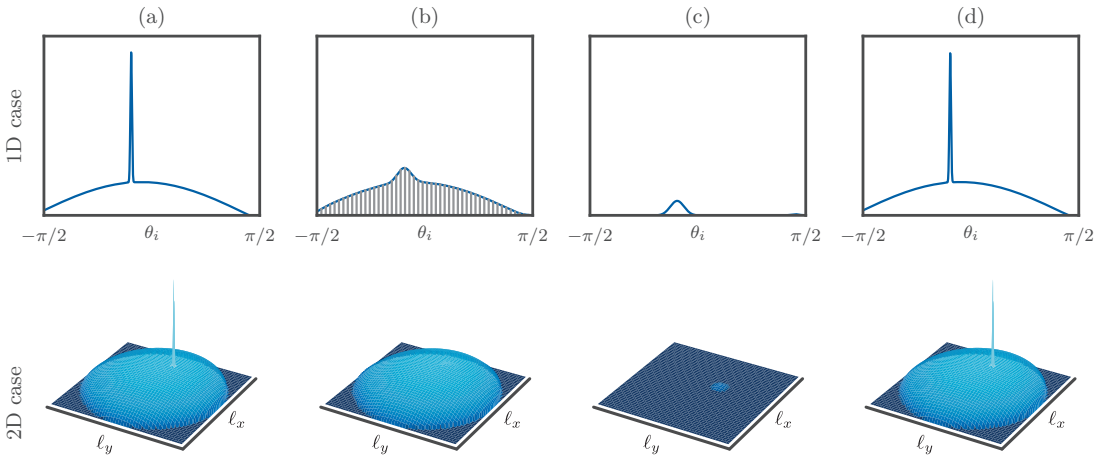


Figure 3.6: Example of the estimation of the BRDF parameters with FRI in the 1D (top) and 2D (bottom) cases. (a) original BRDF ρ; (b) signal lowpass-filtered with a Gaussian kernel ($\sigma = 0.5$ in 1D and $\Sigma = [0.3, 0; 0, 0.3]$ in 2D) and sampled; (c) lowpass specularity after subtraction of the diffuse part; (d) Reconstructed signal: the specularity is estimated with FRI and the diffuse component with photometric stereo. The parameter θ_i represents the inclination of the incoming light vector and (ℓ_x, ℓ_y) are the coordinates of its projection onto the xy-plane.

estimator $\hat{t}_0$ of t_0 is lower-bounded by $\mathrm{CRB}(\hat{t}_0)$. In the AWGN case and with a single unknown parameter, the CRB is given by

$$\mathrm{CRB}(\hat{t}_0) = \frac{\sigma_\epsilon^2}{\sum_{n=0}^{2M-1} \left(\frac{\partial y_s[n]}{\partial t_0} \right)^2}.$$

Results of the simulation can be seen in Figure 3.5 for $M = 2$ and 8; we consider the FRI-based estimator to be fully efficient for an SNR above 10dB. Note that our FRI implementation is coupled with ESPRIT [158] to reduce the effect of the noise.

3.3.2 Application to specularity estimation

We show here how to apply the FRI theory to reconstruct a BRDF from its filtered version and discuss the technicalities of the application of the algorithm.

First, we reduce the number of dimensions of the BRDF from 4 to 2 by fixing the viewing angle $\boldsymbol{\omega}_o$. Furthermore, we project the normalized incoming light vector $\boldsymbol{\omega}_i$ onto the xy-plane, so that we can work in a Euclidean space, where our theory has been developed. We call the 2 components of the projection ℓ_x and ℓ_y (cf. Figure 3.2).

As stated in Section 3.2.2, the extended light source approach allows us to smooth the specularity. The resulting acquired signal can be seen as the convolution of a Dirac pulse with two kernels: its (narrow) pulse shape and the extended light kernel. While we have some control on the shape of the extended light kernel (if it is not known, we can at least measure it), the specular pulse shape is *unknown*. Fortunately, both kernels are symmetric, which makes the specular peak a suitable candidate for our algorithm.

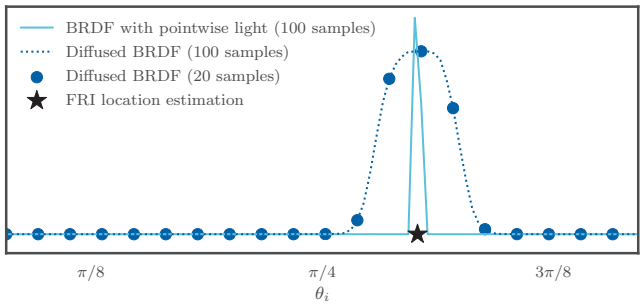


Figure 3.7: 1D BRDFs of a mirror with pointwise (continuous line) and extended (dashed line) light source, where θ_i is the inclination of the light direction. The location ($\star$) is estimated from 20 samples ($\bullet$).

It should be noted that the BRDF of most materials is a mixture of a diffuse and a specular component. Since we are only interested in the specular part, we need to subtract the diffuse component from the acquired signal; luckily, the diffuse component is wide and can be easily estimated, for instance with a few pointwise light sources.

Simulations

We run numerical experiments in 1 and 2 dimensions to demonstrate the feasability of our method. First, we generate a BRDF function ρ that follows a Ward distribution with $k_d = 1, k_s = 1/1000$, a random surface normal vector $\boldsymbol{n}$, and $\alpha = 0.008$, which corresponds to a relatively narrow specular peak (see Figure 3.6a). Note that the specular peak of the Ward model is not a separable kernel, but our experiments have shown that it is close to it (it can be approximated with a Gaussian function, which is separable).

The signal ρ is then convolved with a Gaussian pulse to simulate the effect of a diffuse light and then sampled at a relatively low rate: 50 samples are taken in the 1D case and 40 samples in each dimension in the 2D case (Figure 3.6b). Recall that the theory states that the smoothing kernel should be periodic; this is not the case in our experiments. This difference can however be neglected as the kernels we use decay fast enough.

The diffuse component parameters (surface normal $\boldsymbol{n}$ and diffuse albedo k_d) are estimated with photometric stereo [62] from a lowpass version of ρ. After we subtract the (filtered) diffuse component, we are left with a low-frequency version of the specular peak (Figure 3.6c). We run the FRI algorithm on its DFT coefficients to obtain the location of the specularity.

We then estimate the width of the specularity; this is a nonlinear problem which we tackle with the Levenberg-Marquardt algorithm. Recall that this procedure requires us to know the extended light shape. In practice, it can be measured by lighting a purely specular material with the desired light kernel from several locations. The rendering equation (3.1) shows that when ρ is a Dirac pulse, L_o directly provides pointwise samples of the shape of the light.

Finally, the BRDF is reconstructed using the recovered parameters (Figure 3.6d). Visually, the estimation matches closely the original shape. Quantitatively, we measure an SNR of 43.4 dB in the 1D case and 52.2 dB in the 2D case.

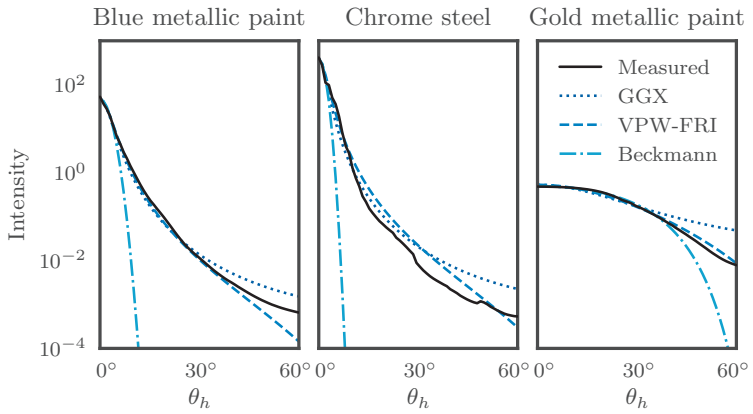


Figure 3.8: Modeling of a few glossy materials from the MERL BRDF database using a GGX distribution, Beckmann distribution and a VPW pulse.

BRDF acquisition with light dome

To confirm the above results, we perform a practical acquisition with our light dome with moving light source (see Figure 3.4b); We propose to localize the specularity of a flat mirror (its diffuse component is considered nil and it does not produce inter-reflections). The shape of the specular peak is close to a Dirac pulse, which is impossible to acquire with pointwise light sources and hence justifies the use of an extended sampling kernel. We position it with an angle with respect to the camera plane such that t_0 does not correspond to an inclination of 0.

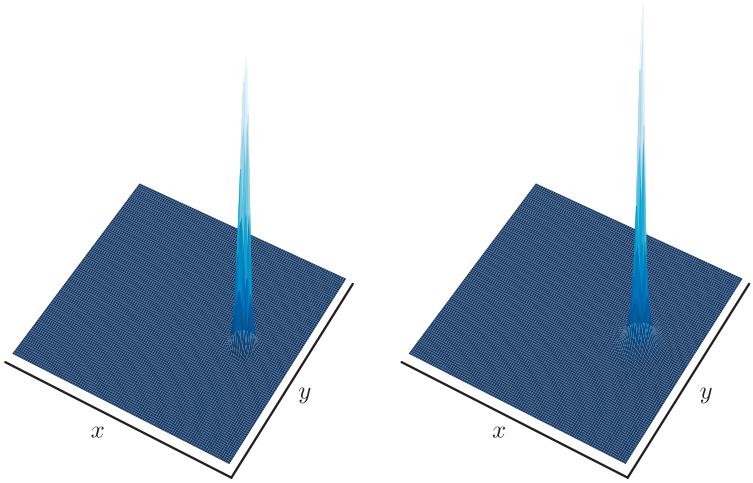
We focus the camera on the surface of the mirror and take 100 photographs of it; for each capture, we vary the inclination angle θ_i of the illumination. Moreover, we use both (almost) pointwise and extended light sources; the extended kernel is obtained by placing a diffusion filter of unknown shape in front of the light. For a single pixel selected in the middle of the frame, the two captured BRDFs are shown in Figure 3.7. By taking only 20 samples from the filtered BRDF, we can effectively recover the location of the specularity with FRI; this is clearly impossible from 20 pointwise samples.

3.4 VPW modeling of the BRDF

The main limitation of the above approach is that it can only represent one specularity. Furthermore, it can only recover the location of the specularities, but does not provide information about their width. To account for multiple specularities and various shapes, we need another model. In what follows, we demonstrate the practicality of the VPW-FRI model for representing and sampling specularities.

3.4.1 Modeling glossy materials

First, we show that the VPW parametric model is adequate to characterize glossy materials and that a single VPW pulse is appropriate to model the specularity ρ_s; indeed, we observe that if we formulate it with respect to $\tan(\theta_h)$, its shape is similar to the GGX distribution, up to



(a) Original signal (b) Estimated signal

Figure 3.9: 2D fit of a specularity: (a) original synthetic specularity generated according to the Beckmann distribution and (b) its estimation using a VPW pulse. For this experiment, we used all DFT coefficients corresponding to positive frequencies—in this case 50 in each dimension—in the estimation algorithm.

a normalization factor and a square root term.[3] We compare in Figure 3.8 the fitting of these models to a few measured BRDFs from the MERL BRDF database [115]. In this experiment, and as we neglect the Fresnel and geometry terms, we focus on θ_h in the range $[0°, 60°]$. Within this range, we see that the VPW model compares favorably to the Beckmann and GGX distributions and presents a good trade-off between the heavy tails of the GGX distribution and the relatively narrow Beckmann distribution.

3.4.2 Specularity localization on the plane

Next, we demonstrate on synthetic data how our algorithm and model can be used to estimate the width and location of a specular peak projected onto the xy-plane. For this experiment, we fix the viewing angle and only study the variation with respect to the incoming light angle. The pulse in Figure 3.9a is a specularity generated according to a Beckmann distribution with $\alpha = 0.003$ and $k_s = 0.005$; we fit a VPW pulse to it and estimate its parameters using the FRI procedure coupled with a 2D version of Cadzow's algorithm. As we can see, a single VPW pulse provides a good approximation of its shape, as shown in Figure 3.9b.

3.4.3 Specularity localization on the sphere

Finally, we propose to perform a practical acquisition from our dome with movable arm presented in Figure 3.4b. Depending on the number of angular samples, the acquisition process can be lengthy and tedious. Moreover, most light domes are composed of a limited number of lights at fixed positions; this motivates the need for a method that can faithfully estimate the BRDF from as few samples as possible.

[3] To account for this difference, we actually fit the *square root* of the measurements to the VPW pulse.

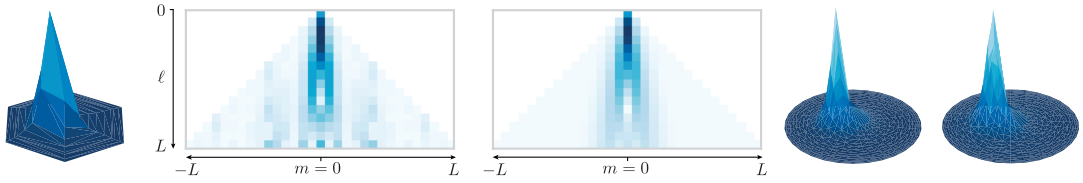


Figure 3.10: Specularity estimation using spherical VPW-FRI algorithm: (a) 18 samples of the acquired BRDF; (b) the spectrum computed from the 18 samples; (c) the spectrum estimated using VPW-FRI with $K = 1$; (d) the corresponding reconstructed spherical signal; (e) the same BRDF acquired at a higher sampling rate (450 samples) for comparison.

Using a Nikon D810 full-frame DSLR camera, we acquired samples of the reflectance field of a painting using $N = 18$ different light positions; we discretized the azimuth into 6 uniform bins and the elevation into 3 uniform bins. In practice the light is not a perfect point source; this fact, combined with the acquisition system, acts as a lowpass filter that eliminates the high frequency components. Using Theorem 2.3, we can approximate the SH coefficients from the spatial samples.

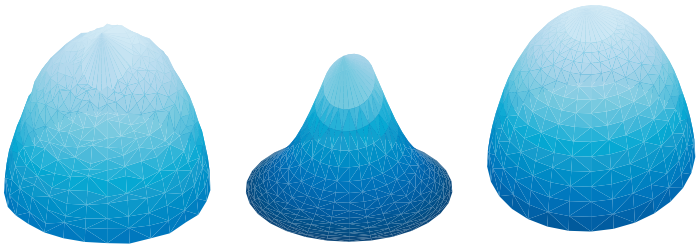
We then use spherical VPW-FRI to estimate the parameters of the pulses from these coefficients; for most BRDFs, one or two pulses suffice to represent the specularity. As a consequence, we can obtain a good approximation of the specular peak even with a relatively low number of samples on the sphere. This idea is illustrated in Figure 3.10: from only 18 spherical samples, we can approximate the spectrum of the reflectance field—see Figure 3.10b—using (2.13). Moreover, from these SH coefficients, the FRI algorithm allows us to estimate the spectrum, as illustrated in Figure 3.10c. Since VPW-FRI is a parametric model, its spectrum can have infinite support. This enables a continuous representation of the reflectance function in the spherical domain; Figure 3.10d shows the estimated VPW pulse sampled at 450 locations on the hemisphere. For the sake of comparison, we also acquired a denser version of the reflectance field—shown in Figure 3.10e; we observe that the VPW-FRI approximation from 18 samples is extremely close to the ground truth.

To summarize, the algorithm allows us to accurately estimate the specularity from a few samples; this reduces the acquisition time as well as storage requirements. Furthermore, given this continuous VPW representation, we can render objects under virtually any new illumination.

3.4.4 Practical considerations

We already mentioned in Section 2.5 that the method does not exploit the entire spectrum but only uses its two outer diagonals, i.e. the SH coefficients for $m = \ell$ and $m = \ell - 1$. Furthermore, as depicted in Figure 3.10b, the energy of a typical BRDF signal is concentrated along the central column of the spectrum, and these coefficients are unfortunately not used by the currently proposed algorithm. To achieve better spectrum usage, we add a nonlinear optimization routine that minimizes the squared error between the model and the measurements over the full spectrum. As the objective function is non-convex, this method relies on the VPW-FRI parameters to provide a good initialization. Moreover, to make the VPW-FRI estimation more robust, we first denoise the SH coefficients with Cadzow's algorithm.

Another limitation is model mismatch. While our experiments have demonstrated that VPW



(a) Original (b) VPW model (c) Cosine model

Figure 3.11: Example of model mismatch: (a) a measured BRDF composed only of a Lambertian part, and (b) its spherical VPW approximation. Simpler models such as (c) a cosine can be used instead.

pulses are adequate to represent specular peaks, there are other components in the BRDF that are not easily approximated by VPW pulses: the Lambertian component and hard cast shadows are such examples. The mismatch between the Lambertian component and its VPW approximation is illustrated in Fig 3.11. Fortunately, the Lambertian part is much less challenging to detect and represent than the specular part, since it can easily be approximated using a constant value modulated with the cosine of the azimuth of the incoming light angle θ_i. Hard shadows are more difficult to cope with, as they also contain high frequency components, and thus provide an interesting topic of investigation for future work.

3.5 Conclusion

We applied FRI-based methods to the problem of estimating BRDFs, with emphasis on the detection of specularities. First, we proposed a new FRI-based technique to estimate the location of a single Dirac pulse convolved with an unknown symmetric sampling kernel. We tested our proposed method in numerical simulations and in a practical experiment: in both cases, we showed that the specularity is well approximated from a small number of samples. Second, we demonstrated that with the VPW-FRI framework from Chapter 2, we can accurately retrieve and represent the shape of specularities from just a handful of measurements. In these applications, we stressed the importance of the data matching the underlying parametric model: when this condition was satisfied, results were very convincing.

Chapter 4

Wave Theory and Interferences: Lippmann Photography*

It doesn't matter how beautiful your theory is, it doesn't matter how smart you are. If it doesn't agree with experiment, it's wrong.

The Feynman Lectures on Physics
Richard Feynman

4.1 Introduction

Color photography has been around for about 150 years, since James Clerk Maxwell and Thomas Sutton proposed and experimented with the three-color method for the acquisition of chromatic information via three separate channels. This approach, which mimics the way the human eye sees, is still the main ingredient of almost all analog and digital contemporary color photography techniques, and is nowadays embodied by color film with dyes and color filter arrays (CFAs).

Leveraging the wave nature of light, Gabriel Lippmann proposed in 1891 an interesting alternative approach. Inspired by preliminary work by Herschel [76], Becquerel [19, 20], Wiener [197],

*The material in this chapter is the result of joint work with Arnaud Latty, Michalina Pacholska, Martin Vetterli and Adam Scholefield [9, 12]. Author contributions: GB, MV, and AS designed research; MP contributed to numerous meetings and brainstorming sessions; GB, AL, and AS performed research; GB and AL performed experiments; and GB analyzed data, ran simulations and wrote the chapter.

Figure 4.1: Principle of Lippmann photography: the scene on the left is focused with a photographic lens on the plane of the emulsion. At the back of the emulsion is a mercury layer; when the light field hits the mercury layer, it reflects, and interferes with itself. Its spatially varying intensity creates patterns that are captured inside the depth of the emulsion (see the close-up).

and Zenker [206], he effectively created the first spatial interferometer [102, 103] by generating and recording standing waves into the depth of photosensitive material.[1] Lippmann achieved this feat by placing a mercury mirror adjacent to a photographic emulsion made of transparent gelatin (or albumen) and silver halides, as depicted in Figure 4.1. One of the challenges was to use extremely fine silver halide grains—around 10 nm—to enable the recording of the interferences corresponding to the visible light range. The interference patterns created by reflection expose the emulsion differently at different depths. For example, in the simple case of monochromatic light—that is with light made of a single wavelength—the interference patterns have the shape of a sinusoid, effectively creating a Bragg grating inside the photographic plate. Once it has been exposed, the plate is developed and possibly fixed via standard techniques.

Viewing a Lippmann photograph is achieved by illuminating it with a white or flat-spectrum light source. After interacting inside the plate, it can be shown that the reflected spectrum resembles the original spectrum of the scene. The colors reveal themselves only under well chosen illumination and viewing angles, namely when the direction of viewing corresponds to the direction of illumination mirrored with respect to the surface normal. Witnessing a Lippmann plate is a special experience in itself that is difficult to describe in words or even with static images. Nevertheless, we show in Figure 4.2 a trichromatic approximation of the appearance of a plate under both ambient illumination and adequate viewing conditions.[2]

Unfortunately, the technique has been almost completely forgotten today and didn't make it into modern camera designs, which are almost exclusively based on 3-layer emulsions in analog cameras, or CFAs in digital cameras; one of the reasons is that the fineness of the silver halide crystals decreases the photographic sensitivity and requires excessively long exposure times. Indeed, the three color-based films and emulsions in analog cameras can be made with much larger grains than those suitable for Lippmann photography and therefore achieve higher sensitivity. Also, Lippmann plates are unique originals and essentially cannot be copied, in contrast to conventional photography, where several copies can be created from a negative; this hampered

[1] Many more actors were involved in the development of interference photography. For a more exhaustive history about it, we direct the reader to [43, 120].

[2] A video of the same plate under different illuminations can be found at `https://go.epfl.ch/lippmann`.

(a) Ambient illumination (b) Proper illumination

Figure 4.2: A Lippmann plate, which has been created by photographer Filipe Alves with a home-made albumen emulsion: under (a) ambient illumination and (b) directed light whose incoming direction is the mirror of the viewing direction with respect to the surface normal.

the democratization of the Lippmann process.

Nevertheless, Lippmann photochromes strongly influenced and paved the way to other major fields, such as holography. In fact, Lippmann photography and holography are both based on the concept of a light field interfering with a reference wave: in the case of Lippmann, the reference wave is the reflection of the image forming wave in a mirror, whereas in holography it is a coherent light beam that is used to illuminate the scene. Another fundamental difference between the two techniques is the fact that, in Lippmann photography, the phase is lost in the process since the mirror acts as a phase locker.

Unlike three-color methods, Lippmann's approach has the advantage of recording the entire spectrum of visible light. Another strength is that, unlike CFA-based techniques, the process doesn't trade off spatial resolution to enable the capture of colors, since the spectrum is acquired along a third dimension. Lippmann was awarded the Nobel Prize in 1908 for his discovery, which was well received by the scientific community and enabled fruitful research in the early 20th century: among the scientists who contributed to refine the technique and its understanding, Neuhauss [131], Lehmann [100], and Cajal [154] managed to observe the interference patterns produced using quasi monochromatic light under optical microscopes by swelling the gelatin. To avoid the use of a toxic element as reflector, Rothé [157] showed that the interface between the gelatin and air generates reflections that are sufficiently strong to create interferences, eliminating the need for mercury. The Lumière brothers [105, 106], Wiener [198] and Ives [81] were also prominent actors who improved the process, mainly by refining the chemical processes used to make and develop Lippmann plates.

More recently, a few research groups investigated it, but Lippmann photography remains a niche topic. Bjelkhagen studied many practical aspects of the technique such as the recording with holographic plates and the use of the method for security purposes [28–30]. Phillips et al. [147, 148] precisely described the scattering of the silver grains as well as their size limitations, and established constraints regarding the aperture (f-number) of the lens for a successful Lippmann photograph. Fournier et al. [63, 64, 114] studied the creation of extremely fine grain emulsions and analyzed the structure and spectrum reproduction of historical and contemporary photographic plates. Finally, Nareid and Pederson [125, 126] introduced a mathematical model

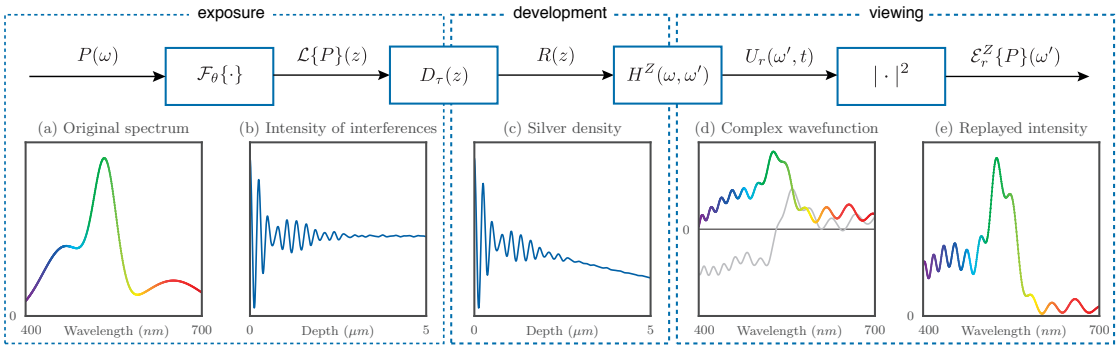


Figure 4.3: Lippmann photography pipeline: (a) the original spectrum $P(\omega)$ undergoes a linear transform and (b) its interference patterns $\mathcal{L}\{P\}(z)$ are captured inside the plate as a distribution of metallic silver particles; the plate is then developed and (c) its reflectivity $R(z)$ at depth z is proportional to $\mathcal{L}\{P\}(z)$, modulated by a window $D(z)$; when the plate is illuminated, the reflected wavefunction (d) $U_r(\omega', t)$ is similar to filtering $P(\omega)$ with $H^Z(\omega, \omega')$; finally, the observed spectrum is given by the time-averaged intensity $\mathcal{E}_r^Z\{P\}(\omega')$ of $U_r(\omega, t)$.

based on local changes in the refractive index of the emulsion.

Motivated by these works and the fact that the largest collection of historical Lippmann plates sits at Musée de l'Elysée[3], a mere 7 km away from our school, we started investigating the technique. On close inspection of some plates with a hyperspectral imaging device, we realized that the synthesized spectra contain some oscillations, which had never been documented previously, and were not predicted by any model. Moreover, existing theory predicts that the choice of reflector (i.e. mercury or air) should only alter the intensity of the reflected colors, but when we ran controlled experiments, we discovered that it has a more significant influence on the resynthesized colors, leading to distortions in the reflected spectrum. Although the Lippmann process is conceptually simple to understand, it is lacking a rigorous treatment in the literature. On one hand, a large body of work only handles the monochromatic case. On the other hand, the details of the polychromatic case are not fully understood from a theoretical point of view. For instance, a recurring claim in the literature [30, 103] is that the Lippmann process enables perfect reconstruction of the original spectrum.

We address these questions by presenting a rigorous treatment of Lippmann's process, adopting a signal processing perspective; in fact, Lippmann photography is explained with Fourier analysis and many of its principles are based on signal processing concepts. The different steps of the Lippmann pipeline are illustrated in Figure 4.3. Their mathematical formulation helps us characterize the different artifacts that appear in the procedure. We show in this chapter that even though the spectra reflected by Lippmann plates are a close approximation, they are not identical to the original spectra, even when assuming plates with infinite thickness. Furthermore, we verify that the thickness of the plate acts as a lowpass filter and in addition generates oscillations in the synthesized spectra. We also demonstrate how the choice of reflector alters the synthesized spectrum and results in what we call a *skewing* effect.

These observations lead naturally to the question of inverting these distortions and recovering the original spectrum of the scenes captured. This could potentially enable us to transform

[3]http://www.elysee.ch

historical Lippmann photographs into hyperspectral windows to the past, revealing spectral data of scenes and objects dating from before the advent of hyperspectral imaging. Interestingly, we show that most of the inaccuracies in the color reproduction can in fact be inverted, and that the original spectrum can be fully recovered under the assumption that its bandwidth is lower than the one determined by the thickness of the plate. The making of our own Lippmann photographs, detailed in Appendix 4.A, is used throughout the chapter to demonstrate and validate the different effects studied.

4.2 Analysis: Recording a Lippmann plate

Lippmann photography essentially works by recording the Fourier cosine transform of the light spectrum in the depth of a photographic emulsion. The emulsion is usually made of gelatin, and contains extremely fine silver halide grains that react to exposure with light, as well as color dyes to make the plate sensitive to all visible wavelengths.[4] During the recording, the emulsion is put in direct contact with a mirror, creating interference patterns inside it. We formalize the process in this section. For simplicity, we first introduce the 1-dimensional case, that is when the waves propagate along the z-axis. Moreover, we only describe the phenomenon for a single 'pixel', but a Lippmann photograph in fact consists of a continuum of varying interferences in the xy-plane [114]; the generalization to three dimensions is discussed in Section 4.4.

4.2.1 Standing waves

Throughout this chapter, we model light fields as coherent plane waves. The 1D wavefunction $U_0(z,t)$ of an incident light wave with heterogeneous spectrum can be decomposed into an infinitesimal sum of individual sine waves, each having an amplitude $A(\omega)$:

$$U_0(z,t) = \int_0^\infty A(\omega) e^{j\left(\omega t - \frac{\omega z}{c}\right)} d\omega,$$

where z and t are the spatial and temporal components, respectively, and $c = c_0/n$ is the speed of light in the medium (with c_0 being the speed of light in vacuum and n the refractive index of the medium). Note that plane waves are not actual physical waves, since they extend to infinity in the directions other than the direction of propagation. They are, however, a reasonable approximation if the size of the imaged object is much larger than the wavelength, which is true for artistic photography.

The power spectral density of U_0 is constant in space and is simply given by

$$P(z,\omega) = P(\omega) = A^2(\omega).$$

By convention, we set $z = 0$ to be the position of the mirror. Upon hitting it, the wavefunction U_0 is reflected back. We define the reflection coefficient of the mirror as $r = \rho e^{j\theta}$, where ρ is the attenuation factor and θ the phase shift. This leads to the following expression for the reflected wave:

$$U_1(z,t) = r \int_0^\infty A(\omega) e^{j\left(\omega t + \frac{\omega z}{c}\right)} d\omega.$$

[4] Since Lippmann recorded his first photographs, other materials such as photopolymers [27, 64] have been used to record the interference field. Dichromated gelatin [162] is an example of a photopolymer used in holography.

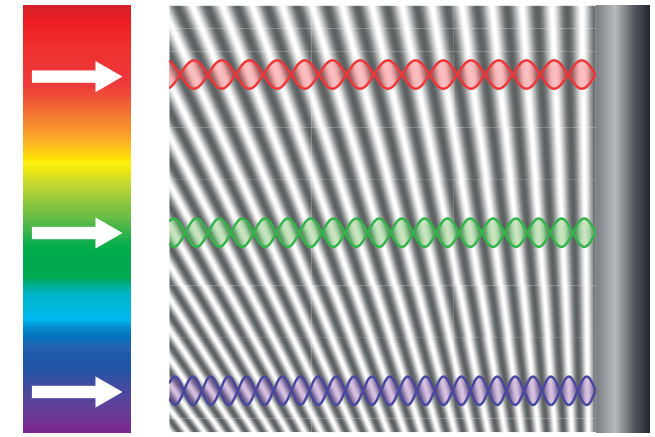


Figure 4.4: Lippmann photography with monochromatic light: monochromatic waves enter the plate and the interference with their reflected version creates standing waves in the photographic emulsion. Note that this illustration is not to scale.

The resulting wave is then composed of both the incoming and the reflected fields:

$$\begin{aligned} U(z,t) &= U_0(z,t) + U_1(z,t) \\ &= \int_0^\infty A(\omega) e^{j\omega t} \left(e^{-j\frac{\omega z}{c}} + r e^{j\frac{\omega z}{c}} \right) d\omega. \end{aligned}$$

The time-averaged intensity of this wavefunction is therefore given by

$$\begin{aligned} \mathcal{E}\{P\}(z) &= \int_0^\infty A^2(\omega) \left| e^{-j\frac{\omega z}{c}} + r e^{j\frac{\omega z}{c}} \right|^2 d\omega \\ &= \int_0^\infty P(\omega) \left(1 + \rho^2 + 2\rho \cos\left(\frac{2\omega z}{c} + \theta \right) \right) d\omega. \end{aligned} \tag{4.1}$$

What is interesting is that the intensity varies spatially in z and takes the form of *standing waves*[5], or more generally *partial* standing waves when $\rho < 1$. Since this is happening inside the depth of a photosensitive plate, these interference patterns expose the photographic emulsion differently at different depths. In the case of monochromatic light, the interference pattern takes the form of cosines, whose frequency of oscillation depends on the wavelength (see Figure 4.4).

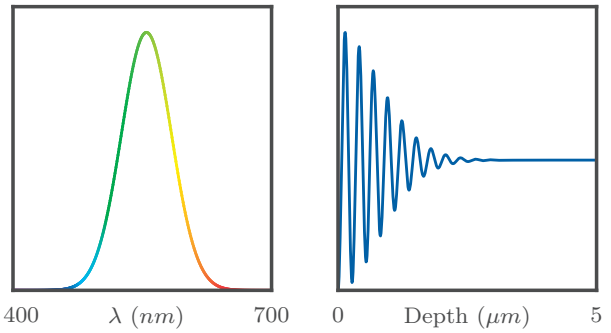
To simplify notation, we introduce the *Lippmann operator* $\mathcal{L}_r$. Given a power spectral density $P(\omega)$, we define it as

$$\mathcal{L}_r\{P\}(z) = (1+\rho^2)\mathcal{F}_\theta\{P\}(0) + 2\rho\mathcal{F}_\theta\{P\}(z), \tag{4.2}$$

where $\mathcal{F}_\theta$ is the generalized Fourier cosine transform operator:

$$\mathcal{F}_\theta\{P\}(z) = \int_0^\infty P(\omega) \cos\left(\frac{2\omega z}{c} + \theta \right) d\omega. \tag{4.3}$$

[5]Technically speaking, standing waves only form in the monochromatic case. In the polychromatic case, we can think of the shape of the resulting wave as an infinite sum of standing waves.



(a) Spectrum (b) Interference patterns

Figure 4.5: Example of (a) a Gaussian spectrum and (b) the interference pattern generated by its reflection on a perfect mirror ($r = -1$).

In the literature [30, 103, 126], perfect reflection is often assumed, that is either $r = -1$ or $r = 1$, depending on whether there is a phase inversion at the interface:

$$\mathcal{L}_1\{P\}(z) = 2\int_0^\infty P(\omega)\left(1 + \cos\left(\frac{2\omega z}{c}\right)\right)d\omega,$$
$$\mathcal{L}_{-1}\{P\}(z) = 2\int_0^\infty P(\omega)\left(1 - \cos\left(\frac{2\omega z}{c}\right)\right)d\omega.$$

We choose to maintain the seemingly more sophisticated model from (4.2) as it allows for more precise analysis; as we will see later in Section 4.5.3, the value of r has a strong influence on the resulting spectrum.

To summarize, the interference patterns inside a Lippmann plate are standing waves and are given by the sum of a constant and an oscillating term: this is showcased in Figure 4.5, which illustrates the pattern created by a Gaussian spectrum with $r = -1$.

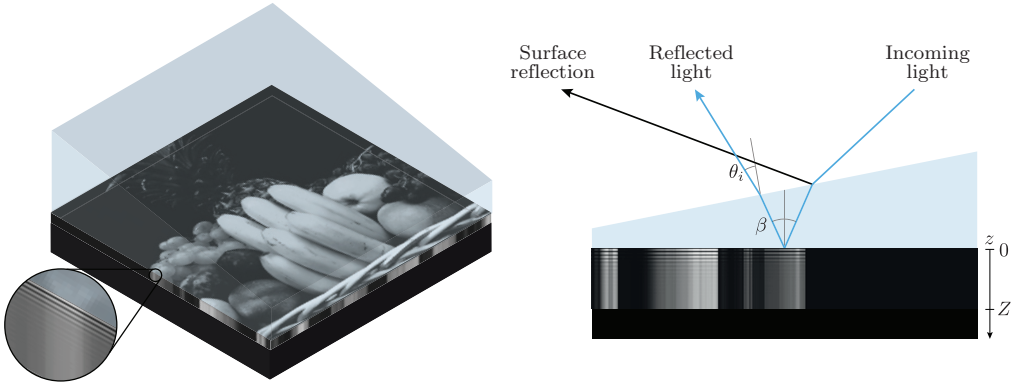
4.2.2 Development

After the plate has been exposed, the next step is to develop it. From a chemical point of view, the exposure to light has broken some of the bonds between silver particles and halide crystals. Once free, these metallic silver particles form tiny specks that combine into a *latent image*, and their density is proportional to the intensity of the exposing standing waves.[6]

Loosely speaking, the role of development is to increase the concentration of metallic silver in the exposed areas. It is important to note that the development is not uniform throughout the depth of the plate and deeper layers are generally less developed than the top layers. We propose to model the development influence as a modulation of the recorded intensity; this is the topic of Section 4.5.2.

After the development bath, a fixing process is optionally applied to wash out dyes, halide crystals and silver halides [75, 173], so that the plate is no longer sensitive to light. We refer

[6] In practice, the response of a photographic plate to light is only linear in a certain regime, and it reaches saturation after a certain point; it can be described more accurately with the so-called *Hurter and Driffield curves*. See Appendix 4.A.2 for more details.



(a) Lippmann plate with prism (b) Light paths

Figure 4.6: Lippmann plate prepared for visualization after development; (a) the glass plate has been painted in black to minimize back reflections and a prism has been mounted on the face that was in contact with the mercury layer to enhance contrast; (b) the light rays reflecting at the surface of the prism do not coincide with the rays reflected from the plate.

the reader to Appendix 4.A.3 for more details about the development process and the different chemical recipes.

4.3 Synthesis: Viewing a Lippmann plate

After exposure, the mirror is removed. To visualize a processed plate, we turn it upside-down and illuminate it with a light source with a power spectral density denoted as $L(\omega')$. The incoming light is scattered by the silver particles inside the plate and the silver particle distribution creates local Bragg networks such that, for specific wavelengths, the reflected light waves add constructively. From a higher level, these networks can be thought of as elementary partially reflective mirrors, whose reflectance is proportional to the intensity of the exposing interference field.

Remark that here we strictly describe the interaction of the light *inside* the plate, but do not take into account the reflection of the light at the surface of the emulsion. This reflection is generally strong compared to the optical wave reflected from the plate. To cope with this issue, a prism with a refractive index relatively close to that of gelatin is usually attached to the top of the plate; this prevents the surface reflection from mixing with the internal reflections (see Figure 4.6).

4.3.1 Analytical model

Gabriel Lippmann chose to explain his recording technique in an analytical fashion, which led to an equation which he recognized as being a Fourier integral. We follow the same path here, but adopting a more rigorous style. In Lippmann's original paper [103], it is assumed that

the *reflectance* at depth z in the plate is proportional to the intensity of the exposing wave.[7] This model makes a couple of additional assumptions that simplify the behavior of light inside the plate. First, all but the first-order reflections are neglected. In other words, this assumes that each of the mirrors formed by the metallic silver particles only reflect direct light and we neglect all inter-reflections. Second, it is supposed that the amount of light that hits each layer is always the same regardless of the thickness. This is clearly not the case in reality: deeper layers cannot receive as much light as shallow layers, since some light is already reflected by these shallow layers. Nevertheless, these assumptions are reasonable when the amount of reflected light remains relatively small. More importantly, they enable us to derive a closed-form solution of the Lippmann procedure, which we describe below. In Section 4.3.2, we relax these assumptions and quantify the estimation error of this analytical model using the theory of wave-transfer matrices.

Modeling

Defining the depth of the entrance of the plate—i.e. the air-gelatin, or prism-gelatin interface—as $z = 0$ (see Figure 4.6b), we denote the amount of light that is reflected at depth z by a *reflectance function* $R(z)$ that is proportional to the intensity of the exposing interference field:

$$R(z) = \epsilon \mathcal{L}_r \{P\}(z). \tag{4.4}$$

Note that in practice, the development and processing of the plates is a non-linear process that is difficult to characterize; here, like Lippmann, we assume linearity, which leads to a closed form analytic expression. The parameter ϵ controls the amount of light exiting the plate, which typically cannot be larger than the quantity of light entering the plate. For now, we assume it to be constant; later, we will modulate it with a decaying window that is dependent on z, such that it can encompass the effects due to a nonuniform development and processing of the plate.

Each monochromatic component of $L(\omega')$ reflected at depth z travels a round trip distance of $2z$ between the entrance of the plate and the depth z: this corresponds to a phase shift of $2\omega' z/c$. The optical wave reflected from the depth z and measured at the entrance of the plate is therefore given by

$$dU_r(z, \omega', t) = \sqrt{L(\omega')} R(z) e^{j(\omega' t - \frac{2\omega' z}{c})} dz.$$

The total reflected wavefunction U_r is then the integral from 0 to Z over all partial reflected waves:

$$\begin{aligned} U_r(\omega', t) &= \int_0^Z dU_r(z, \omega', t) \\ &= \sqrt{L(\omega')} e^{j\omega' t} \int_0^Z R(z) e^{-j\frac{2\omega' z}{c}} dz. \end{aligned} \tag{4.5}$$

Using (4.2) and (4.4), it is convenient to express this integral as

$$U_r(\omega', t) = \epsilon \sqrt{L(\omega')} e^{j\omega' t} \int_0^\infty P(\omega) H^Z(\omega, \omega') d\omega, \tag{4.6}$$

[7] To be more precise, Lippmann originally assumed that the reflectance was proportional to the amplitude of the wave, which is physically less intuitive.

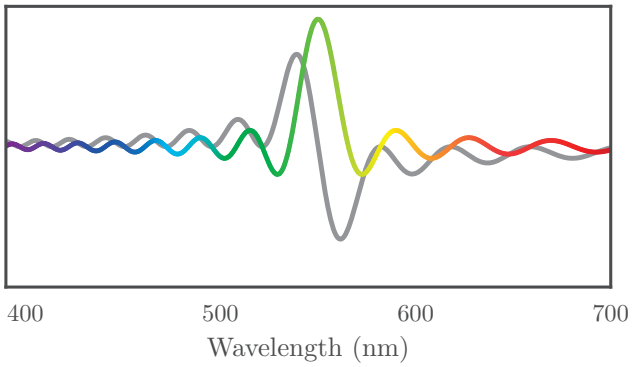


Figure 4.7: The filter $S^Z(\omega' - \omega)$ for $Z = 5\ \mu m$ and $\omega' = \frac{2\pi c}{550\ \text{nm}}$. The real part is represented by the colorful line and the imaginary part by the gray line. Notice that the period of the oscillations is regular with respect to the frequency ω; it appears to be changing here as we show the spectrum as a function of the wavelength, which is inversely proportional to the frequency.

where we define the function $H^Z(\omega, \omega')$ as

$$H^Z(\omega, \omega') = \int_0^Z \left(1 + \rho^2 + 2\rho \cos\left(\frac{2\omega z}{c} + \theta\right)\right) e^{-j\frac{2\omega' z}{c}} dz. \tag{4.7}$$

Observe that expression (4.6) resembles a filtering operation. At this point, it is interesting to consider the hypothetical case where we extend the reflectance $R(z)$ to negative values of z. Even though the use of such a function does not make sense in the context of Lippmann photography—it is obviously impossible to acquire the interference field beyond the mirror—it contributes to a better understanding of the process. Let

$$\begin{aligned} h^Z(\omega, \omega') &= \frac{1}{2}\int_{-Z}^{Z} \left(1 + \rho^2 + 2\rho \cos\left(\frac{2\omega z}{c} + \theta\right)\right) e^{-j\frac{2\omega' z}{c}} dz \\ &= \frac{1}{2}\int_{-Z}^{Z} \left(1 + \rho^2\right) e^{-j\frac{2\omega' z}{c}} dz + \frac{1}{2}\int_{-Z}^{Z} \left(r e^{j\frac{2\omega z}{c}} + r^* e^{-j\frac{2\omega z}{c}}\right) e^{-j\frac{2\omega' z}{c}} dz \\ &= \frac{r}{2} s^Z(\omega' - \omega) + \frac{1 + \rho^2}{2} s^Z(\omega') + \frac{r^*}{2} s^Z(\omega' + \omega), \end{aligned} \tag{4.8}$$

where the filter $s^Z(\omega)$ is defined as

$$s^Z(\omega) = Z\,\text{sinc}\left(\frac{2Z\omega}{c}\right)$$

and $\text{sinc}(x) = \sin(x)/x$ is the *cardinal sine* or *sinc function*. From (4.7), we observe that H^Z is the Fourier transform of a function that is zero for depths $z < 0$. In signal processing terms, we call H^Z the *analytic signal* of h^Z. The frequency domain representation of H^Z can be derived using the Hilbert transform of h^Z:

$$\begin{aligned} H^Z(\omega, \omega') &= h^Z(\omega, \omega') + j\int_{-\infty}^{\infty} \frac{h^Z(\omega, \omega' - \nu)}{\pi\nu}\,\mathrm{d}\nu \\ &= \frac{r}{2} S^Z(\omega' - \omega) + \frac{1 + \rho^2}{2} S^Z(\omega') + \frac{r^*}{2} S^Z(\omega' + \omega), \end{aligned} \tag{4.9}$$

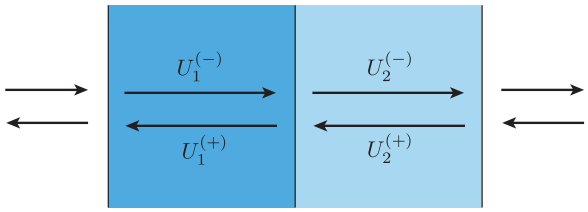


Figure 4.8: Wave-transfer matrix theory: the medium is subdivided into thin layers (2 are shown here); at each boundary, the matrix captures the relationship between the forward and backward waves on each side of the boundary. Illustration inspired from [160].

with

$$S^Z(\omega) = Z\,\mathrm{sinc}\left(\frac{2Z\omega}{c}\right) + jc\frac{1-\cos(2Z\omega/c)}{2\omega}$$
$$= \frac{c}{2j\omega}\left(1 - e^{-2j\frac{Z\omega}{c}}\right).$$

An example of such a function is shown in Figure 4.7. Finally, we do not observe directly $U_r(\omega', t)$ but rather its time-averaged intensity, which is given, for each wavelength ω', by

$$\mathcal{E}_r^Z\{P\}(\omega') = \epsilon^2 L(\omega')\left|\int_0^\infty P(\omega)H^Z(\omega,\omega')d\omega\right|^2. \tag{4.10}$$

This expression is the squared magnitude of an analytic signal and it is often called the *envelope*.

4.3.2 Wave-transfer matrix model

The more rigorous analytic model that we have just described can explain a certain number of phenomena regarding the color reproduction. Before exploring them, we assess the validity of some of the assumptions of this model. In particular, recall that, like Lippmann, our analytic model does not take into account higher-order reflections and light intensity decay inside the plate. To quantify this approximation and its deviation from a physically accurate model, we make use of *wave-transfer matrices*.

Wave-transfer matrices

Sometimes, the interactions of waves inside a medium with varying composition can be complex to characterize analytically. In such cases, wave-transfer (or *transmission*) matrices can be used to describe and numerically simulate a number of stacked optical phenomena. The idea is to decompose a medium into elementary thin layers (see Figure 4.8) and to express each layer and the interface between them with a matrix [160]. Such a matrix $\boldsymbol{M}$ connects the optical waves on each side as follows:

$$\begin{bmatrix} U_2^{(+)} \\ U_2^{(-)} \end{bmatrix} = \underbrace{\begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}}_{\boldsymbol{M}} \begin{bmatrix} U_1^{(+)} \\ U_1^{(-)} \end{bmatrix}.$$

Here, U_1 represents the resulting wave at the entrance of the layer and U_2 the resulting wave on the other side of it. Furthermore, $U_1^{(+)}$ describes a forward resulting wave (traveling toward the first layer) and $U_1^{(-)}$ a backward resulting wave (traveling away from the layer). Similarly, $U_2^{(+)}$ characterizes a forward wave that travels toward the second layer, and so on...

Wave-transfer matrices are used to characterize a wide range of optical components and phenomenons such as mirrors, beam splitters, reflection, refraction, or propagation.

Application to Lippmann's theory

Lippmann plates can be approximated by a succession of two basic layers: a transport layer and a semi-reflective layer. An elementary semi-transparent reciprocal mirror at depth z can be described by the matrix [160]

$$\boldsymbol{M}(z) = \frac{1}{T(z)} \begin{bmatrix} 1 & jR(z) \\ -jR(z) & 1 \end{bmatrix},$$

where $T(z)$ is the *transmittance* and $R(z)$ is the *reflectance* as defined in (4.4). The two quantities are such that $|T(z)|^2 + |R(z)|^2 = 1$; that means we assume that all the light is either transmitted or reflected, but not absorbed—absorption occurs when the sum of squares is strictly less than 1. This is an approximation as in practice, silver particles do absorb some amount of light. Furthermore, the propagation in a homogeneous medium of refractive index n over a distance Δz can be represented by the following matrix operation:

$$\boldsymbol{P}(\Delta z, k_0) = \begin{bmatrix} e^{-j\varphi} & 0 \\ 0 & e^{j\varphi} \end{bmatrix},$$

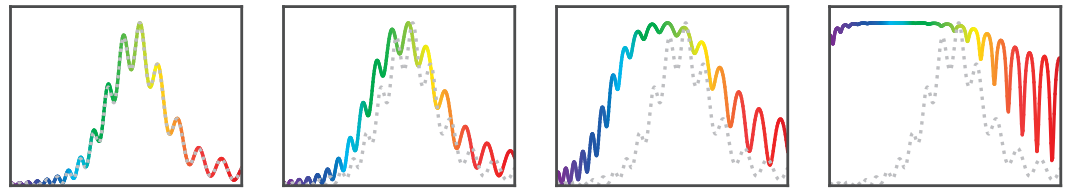
where Δz is the distance between two mirrors and $\varphi = nk_0\Delta z$ is the phase difference (with k_0 the wave number in vacuum).

With this notation at hand, we can represent the total transmitted and reflected light by combining elementary mirrors and propagation matrices. We assume that the plate is subdivided into L thin layers located at depths $z = 0, \Delta z, 2\Delta z, \ldots, (L-1)\Delta z$. Then the resulting wave-transfer matrix $\boldsymbol{M}_{tot}(k_0)$ for a given wavenumber k_0 is computed as

$$\boldsymbol{M}_{tot}(k_0) = \prod_{\ell=0}^{L-1} \boldsymbol{P}(\Delta z, k_0)\boldsymbol{M}(\ell\Delta z).$$

In our model, the amount of light reflected is controlled by the parameter ϵ. Simulations for different values of ϵ are shown in Figure 4.9; we observe that for low reflectance values, the analytic expression developed in Section 4.3.1 matches the spectra obtained with wave-transfer matrices. However, when the reflectance becomes more important, we enter a saturation regime that is not predicted by the analytical model. Fortunately, our experience have shown that in Lippmann photography the reflectance is relatively low and we are in a regime that is well below saturation.

Note that oscillations are present in the reflected spectra of Figure 4.9. They are due to the finite thickness of the plate; we will study this effect in more details in Section 4.5.1.



(a) $\epsilon = 10^6$ (b) $\epsilon = 5 \cdot 10^6$ (c) $\epsilon = 10^7$ (d) $\epsilon = 2 \cdot 10^7$

Figure 4.9: Saturation curves for the reflection-based model on the spectrum of Figure 4.5 for different values of ϵ. The dotted line represents the spectrum predicted by the Lippmann assumption while the spectral line is the one obtained with wave-transfer matrices.

4.4 The 3D case: Oblique illumination and viewing angles

In Sections 4.2 and 4.3, we considered 'one pixel' Lippmann photographs, that is the recording of a plane wave propagating in the z direction and a plate perpendicular to z. In this section, we extend the theory to the recording and viewing of Lippmann plates along arbitrary directions. It is known that when the viewing angle is not perpendicular to the plate, the rendered spectrum suffers a color shift toward the blue [151]. Using our model, we can easily explain this phenomenon too; additionally, we demonstrate what happens when the plate is not perpendicular to the wavefronts during the exposure. This is also what happens in practice, as the light rays of the scene hit the image plane with a small angle after passing through the photographic lens. Interestingly, Ives proposed to leverage this phenomenon to create interferential filters using Lippmann plates [82].

4.4.1 Recording

Suppose that we have a plane wave traveling along the direction $\boldsymbol{d} = [d_x, d_y, d_z]^\top$, where $\|\boldsymbol{d}\| = 1$. It can be described at position $\boldsymbol{r} = [x, y, z]^\top$ and time t as

$$U_0(\boldsymbol{r}, t) = \int_0^\infty A(\omega) e^{j(\omega t - \frac{\omega}{c}\boldsymbol{d}^\top \boldsymbol{r})} d\omega.$$

We assume that the wave is reflected on a mirror with reflection coefficient r that is perpendicular to the z-axis. As in the 1D case, the wave U_0 interferes with its reflection, resulting in the expression

$$\begin{aligned} U(\boldsymbol{r}, t) &= U_0(\boldsymbol{r}, t) + r \int_0^\infty A(\omega) e^{j\frac{\omega}{c}(ct - d_x x - d_y y + d_z z)} d\omega \\ &= \int_0^\infty A(\omega) \left(e^{-j\frac{\omega d_z z}{c}} + r e^{j\frac{\omega d_z z}{c}} \right) e^{j\frac{\omega}{c}(ct - d_x x - d_y y)} d\omega. \end{aligned}$$

The last step is to compute the intensity of U. Observe that due to the position of the mirror, the interferences patterns are only created along the z-axis and not along the x- and y-axes. Let $d_z = \cos(\alpha)$, where α is the angle between $\boldsymbol{d}$ and the z-axis. We can generalize the Lippmann

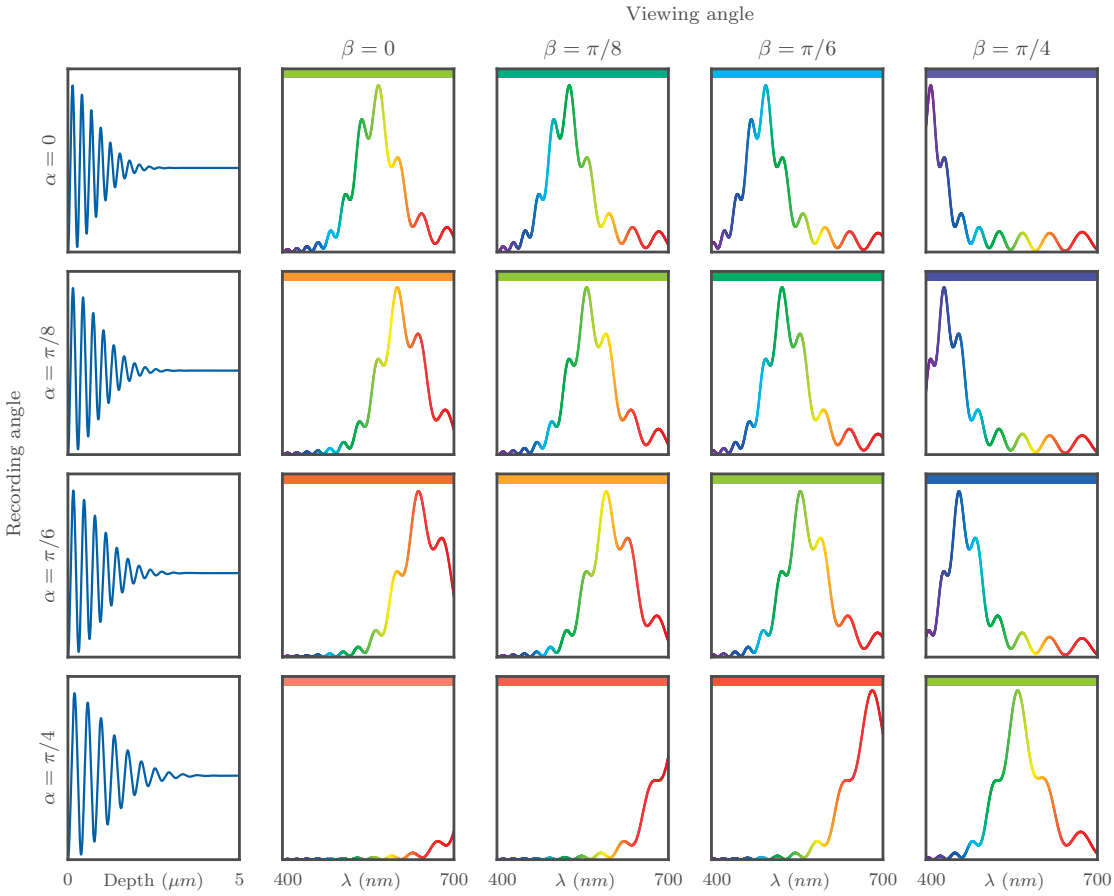


Figure 4.10: Reflected waves given the spectrum from Figure 4.5 for different combinations of recording angle α and viewing angle β. The first column represents the interference patterns captured in a plate of thickness 5 μm and with $r = -1$. For every spectrum we display the corresponding RGB color on top of the graph. When $\alpha = \beta$, the synthesized spectrum has a shape similar to the original one, the only difference is that the period of the oscillations increases with the angle: this is due to the fact that oblique angles artificially decrease the relative thickness of the plate.

operator as

$$\begin{aligned}\mathcal{L}_{r,\alpha}\{P\}(z) &= \int_0^\infty P(\omega)\left(1+\rho^2+2\rho\cos\left(\frac{2\omega z}{c}\cos(\alpha)+\theta\right)\right)d\omega \\ &= (1+\rho^2)\mathcal{F}_0\{P\}(0)+2\rho\mathcal{F}_\theta\{P\}\left(z\cos(\alpha)\right).\end{aligned}$$

The result is surprisingly close to (4.2). The only difference is the $\cos(\alpha)$ factor, which creates a dilation of the interference patterns.

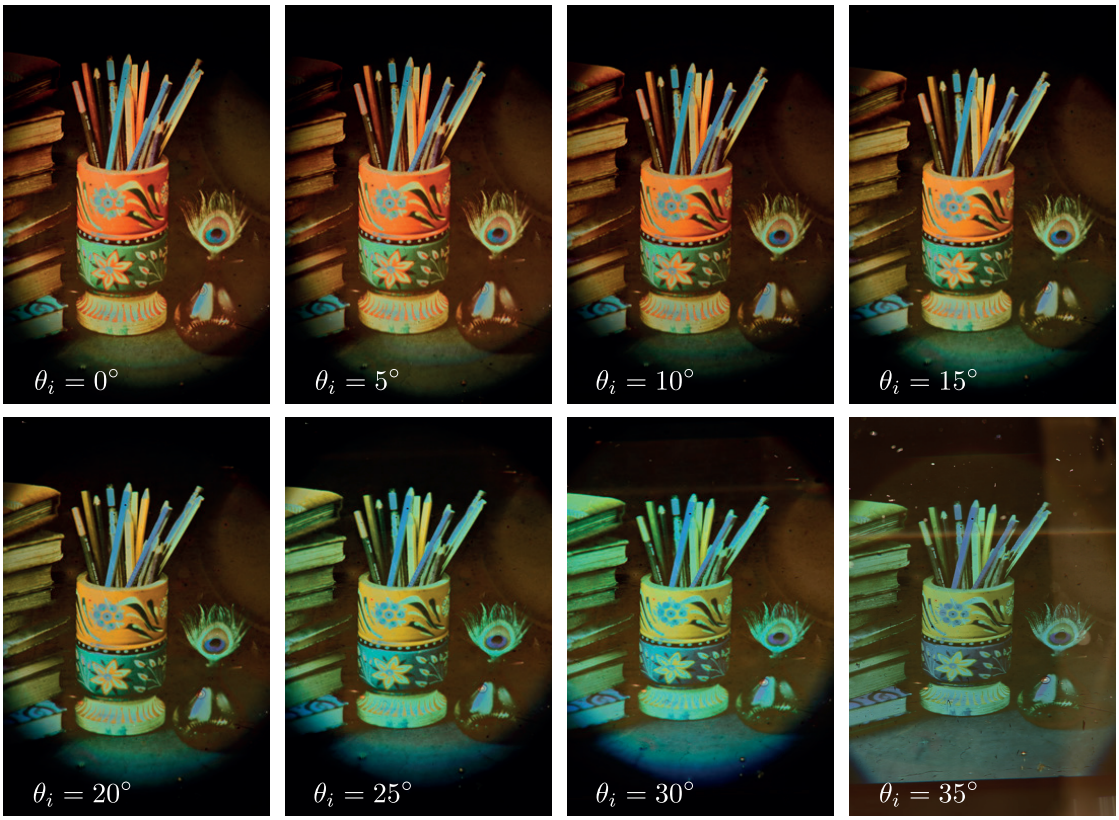


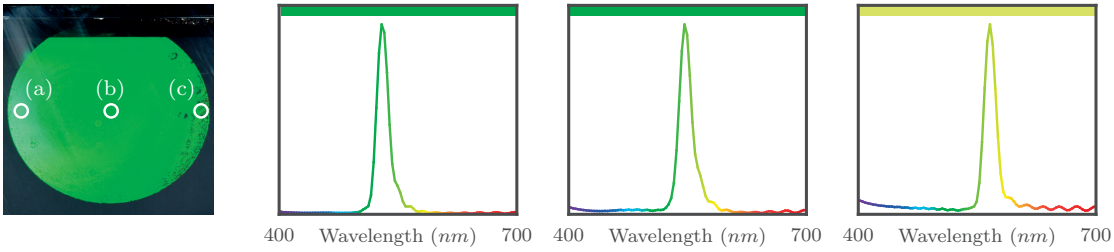
Figure 4.11: Spectrum shift due to different viewing angles: the plate from Figure 4.2 is photographed under eight different angles θ_i, ranging from 0° to 35°. Note that θ_i is measured with respect to the surface normal of the prism (which has an angle of 10°) and not the plate (see Figure 4.6b). We observe that the image suffers from perspective distortion and the scene appears shorter for larger angles. Remark that the images have been acquired with an RGB camera, which cannot reproduce the richness and complexity of the spectra.

4.4.2 Viewing

The dual problem to recording a Lippmann plate at a specific angle is observing it from a different angle: we now assume that we do not necessarily watch the Lippmann plate under the angle $\boldsymbol{\omega}_o = [0, 0, 1]^\top$. Assuming that the silver particles behave as purely specular bodies[8], the only contribution of the light is from the angle $\boldsymbol{\omega}_r$, which is defined as the mirror of $\boldsymbol{\omega}_o$ with respect to the surface normal $\boldsymbol{n} = [0, 0, 1]^\top$.

Observe that the round-trip distance covered by the light between the entrance of the plate and a depth z is given by $2z/\cos\beta$, where β is the angle between $\boldsymbol{\omega}_o$ and $\boldsymbol{n}$ (see Figure 4.6b). The second main difference is that, since the viewing angle is oblique, the integral path changes. Therefore, we should generalize the reflectance function $R(z)$ of the plate to be spatially varying in x and y. However, since the thickness of the plate is relatively small, we can assume that the

[8]We measured the bidirectional reflectance distribution function (BRDF) of Lippmann plates and we verified that a perfectly specular model is a good approximation to characterize the reflectance function of the silver particles.



(a) Original plate (b) 531 nm (c) 547 nm (d) 566 nm

Figure 4.12: Spectrum shift due to different incident angles: (a) a Lippmann plate exposed under a monochromatic light (531 nm) arriving at different incident angles. The extreme left of the disk corresponds to an angle of 0° with respect to the surface normal. As we move towards the right side of the disk, the incident angle increases. After development, the plate is photographed with a hyperspectral camera at the marked spots. The resulting spectra are shown in (b), (c), and (d), along with the measured wavelength of the peak.

spatial variation is negligible; this allows us to simplify notation. We further assume spatially-uniform lighting and factor the term $L(\omega')$ outside of the integral. With these assumptions, the reflected wavefunction (4.5) can be rewritten as

$$U_\beta(\omega', t) = \sqrt{L(\omega')} e^{j\omega' t} \int_0^Z R(z) e^{-j \frac{2\omega' z}{c \cos \beta}} dz.$$

The key difference with (4.5) is the term $\cos\beta$ in the exponential. Following a similar reasoning as in Section 4.3.1, we obtain

$$U_\beta(\omega', t) = \epsilon \sqrt{L(\omega')} e^{j\omega' t} \int_0^\infty P(\omega) H^Z \left(\omega, \frac{\omega'}{\cos \beta} \right) d\omega.$$

Therefore, the wavelengths λ' of the original spectrum will be scaled down by a factor $\cos\beta$. When represented with respect to the wavelength, spectra appear to undergo a color shift towards the blue colors. Figure 4.10 shows the reflected spectrum for a few pairs of recording/viewing angles. Interestingly, for $\alpha = \beta$, the two effects neutralize each other. Yet, there is a notable difference in the sense that the oscillations do not have the same period. In fact, their period increases as α and β increase. This is due to the fact that the effective thickness of the plate is scaled down by a factor $\cos\alpha$.

4.4.3 Experiment

The effect due to different viewing angles can be easily verified by photographing a Lippmann plate from different view points. Figure 4.11 shows the plate from Figure 4.2, photographed under eight different angles θ_i. These angles are measured with respect to the surface normal of the prism mounted on the plate and they range from 0° to 35°. As predicted by our analytical model, we observe a shift of the spectrum toward the blue colors.

Similarly, the shift caused by different incident angles can also be observed experimentally. Figure 4.12 shows a plate that we created with a laser beam. The beam was first expanded and then made divergent with a convex lens in such a way that the rays hit the plate with increasingly

large incident angles. More precisely, the rays hit the left side of the plate perpendicularly with respect to the surface normal, and the further to the right, the larger the incident angle. The plate has then been developed and photographed with a hyperspectral camera. As showcased in Figure 4.12b-d, we can observe a spectrum shift towards the red colors as predicted by our model.

4.5 Color and spectrum reproduction

We now use our mathematical model to study the effects and artifacts generated by Lippmann's procedure. A common misconception in the literature is that the Lippmann method enables perfect reconstruction of the original spectrum, at least from a theoretical point of view; that is, assuming a plate of infinite thickness, perfectly uniform development and a re-illumination with a flat spectrum. In fact, Lippmann himself made such a claim in his 1894 paper, using the premises of Fourier analysis [103]:

"en d'autres termes, la couleur de l'image est la même que celle de l'objet."[9]

More recently, H. Bjelkhagen, an expert in Lippmann photography and holography, wrote [30]:

"The light is reflected from these fringes, creating precisely the colours that correspond to the original ones that had produced them during the recording."

Both these claims are backed up with equations that unfortunately contain errors. We see in this section that even though colors are reproduced relatively accurately, the synthesized spectrum is not the same as the original one, even in a 'perfect' world, assuming an infinitely thick plate and no windowing due to the processing. To gain some intuition, we refer to Figure 4.13a, which shows monochromatic and Gaussian spectra, and Figures 4.13b-d, which illustrate their reproduced spectra as predicted by (4.10) for plates of differing thicknesses. We notice that, while the overall shape of the synthesized spectra is relatively accurate, it contains several artifacts. To simplify the analysis of these distortions, we assume that the illuminating light source has a constant power spectral density[10] and we rewrite the expression (4.10) as

$$\begin{aligned}\mathcal{E}_r^Z\left\{P\right\}(\omega') &= \left|\int_0^\infty P(\omega)H^Z(\omega,\omega')d\omega\right|^2 \\ &= \left|\frac{r}{2}P^Z(\omega') + \frac{1+\rho^2}{2}Q^Z(\omega') + \frac{r^*}{2}(P^Z)^*(-\omega')\right|^2,\end{aligned}$$

where

$$\begin{aligned}Q^Z(\omega') &= S^Z(\omega')\int_0^\infty P(\omega)d\omega, \\ P^Z(\omega') &= \int_0^\infty P(\omega)S^Z(\omega-\omega')d\omega.\end{aligned}$$

Aided by Figure 4.13, we identify four effects that are worth further investigation.

[9] "in other words, the color of the image is the same as that of the object." (translated by the author)

[10] In practice, this is not the case, but we can always measure it and calibrate our measurements accordingly.

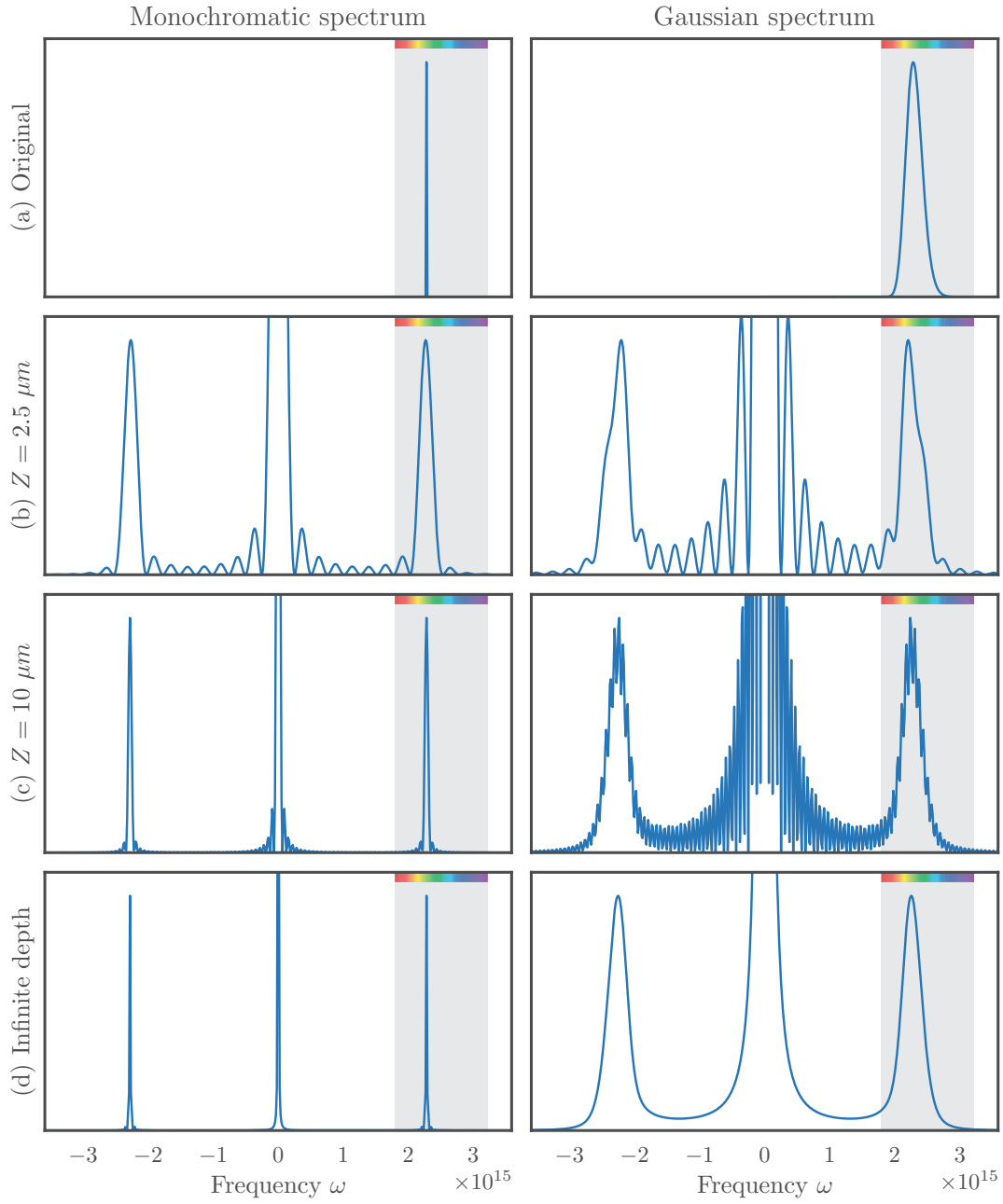


Figure 4.13: Influence of the thickness of the plate on the reflected spectrum. Given (a) a spectrum recorded in a plate of thickness Z, we show the reflected spectrum for (b) $Z = 2.5\ \mu$m, (c) $Z = 10\ \mu$m, and (d) $Z = \infty$. The value Z controls two effects: the sharpness of the synthesized spectrum (this is especially noticeable with monochromatic light that has high frequency components) and the frequency $Z/(\pi c)$ of the ringing artifacts. Shaded areas represent visible light.

1. First, recall that we do not observe directly $U_r(\omega', t)$, but rather its corresponding time-averaged intensity, which is proportional to the squared magnitude of $U_r(\omega', t)$. As a result, the filtered power spectral density P is squared; this contributes to increase the contrast in the synthesized colors and explains why Lippmann plates tend to exhibit a wider tonal range as well as strong and bold colors.

2. Second, S^Z acts as a lowpass filter on the spectrum. This effect confirms early experiments conducted by Ives [81], who reported that the thickness of the plate controls the spectral resolution. Additionally, the imaginary part of S^Z distorts the overall shape of the spectrum: since we observe the squared magnitude of the analytic signal U_r, its real and imaginary parts are blended together. This results in broader spectra.

3. Third, the term $Q^Z(\omega')$ contains some strong oscillations that leak into the visible spectrum, as illustrated in Figures 4.13b-c. As seen in Figure 4.5, the interference patterns tend to stabilize around a value proportional to the energy of $P(\omega)$. The relative strength of the oscillations is influenced by this value: a signal with its energy spread over the entire visible spectrum will have more pronounced relative oscillations.

4. Finally, the heavy tails of the term $Q^Z(\omega')$ and in particular its imaginary component also influences the visible spectrum. Perhaps surprisingly, this effect depends on the reflection coefficient r.

To the best of our knowledge, points 2 to 4 have never been documented. We will now investigate these concepts in more detail.

4.5.1 Thickness of the plate

The recording of interference patterns in the plate is in fact a generalized Fourier cosine transform of the original spectrum, where lower frequencies are located at the entrance of the plate and higher frequencies deeper in the plate. From that perspective, it is clear that the thickness of the plate has a lowpass effect on the synthesized spectrum. Not only are the high frequencies clipped in the synthesized spectrum, but also the sharp cutoff at the end of the plate results in ringing artifacts due to the Gibbs phenomenon, which can be observed in Figure 4.13.[11]

Another direct consequence of the finite thickness of the plate is the middle component $Q^Z(\omega)$, which also generates oscillations whose frequency depends on Z.

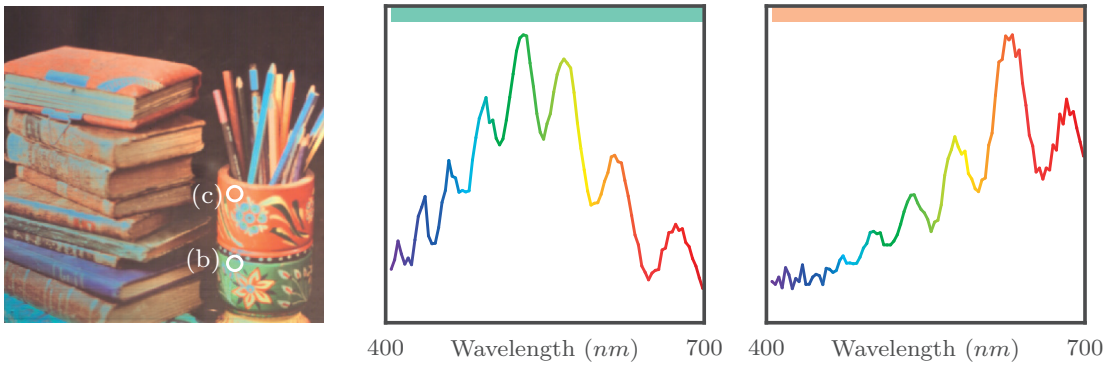
Plate with infinite thickness

When we assume that the plate has infinite thickness, we have

$$S^\infty(\omega) = \lim_{Z\to\infty} S^Z(\omega) = \frac{\pi c}{2}\delta(\omega) + \frac{jc}{2\omega}. \tag{4.11}$$

As mentioned, it is often claimed in the literature [30, 103] that an infinite thickness is a condition for a perfect recovery of the spectrum (up to the square factor described above). While this might

[11] To gain better intuition regarding the influence of components located outside of the visible spectrum, we choose to display spectra as a function of the frequency—as opposed to a function of the wavelength as is often the case.



(a) Lippmann plate (b) green spectrum (c) red-orange spectrum

Figure 4.14: Hyperspectral acquisition of the Lippmann plate from Figure 4.2: (a) close-up of the plate in RGB colors; (b) a green spectrum; (c) a red-orange spectrum measured on the pot. The spatial location of the spectra displayed is shown in (a) and the uniform band on top of each spectrum plots represents the corresponding RGB color.

seem true for a monochromatic spectrum, it is in general not the case. Given (4.11), the reflected wavefunction is given by

$$U_r(\omega', t) = \epsilon\sqrt{L(\omega')}e^{j\omega' t}\left(\frac{r}{2}P^{\infty}(\omega') + \frac{1+\rho^2}{2}Q^{\infty}(\omega') + \frac{r^*}{2}(P^{\infty})^*(-\omega')\right).$$

We see that the Hilbert component of the signal is still present in the reconstruction. In other words, when Z goes to infinity, $U_r(\omega', t)$ contains the analytic signal P^{∞} of P, defined as the sum of P and its Hilbert transform:

$$P^{\infty}(\omega) = \frac{\pi c}{2}P(\omega) + \frac{j\pi c}{2}\int_{-\infty}^{\infty}\frac{P(\omega-\nu)}{\pi\nu}\,\mathrm{d}\nu.$$

Moreover, the term $Q^{\infty}(\omega')$ located around frequency zero does not vanish and its imaginary part still alters the synthesized spectrum by boosting its tails in the visible range. In this case, the amplitude of $Q^{\infty}(\omega')$ is controlled by the energy of the spectrum. In the example of Figure 4.13, we see that the Gaussian spectrum has more energy than the monochromatic one, therefore the influence of $Q^{\infty}(\omega')$ is more pronounced. The observed intensity is then proportional to the squared magnitude of $U_r(\omega', t)$.

Experiment

As shown in Figure 4.14, we performed a hyperspectral acquisition of the plate from Figure 4.2. It is clear that the spectra exhibit the oscillations predicted by our reflection-based theoretical model. Furthermore, we can approximate the thickness of the plate from these oscillations. The spectra from Figures 4.14b and 4.14c present 8 oscillations in the visible frequency range $\Delta\omega$, so they have a period of $T = \frac{\Delta\omega}{8}$. Comparing this value with the expression for $S^Z(\omega)$, we deduce that $\frac{2\pi}{T} = \frac{2Z}{c}$, or $Z \approx 3.7$ μm. This value is in line with the reported plate's thickness.

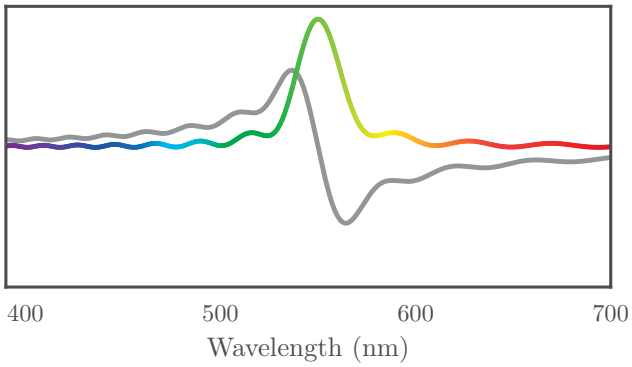


Figure 4.15: The filter $S_\tau^Z(\omega' - \omega)$ for $Z = 5\ \mu m$, $\tau = 2$ and $\omega' = \frac{2\pi c}{550\ \text{nm}}$. Compared to Figure 4.7, the oscillations are less pronounced.

4.5.2 Processing and windowing effect

Like the thickness of the plate, the development, and more generally the processing, also influences the shape of the resulting spectrum. For instance, as the developer diffuses into the depth, the bottom of the plate will be less affected by the development than the top.

The concentration $D(z,t)$ of the developer—at depth z and after time t—obeys the diffusion equation, whose solution is given by the complementary error function [140]:

$$D(z,t) = 1 - \operatorname{erf}\left(\tau_0 \frac{z}{\sqrt{t}}\right), \tag{4.12}$$

for some constant τ_0. The cumulated effect of the diffusion can then be represented by integrating (4.12) over the development time T. For ease of calculation, we approximate this expression with a decaying exponential:

$$D_\tau(z) \approx e^{-\frac{\tau z}{Z}},$$

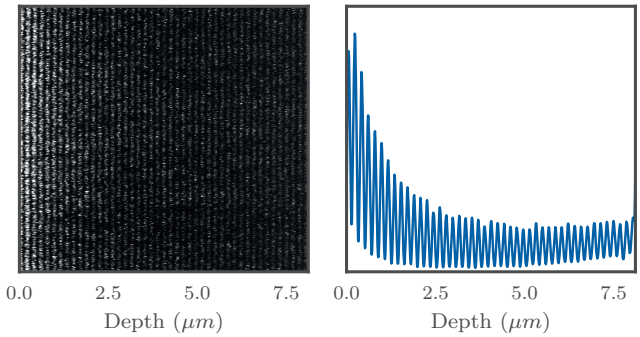
where τ is the decay rate. Taking into account the development, we can replace the expression (4.4) for the silver density $R(z)$ with

$$R(z) = \epsilon D_\tau(z) \mathcal{L}_r \{P\}(z).$$

Note that this windowing function can also encompass other effects; for instance, the drying and hardening of the plate, the light intensity decay and the decrease of fringe visibility inside the plate can all be modeled by a decaying windowing function.

We can also generalize the filter S^Z such that it depends on the decay parameter. Following a similar reasoning as in (4.8) and (4.9), we can compute a closed-form formula for this generalized filter:

$$\begin{aligned} S_\tau^Z(\omega) &= \int_0^Z D_\tau(z) e^{-j\frac{2\omega' z}{c}} dz \\ &= \frac{cZ}{c\tau + 2j\omega Z}\left(1 - e^{-2j\frac{Z\omega}{c} - \tau}\right). \end{aligned}$$



(a) FIB image (b) Column averaging

Figure 4.16: Capture of the interference patterns: (a) Electron microscope image of a slice of a Lippmann plate created with a 531 nm laser. After fixing, we can observe interference patterns created by the varying distribution of the metallic silver grains. These patterns are clearly visible after averaging the columns (b).

An example of such a filter is shown in Figure 4.15. Due to its decaying nature, the development curve attenuates the high frequencies and has a smoothing effect on the oscillations created by the sharp cutoff at the end of the plate.

Experiment

To underline the effect of the development, we propose to look directly at the distribution of developed silver inside a plate. The variations in the distribution of metallic silver are complicated to observe with optical microscopes since the size and spacing of the patterns are below their resolving power. At the beginning of the last century, physicists found workarounds by first soaking the plates in water, essentially dilating the plate by an order of magnitude [81]. Another manoeuvre was to slice the plates with an angle, thereby expanding the length of the observed face and the shape of the silver patterns.

Nowadays, we can use electron microscopy to obtain sub-wavelength resolution imaging. Electron microscopes cannot distinguish between undeveloped silver halide and metallic silver particles. To see interference patterns, it is necessary to wash out the remaining silver halide; this can be done by fixing the plate. Figure 4.16 shows an example of the metallic silver particles density resulting from the exposure of a plate to a 531 nm laser.

The interference pattern is clearly noticeable in the micrograph from Figure 4.16a as well as when averaging it across pixels located at the same depth in Figure 4.16b. Additionally, we can observe in Figure 4.16b a decay as the depth increases, and while we cannot precisely identify which phenomenon (development, drying, light absorption, ...) causes it, we notice that it can be modeled, fairly accurately, with an exponential function.

4.5.3 Dependance on the reflection coefficient

Another parameter that alters color rendition is the reflection coefficient $r = \rho e^{j\theta}$. We voluntarily described our model as a function of r and did not restrict it to $r = 1$ or $r = -1$ as is commonly done in the literature. We now investigate the effects of changing it. In Lippmann photography,

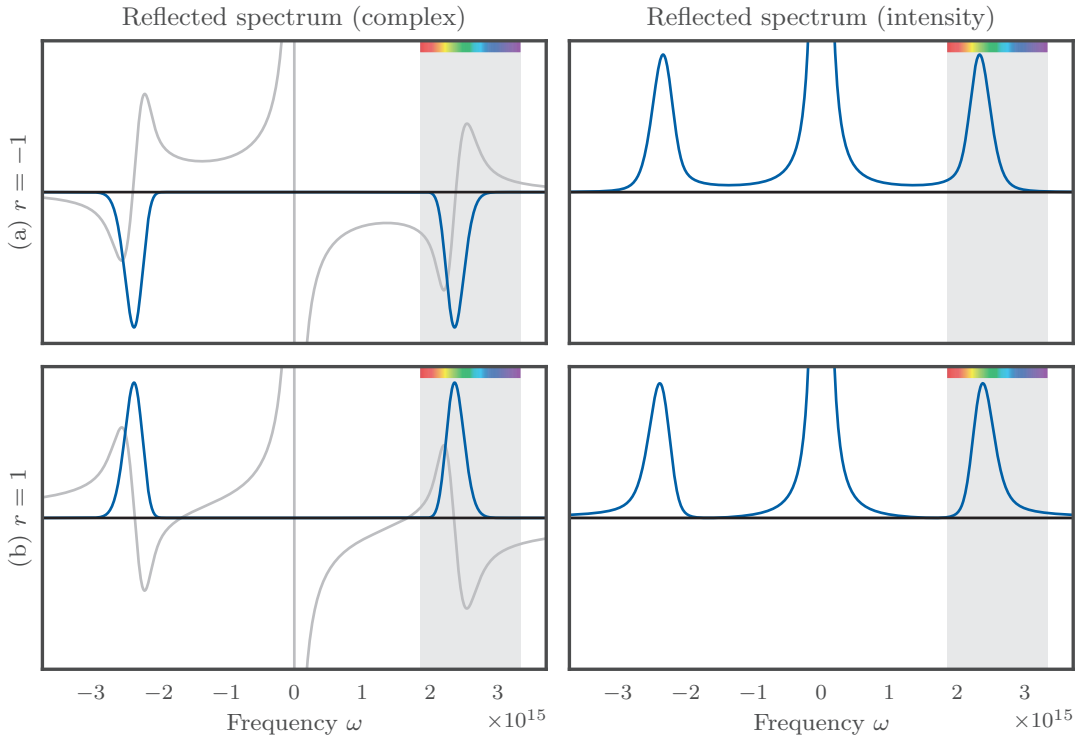


Figure 4.17: Skewing effect on the reflected spectrum by an infinite plate that recorded the Gaussian spectrum from Figure 4.5 is for (a) $r = -1$ and (b) $r = 1$. The left column shows the complex wavefunction (the blue curve is the real part and the gray curve the complex part) and the right column its observed intensity. We observe that the top reflected spectrum ($r = -1$) exhibits a negative skew and the bottom reflected spectrum ($r = 1$) a positive skew.

there are essentially two mediums that are used to create interferences: mercury and air. For an interface between glass and mercury $\rho = 0.71$ and $\theta = -148°$, and for an interface between glass and air $r = 0.2$, or equivalently, $\rho = 0.2$ and $\theta = 0$. It is evident that different values of r will generate different interference patterns inside the plate. For example, with $r = 1$ a peak is formed at the entrance ($z = 0$) of the plate, whereas with $r = -1$, a node is located at $z = 0$. What interests us here is how these different values influence the synthesized spectra. Let us observe the expression for H^Z in (4.9). The side components P^Z and $(P^Z)^*$ are modulated with r and r^*, hence they undergo a rotation in the complex plane that is determined by the value of θ. On the other hand, the center component Q^Z is modulated by $1 + \rho^2$ and is never rotated. This means that these two kinds of components interfere with each other differently for different values of r.

Figure 4.17 shows what we call a *skewing* effect on the spectrum. For $r = -1$, the imaginary part of $S^Z(\omega')$ and $-S^Z(\omega' - \omega)$ add up constructively for red colors and destructively for blue colors, as seen in Figure 4.17a. On the other hand, for $r = 1$, Figure 4.17b shows the opposite phenomenon, namely the spectrum is skewed such that it favors blue colors over red colors. In practice, when the reflection is due to the interface with air, we have $r = 0.2$, so we should

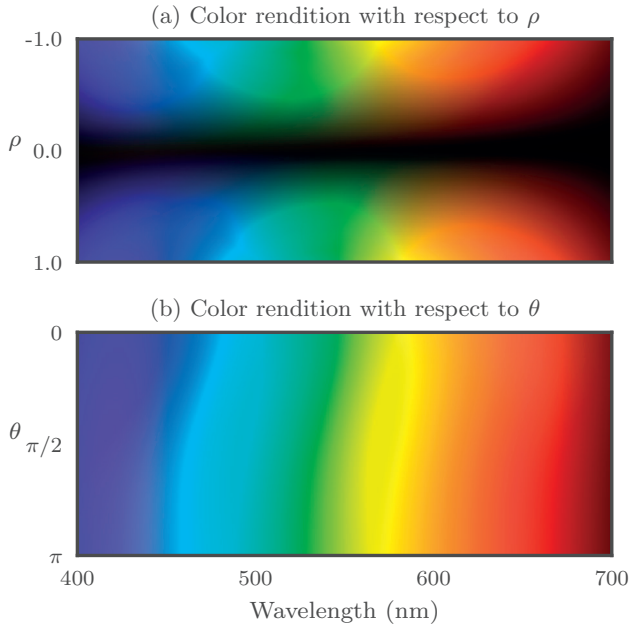


Figure 4.18: Skewing effect on the spectrum: Given an original Gaussian spectrum centered at the different visible wavelengths and with $\sigma = 30 \cdot 10^{-9}$, we show the reproduced colors for (a) varying amplitudes ρ and (b) phases θ of the reflection coefficient r. The spectra are first converted to the CIE 1931 XYZ color space and then to RGB.

expect stronger blues, to the detriment of the reds since $\theta < \pi/2$. With mercury we have $\rho = 0.7$ and $\theta = -148°$ and therefore, red colors should predominate since $\theta > \pi/2$. Given a Gaussian spectrum centered around a specific wavelength, Figure 4.18 illustrates the expected reproduced colors for different values of ρ and θ. It suggests that bright reds are more challenging to reproduce with air, and blues are harder to represent with mercury. We also notice a shift of the reflected spectra: for instance, greens should appear blueish with air and yellowish with mercury. Moreover, a smaller value $|\rho|$ leads to darker colors; this explains why, in general, plates made with mercury appear brighter than those made with air.

As a final remark, note that the strength of the skewing effect is correlated with the spread of the energy of the original spectrum. A spectrum that is concentrated around a single frequency will have a relatively much weaker center component, and as a consequence the skewing effect is less noticeable. In the extreme case where the original spectrum is a Dirac, the tails of the center component are not strong enough to influence the visible spectrum, as illustrated in Figure 4.13c-d.

Experiments

Monochromatic light To analyze the skewing effect, we decomposed white light through a prism with a setup similar to Newton's crucial experiment (see Figure 1.5). This experiment was also famously created on photosensitive plates with a piece of glass coated with silver by Wiener in 1890 [197]. All visible wavelengths are represented and the light of the spectrum obtained

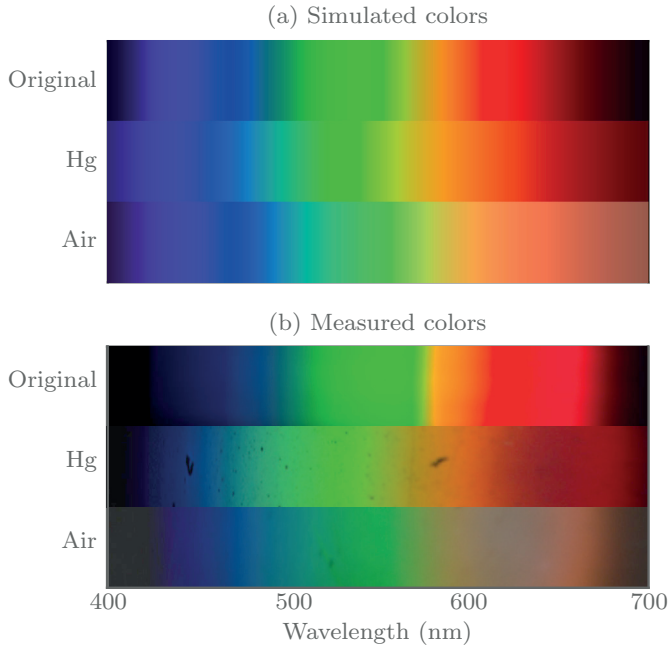


Figure 4.19: Color reproduction experiment: (a) simulated colors and (b) measured colors. To measure the colors, we first decomposed white light into monochromatic colors with a prism and recorded Lippmann photographs using mercury (Hg) and air reflectors. The resulting reflected spectra and the original spectrum were then recorded using a digital camera. The black speckles in the plate with mercury are due to tiny particles of dust floating on the mercury surface.

from this decomposition is locally monochromatic. Under such light, we exposed two Lippmann plates: one using mercury as reflector and one using the interface with air. We then developed these two plates and photographed them. For reference, we also photographed the light spectrum coming from the prism: the original spectrum along with the spectra of the two plates are shown in Figure 4.19b.

To quantify the accuracy of our model, we also run a simulation of this experiment using our analytical model, illustrated in Figure 4.19a. To be as close as possible to the experiment, we set the thickness of the plate to $Z = 7 \cdot 10^{-6}$ μm and combined it with a development window $D(z) = 1 - \operatorname{erf}(3z/Z)$. We observe that the measured colors closely resemble the ones from our simulations. In particular, the purple colors are weaker in plates made with mercury while the plates made with air struggle at reproducing bright reds. Also, the reproduced spectra are subject to some shifts: for instance, the green colors tend to have a stronger red tint when using mercury and a more dominant blue component with air.

Gaussian-shaped light We run a second experiment by placing Gaussian-shaped bandpass filters having a bandwidth of approximately 10 nm in front of a halogen light source. As before, we create two plates: one with mercury and one with air as reflectors. Photographs of the resulting plates along with the color of the filters are shown in Figure 4.20a. For this experiment, we also acquired hyperspectral measurements of the original spectra as well as the spectra reflected by

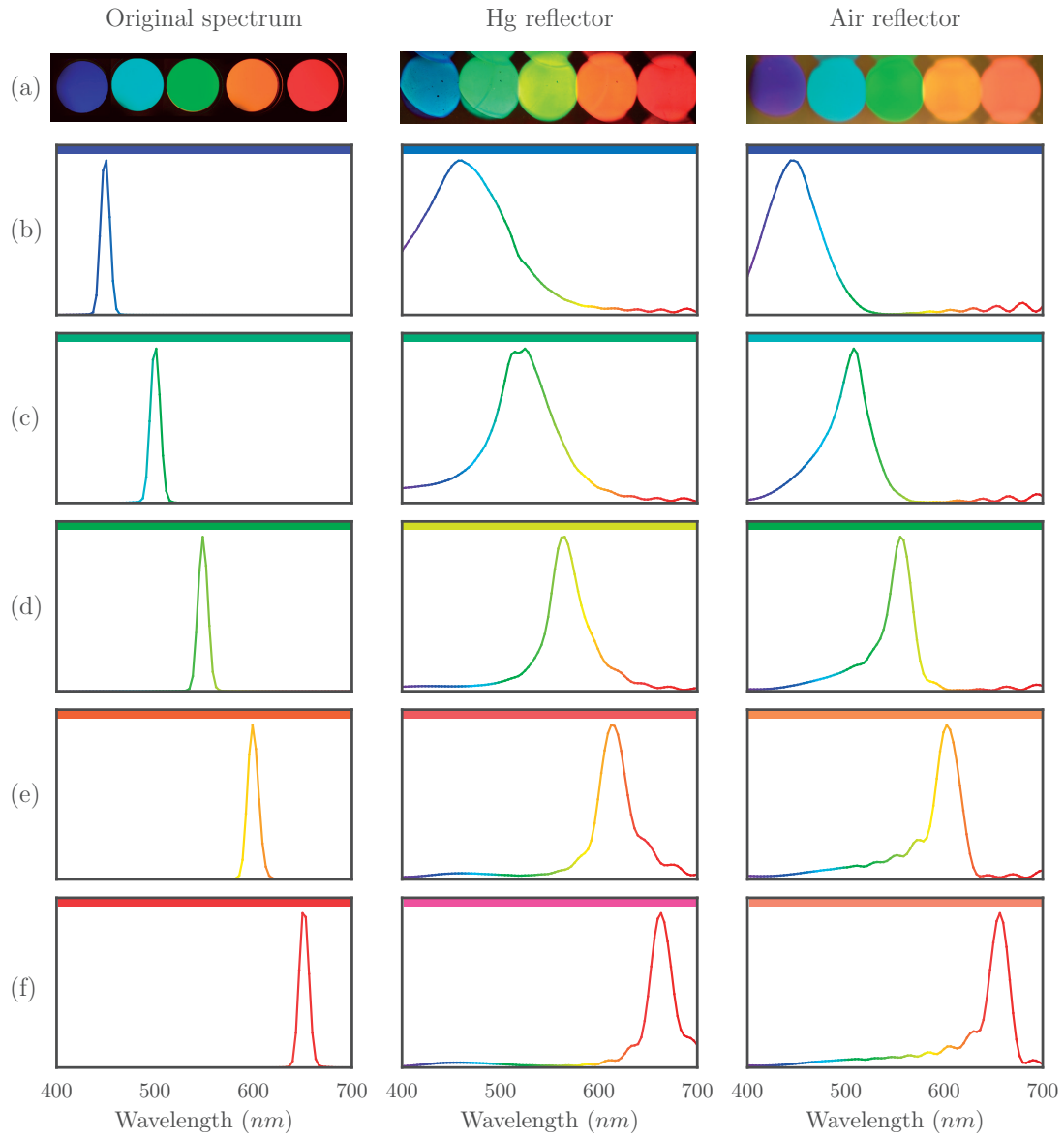
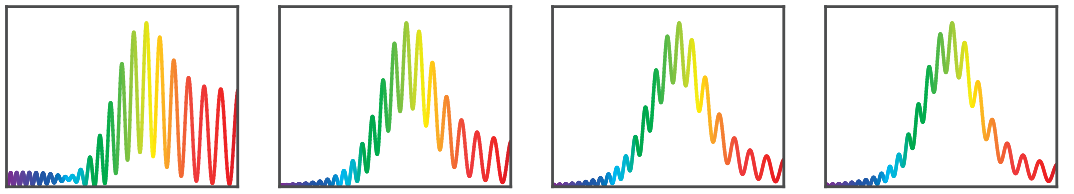


Figure 4.20: Skewing effect on Gaussian-shaped spectra: (a) the left column shows the five filters used to expose the plates, while the middle and right columns show the resulting plates after development. Additionally, we show the spectral data corresponding to each of these five bands: (b) 450 nm, (c) 500 nm, (d) 550 nm, (e) 600 nm, and (f) 650 nm.

these plates (see Figures 4.20b-f). We see that the reproduced colors are in agreement with our predictions. Indeed, the spectra of the plates realized with a mercury reflector exhibit a negative skew while those made with an air reflector have a positive skew. Moreover, the synthesized spectra appear broader that the original one; the main reason is that the imaginary part of the reflected wave combines with the real part in the squared magnitude operation, resulting in



(a) $r = -0.1$ (b) $r = -0.25$ (c) $r = -0.5$ (d) $r = -1$

Figure 4.21: Reflected spectra based on the Gaussian spectrum from Figure 4.5, for $r = -0.1, -0.25, -0.5, -1$. We observe that the oscillations are proportionally more important in spectra with lower intensities. The spectra have been rescaled such that they appear similar, but in practice the proportion of reflected light is higher for higher absolute values of r.

larger spectra.

Regarding the oscillations, we notice two effects: first, they are not so significant, which suggests that they are damped by a development curve strongly decaying with the depth; second, their amplitude is slightly larger in the plates made with air. This latter effect is studied in greater detail in the next section.

Strength of oscillations

Besides skewing the reflected spectrum, the reflection coefficient also has some influence on the strength of the oscillations. When $\rho < 1$, the interferences patterns take the form of partial standing waves. The baseline term in (4.1) is multiplied by the value $1+\rho^2$ and the oscillating part by the value 2ρ. Subsequently, the baseline becomes relatively stronger than the relevant part of the signal as ρ becomes smaller. This baseline is then directly mapped to the center component $Q^Z(\omega)$ in the reflected spectrum. As a result, the spectrum contains stronger oscillations for smaller values of ρ. This effect is illustrated in Figure 4.21. It also explains why the oscillations are more intense in plates made with air reflectors than those made with mercury reflectors.

4.6 An alternative: A refractive index based model

In the previous sections, we assumed that the creation of colors was due to the direct reflection of light from the silver particles. A second school of thought was first suggested by Lippmann [103] and further discussed in an article by Wiener [198]; they postulated that the *refractive index* $n(z)$ at depth z of the exposed plate could be modified proportionally to the intensity of the light field, or more precisely to the varying part of the intensity of the interference field:

$$n(z) = n_0 + \gamma \mathcal{F}_\theta\{P\}(z),$$

where n_0 is the refractive index of the medium before exposition. Since the density of silver particles in the gelatin is proportional to $\mathcal{F}_\theta\{P\}(z)$, the index of refraction varies accordingly. Because silver has a higher index of refraction than gelatin, the overall refracting index increases when the density of silver is higher. An analysis similar to the one from Section 4.3.1 has been conducted by Nareid and Pederson [126], but focusing on this refractive index based model.

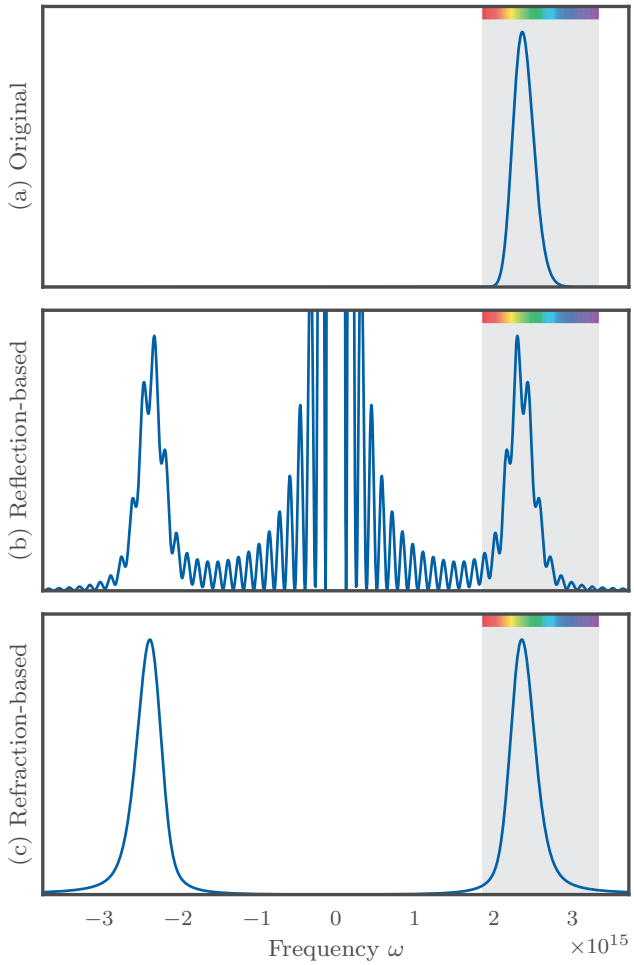


Figure 4.22: Simulation of (a) the spectrum from Figure 4.5 reflected by a plate of width $Z = 5\ \mu\text{m}, r = -1$ and with a constant illumination $L(\omega) = 1$ according to (b) the reflection-based model and (c) the refractive index based model.

Using the so-called first Born approximation [200], and assuming relatively small changes in the index of refraction of the medium, it is possible to obtain a closed-form solution similar to (4.6); indeed, we can extend Nareid and Pederson's analytic model using the principles from Section 4.3.1. This results in

$$U_r(\omega', t) = \gamma \frac{\omega'}{c} \sqrt{L(\omega')} e^{j\omega' t} \int_0^\infty P(\omega) F^Z(\omega, \omega') d\omega, \tag{4.13}$$

where the function $F^Z(\omega, \omega')$ is defined as

$$F^Z(\omega, \omega') = \frac{r}{2} S^Z(\omega' - \omega) + \frac{r^*}{2} S^Z(\omega' + \omega).$$

As in (4.10), the time-averaged intensity of the reflected wavefunction is

$$\mathcal{E}_r^Z\left\{P\right\}(\omega') = \epsilon^2 \left(\frac{\omega'}{c}\right)^2 L(\omega') \left| \int_0^\infty P(\omega) F^Z(\omega, \omega') d\omega \right|^2 .$$

4.6.1 Comparison with the reflection-based model

An example of a spectrum and its synthesized version according to the reflection-based and refractive index based models is shown in Figure 4.22. Due to the similarities between the two models, the analysis performed in Section 4.5 can be easily applied to this alternative model. If anything, some parts can be simplified due to the absence of the component located at frequency zero. As a consequence, the reproduced spectra due to each theory differ on a couple of points. First, there is an additional squared dependency on the frequency in (4.13). Another notable difference is stressed in Figure 4.22c: the refractive index based hypothesis does not contain oscillations. Indeed, $F^Z(\omega, \omega')$ does not incorporate the term $Q^Z(\omega')$ which leaks into the visible spectrum. As a consequence, this model is not subject to the spectrum skewing effect. The presence or absence of these effects in a measured plate can help us identify the model and consequently the dominant underlying physical process involved.

Experiments

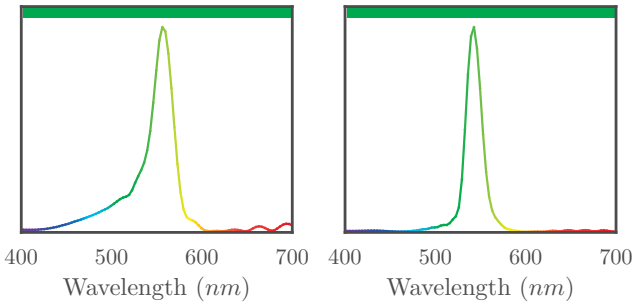
Interestingly, both models are legitimate, each in its own specific domain of application. We claim that depending on the developing procedure, the dominating effect recreating the reflected spectrum can be either due to a variation of the reflectivity or local changes of the refractive index. More precisely, if *bleaching*—a post-development bath which essentially converts the grains of metallic silver back into silver halide crystals—is applied after development, we are in a regime dominated by the refractive index based model. On the other hand, when no bleaching is applied, we are in a reflection-based regime. To illustrate this statement, we created two plates using the 550 nm bandpass filter from Figure 4.20d. One plate was developed using the historical Lumière developer, which was one of the typical Lippmann developers in the early 1900s, and the other one was developed and bleached with holographic recipes.[12] The spectra synthesized by these plates are shown in Figure 4.23. We notice that the spectrum from Figure 4.23a clearly exhibits the skewing effect as well as moderate oscillations; to the contrary, the spectrum from Figure 4.23b is relatively symmetric and does not present any of these effects.

We choose to exclusively focus the rest of our analysis on the reflection-based model, as it can easily be extended to the somewhat simpler refractive index based model.

4.7 Can we invert the Lippmann operations?

In Section 4.5, we concluded that the spectra reflected from Lippmann plates are distorted versions of the originals. If we could invert these artifacts, we would have access to 100 year old hyperspectral measurements of natural scenes by applying hyperspectral imaging techniques on historical Lippmann plates. We investigate below which of these artifacts can be digitally reversed, the goal being to get as close as possible to the original spectrum of the photographed scenes.

[12] For more details about the different types of developers and development techniques, refer to Appendix 4.A.3.



(a) Lippmann development (b) Holographic development

Figure 4.23: Two different development procedures and their effects: (a) the Lumière developer results in spectra that can be explained with the reflection-based model, and (b) holographic recipes produce spectra that can be inferred from the refractive index based model.

4.7.1 Recording

The recording operation is represented by the operator $\mathcal{L}_r\{P\}(z)$ in (4.2) and is made of the sum of two terms: a constant term $(1+\rho^2)\mathcal{F}_\theta\{P\}(0)$ and an oscillating term $2\rho\mathcal{F}_\theta\{P\}(z)$. To remove the influence of the former, we simply subtract the mean μ of $\mathcal{L}_r\{P\}(z)$:

$$\mathcal{L}_r^0(z) = \mathcal{L}_r\{P\}(z) - \mu. \tag{4.14}$$

We are left with the generalized cosine transform from (4.3), for which there exists an analytic solution: as shown in [121], (4.14) can be inverted as

$$P(\omega) = \frac{2}{c\pi\Gamma(\nu+1)} \int_0^\infty (\omega z)^\nu \left({}_1F_1\left(1; 1+\nu; j\frac{2\omega z}{c}\right) + {}_1F_1\left(1; 1+\nu; -j\frac{2\omega z}{c}\right)\right) \mathcal{L}_r^0(z)dz,$$

where $\Gamma(\cdot)$ is the Gamma function, ${}_1F_1(\,\cdot\,;\,\cdot\,;\,\cdot\,)$ is Kummer's confluent hypergeometric function [97], and $\nu = 2\theta/\pi \in [0,1]$. Notice that the above inversion assumes an infinite plate. For plates of finite thickness, the higher frequency components are obviously lost in the process. In this case, perfect inversion is guaranteed for bandlimited signals; otherwise, the above formula gives the projection of $P(\omega)$ onto the subspace of bandlimited functions.

4.7.2 Viewing

The viewing operation is slightly more complex to analyze. Even when we assume an infinitely thick plate, it appears more challenging to recover either the interference patterns or the original spectrum from an observed synthesized spectrum. The observed spectrum given by (4.10) is the squared magnitude of an analytic signal, and its phase information is irremediably lost in the process. The problem of recovering the phase from the magnitude of a signal is known as *phase retrieval* [83]. In general, unless we introduce prior information and further constrain the problem, it does not have a unique solution. The reason for this is that all functions with the same magnitude but arbitrary phase will lead to the same measured intensity. For instance, Figure 4.24 shows two very different interference patterns that generate the same spectrum. Constraints that guarantee unique reconstruction [22] of the phase include *sparse signals* [155],

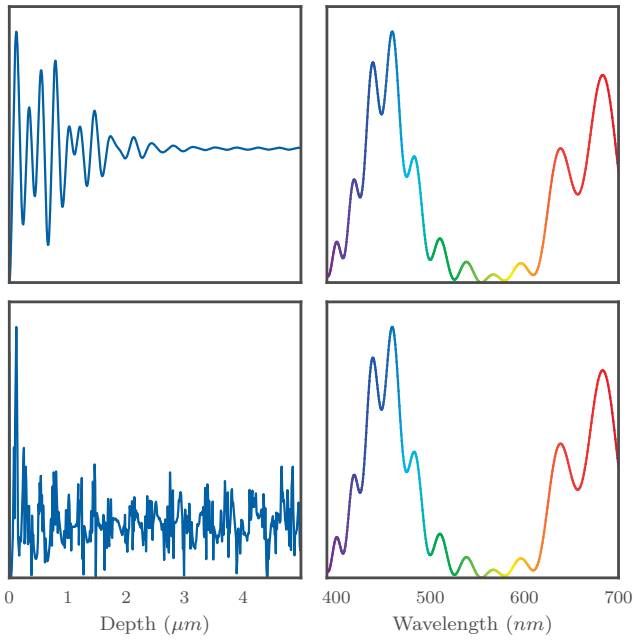


Figure 4.24: Two different interference patterns that generate the same spectrum. The second interference pattern also contains negative values, which is not practical in the Lippmann framework.

that is when the support of observed signals has few nonzero entries. Another prior information that is sometimes considered is the (partial) knowledge of the magnitude of the signal in the primal domain [21]. In Chapter 6, we further explore the phase retrieval problem: under the sparsity assumption, we propose an algorithm which, coupled with the FRI theory, can both recover the phase and super-resolve a signal.

In the Lippmann case however, we can take advantage of two crucial observations. First, $P(\omega)$ is a power spectrum; hence it is real, non-negative and zero for negative frequencies. This means that not all interference patterns are possible. For example, the second pattern in Figure 4.24 does not originate from a valid power spectrum. Second, even though the wavefunction $U_r(\omega, t)$ is complex, it is analytic and thus its real and imaginary parts are fully correlated. These additional pieces of information are very valuable and we see below how to leverage them to fully recover the original spectrum $P(\omega)$. The only restriction we impose is that we only recover a bandlimited version of $P(\omega)$; indeed, we have seen in Section 4.5.1 that the bandwidth of the synthesized spectrum is limited by the thickness of the plate Z. Assuming $P(\omega)$ is bandlmited, we can express it with a linear combination of sinc functions:

$$P^Z(\omega) = \sum_{n \in \mathbb{Z}} p_n \operatorname{sinc}\left(\frac{\pi(\omega - n\Omega_Z)}{\Omega_Z}\right), \tag{4.15}$$

where $\Omega_Z = \pi c/2Z$ is the sampling period controlled by the thickness Z and the entries p_n are lowpass samples of $P^Z(\omega)$.

Spectrum recovery from the complex wavefunction

Before we look at the algorithm for the general case, let us consider the simplified unrealistic setup where we have access to the phase, that is, we assume we measure $U_r(\omega, t)$ instead of $|U_r(\omega, t)|^2$. Furthermore, since the acquisition device only provides us a discretized version of the spectrum, we consider $u_m = U_r(m\Omega_s, 0)$, which is a sampled version of $U_r(\omega, t)$, where Ω_s is the sampling period and $m \in \mathbb{Z}$. We assume that $\Omega_s < \Omega_Z$, which means that we sample the spectrum at a bandwidth that is higher than that determined by the thickness of the plate. This is a realistic assumption since current hyperspectral cameras can now acquire a few hundred spectral bands in the visible spectrum. On the other hand, the thickness of Lippmann plates is usually around 5 μm and rarely above 10 μm; the latter corresponds to a sampling period of $\Omega_Z = \pi c/2Z = 3.14 \cdot 10^{13}$, or 46 samples in the visible spectrum, which is well below the spectral resolution of the best hyperspectral cameras.

The representation of $P(\omega)$ with a finite number of coefficients in (4.15) allows us to devise a matrix-based algorithm. More specifically, assuming $L(\omega')$ and ϵ are equal to 1, we express the sampled reflected wave function u_m using (4.15). From (4.6), we have

$$\begin{aligned} u_m &= U_r(\omega', 0)\big|_{\omega'=m\Omega_s} \\ &= \int_0^\infty P(\omega) H^Z(\omega, m\Omega_s) d\omega \\ &= \sum_{n\in\mathbb{Z}} p_n \int_0^\infty H^Z(\omega, m\Omega_s) \operatorname{sinc}\left(\frac{\pi(\omega - n\Omega_Z)}{\Omega_Z}\right) d\omega \\ &= \sum_{n\in\mathbb{Z}} p_n \Phi_{m,n}, \end{aligned} \tag{4.16}$$

where we define

$$\begin{aligned} \Phi_{m,n} &= \int_0^\infty H^Z(\omega, m\Omega_s) \operatorname{sinc}\left(\frac{\pi(\omega - n\Omega_Z)}{\Omega_Z}\right) d\omega \\ &= H^Z(n\Omega_Z, m\Omega_s). \end{aligned}$$

Since $H^Z(\omega, m\Omega_s)$ is already $\frac{2Z}{c}$-bandlimited, the effect of the convolution with the sinc is void. This result connects the sequence—or infinite vector—of power spectral data p_n with the sequence of measured reflected wave function u_m. In practice, we are only interested in a finite number of elements of these sequences (in general, the ones corresponding to visible light). We call $\mathbb{V}(\Omega)$ the set of indices that correspond to the range of frequencies of interest

$$\mathbb{V}(\Omega) = \left\{ n \mid n \in \mathbb{Z} \text{ and } n\Omega \in \left[\frac{2\pi c}{\lambda_{max}}, \frac{2\pi c}{\lambda_{min}}\right] \right\},$$

where λ_{min} and λ_{max} are the minimum and maximum wavelengths, respectively. For visible light, $\lambda_{min} = 390$ nm and $\lambda_{max} = 700$ nm. We define the vector $\boldsymbol{p} \in \mathbb{R}^N$, whose entries are given by p_n, where $n \in \mathbb{V}(\Omega_Z)$ and N is the size of $\mathbb{V}(\Omega_Z)$. Analogously, we define $\boldsymbol{u} \in \mathbb{C}^M$, whose entries are u_m, where $m \in \mathbb{V}(\Omega_s)$ and M is the size of $\mathbb{V}(\Omega_s)$. This notation allows us to reformulate (4.16) as a matrix-vector operation:

$$\boldsymbol{u} = \boldsymbol{\Phi}\boldsymbol{p}, \tag{4.17}$$

where $\boldsymbol{\Phi} \in \mathbb{C}^{M\in N}$ and $\Phi_{m,n} = H^Z(n\Omega_Z, m\Omega_s)$ for $m \in \mathbb{V}(\Omega_s)$ and $n \in \mathbb{V}(\Omega_Z)$. Note that the truncation of (4.16) into (4.17) leads to an exact result only if we assume that $p_n = 0$ for n

outside of $\mathbb{V}(\Omega_Z)$. This assumption is reasonable, since glass does not let ultra-violet light go through and most modern cameras contain an infrared cut-off filter. Observe that $\boldsymbol{\Phi}$ is a tall, full-rank matrix; thus, we can invert (4.17) by computing

$$\boldsymbol{p} = \boldsymbol{\Phi}^{\dagger}\boldsymbol{u}, \tag{4.18}$$

where $\boldsymbol{\Phi}^{\dagger}$ is the Moore-Penrose pseudoinverse. Since Φ is full-rank, $\boldsymbol{\Phi}^{\dagger} = (\boldsymbol{\Phi}^{*}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{*}$ is a left inverse and $\boldsymbol{\Phi}^{\dagger}\boldsymbol{\Phi} = \boldsymbol{I}$. In other words, we can fully recover the power spectrum from the measured wave function. This result is summarized in the following lemma.

Lemma 4.1 (Spectrum recovery from complex wavefunction) Let $P(\omega)$ be a power spectrum acquired by a Lippmann plate and $U_r(\omega, t)$ the resynthesized complex wavefunction given in (4.6) assuming a flat illumination $L(\omega) = 1$. Furthermore, let $\boldsymbol{u} \in \mathbb{C}^M$ be a sampled version of $U_r(\omega, t)$ such that $u_m = U_r(m\Omega_s, 0)$ for $m \in \mathbb{V}(\Omega_s)$ and $\Omega_s < \Omega_Z$.

If $P(\omega)$ is bandlimited such that (4.15) holds and such that its coefficients p_n are zero for $n \notin \mathbb{V}(\Omega_Z)$, then we can exactly recover it by computing

$$\widehat{\boldsymbol{p}} = \boldsymbol{\Phi}^{\dagger}\boldsymbol{u},$$

and further interpolate it using (4.15). If $P(\omega)$ is not bandlimited, then $\widehat{\boldsymbol{p}}$ gives the closest bandlimited approximation of $P(\omega)$.

Practical consideration Lemma 4.1 assumes that p_n is zero for indices n that are outside the range of visible light; this may not be the case in practice. Since the values u_n depends on p_n for all $n \in \mathbb{Z}$ in (4.16), some artifacts might appear in $\widehat{\boldsymbol{p}}$. To attenuate these artifacts, we assume that the boundary values of p_n are repeated outside of the visible range.

Improving noise resilience When the measurements $\boldsymbol{u}$ are noisy, it can happen that the estimated power spectrum $\widehat{\boldsymbol{p}}$ has negative and/or complex entries, which violates the definition of a power spectrum. We enforce the positivity of $\widehat{\boldsymbol{p}}$ by replacing (4.18) with

$$\widehat{\boldsymbol{p}} = \mathrm{NNLS}(\boldsymbol{\Phi}, \boldsymbol{u}), \tag{4.19}$$

where $\mathrm{NNLS}(\boldsymbol{\Phi}, \boldsymbol{u})$ solves a nonnegative least-square problem [99]. Similarly, we make $\widehat{\boldsymbol{p}}$ real by solving (4.19) over $\mathbb{R}$; this is done by vertically stacking the real and imaginary parts of $\boldsymbol{\Phi}$ and $\boldsymbol{u}$ as follows:

$$\widehat{\boldsymbol{p}} = \mathrm{NNLS}\left(\begin{bmatrix}\Re\mathfrak{e}\{\boldsymbol{\Phi}\}\\ \Im\mathfrak{m}\{\boldsymbol{\Phi}\}\end{bmatrix}, \begin{bmatrix}\Re\mathfrak{e}\{\boldsymbol{u}\}\\ \Im\mathfrak{m}\{\boldsymbol{u}\}\end{bmatrix}\right). \tag{4.20}$$

Phase retrieval

Next, we estimate the reflected complex wavefunction $\boldsymbol{u}$ from its measured spectrum $\boldsymbol{v} = |\boldsymbol{u}|^2$. To that end, we propose an iterative algorithm inspired by [60] that successively enforces the following two properties:

Algorithm 4.1 Spectrum recovery

Input: The measured intensity $\boldsymbol{v}$
Output: The spectrum $\widehat{\boldsymbol{p}}$ such that $|\boldsymbol{\Phi}\widehat{\boldsymbol{p}}|^2 = \boldsymbol{v}$
$\widehat{\boldsymbol{u}}_0 = \mathbb{1}$
$k = 0$
Estimate the thickness of the plate Z and the development parameters ϵ and τ
Construct $\boldsymbol{\Phi}$ from Z and τ
repeat
  $\widehat{\boldsymbol{u}}_{k+1} = \widehat{\boldsymbol{u}}_k \frac{\sqrt{\boldsymbol{v}}}{|\widehat{\boldsymbol{u}}_k|}$
  Compute $\widehat{\boldsymbol{p}}_{k+1}$ by solving (4.20) with $\boldsymbol{\Phi}$ and $\widehat{\boldsymbol{u}}_{k+1}$
  $\widehat{\boldsymbol{u}}_{k+1} = \boldsymbol{\Phi}\widehat{\boldsymbol{p}}_{k+1}$
  Estimate the thickness of the plate Z and the development parameters ϵ and τ
  Update $\boldsymbol{\Phi}$ with the new values of Z and τ
  $k = k + 1$
until convergence
return $\widehat{\boldsymbol{p}}_k$

1. **A consistent spectrum**. A spectrum $\widehat{P}(\omega)$ is *consistent* with respect to the measurements when the magnitude of the corresponding reproduced spectrum is equal to the magnitude of the measured spectrum. More formally, we want $|\boldsymbol{u}| = \sqrt{\boldsymbol{v}}$. To enforce this, we correct our current estimate $\widehat{\boldsymbol{u}}_k$ as follows:

$$\widehat{\boldsymbol{u}}_{k+1} = \widehat{\boldsymbol{u}}_k \frac{\sqrt{\boldsymbol{v}}}{|\widehat{\boldsymbol{u}}_k|}.$$

2. **A valid bandlimited power spectrum**. We call a wavefunction $\widehat{\boldsymbol{u}}_k$ *valid* when it originates from a power spectrum $\widehat{\boldsymbol{p}} \in \mathbb{R}^N$ through (4.17). Mathematically speaking, we would like $\widehat{\boldsymbol{u}}_k$ to be in the range of $\boldsymbol{\Phi}$. We can further ensure that $\widehat{\boldsymbol{u}}_k$ is issued from a real and positive power spectrum by computing $\widehat{\boldsymbol{p}}_k$ from $\widehat{\boldsymbol{u}}_k$ using (4.20). Then, to enforce the validity of $\widehat{\boldsymbol{u}}_k$, we perform the forward transform on $\widehat{\boldsymbol{p}}_k$:

$$\widehat{\boldsymbol{u}}_{k+1} = \boldsymbol{\Phi}\widehat{\boldsymbol{p}}_k.$$

   This ensures that $\widehat{\boldsymbol{u}}_{k+1}$ results from a bandlimited, real and positive power spectrum.

We conjecture that this recovery is unique, under the condition that $P(\omega)$ is a valid and bandlimited power spectrum. Even though we do not provide a formal proof here, extensive simulations have shown that our algorithm always converges to the correct spectrum.

Estimation of thickness and development parameters

In addition to the shape of the original spectrum, there are a few extra parameters that we need to estimate to achieve a faithful reconstruction, namely the thickness of the plate Z, the scaling factor ϵ and the development decay τ. The oscillations present in synthesized spectra provide a very valuable clue regarding the value of the thickness of the plate: from the Fourier transform of $\boldsymbol{v}$, we can identify a peak that corresponds to the frequency of the oscillations created by $Q^Z(\omega)$. Furthermore, we can refine the estimation using our forward model (4.10) and by performing a

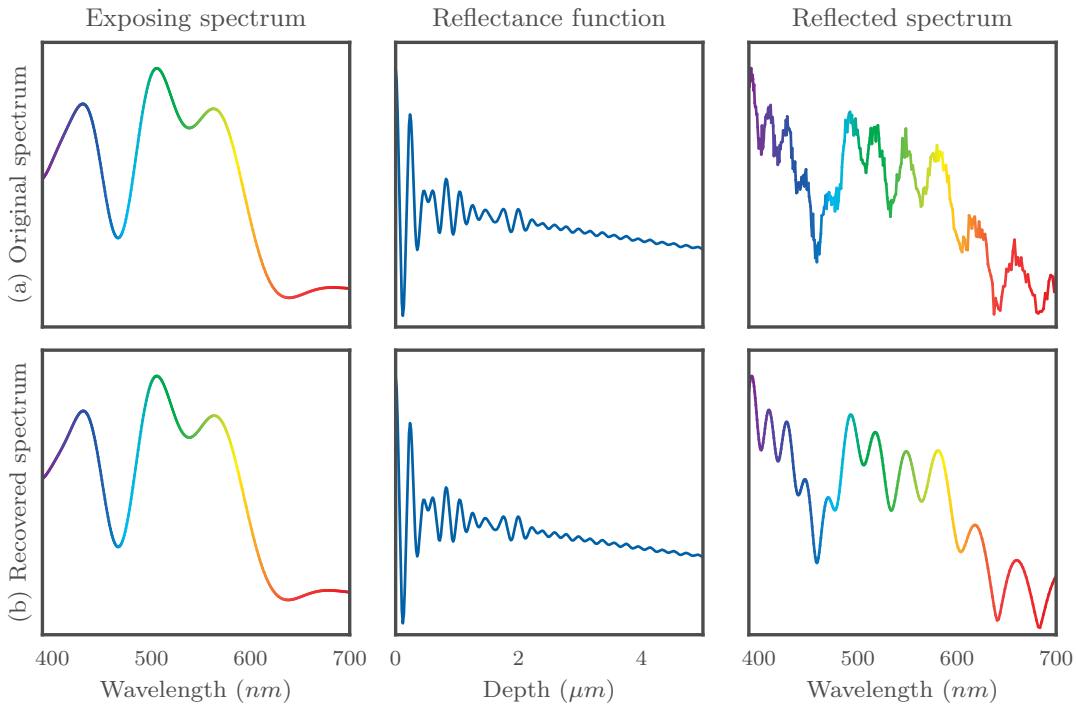


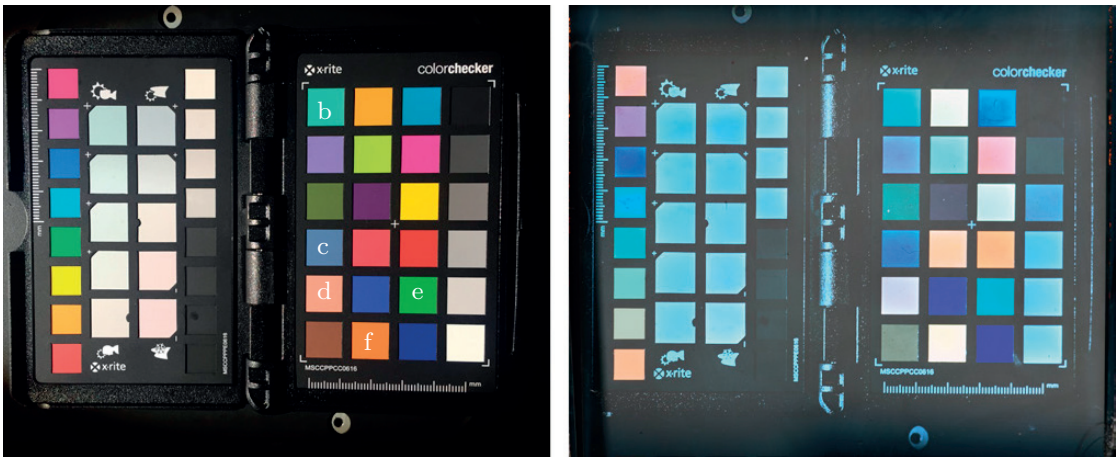
Figure 4.25: Spectrum recovery algorithm: the top row shows the original spectrum, the corresponding reflectance function with some decay due to the development and a reflected spectrum corrupted with Gaussian noise; the bottom row shows the estimated spectrum using Algorithm 4.1 and the corresponding reflectance function and reflected spectrum.

line search on the parameter Z and identifying the value that minimizes the discrepancy between our model and the measurements. We estimate the development parameters τ and ϵ in a similar fashion. Doing so at each round of the algorithm also strengthens the forward model by adapting the filter parameters.

With this, we have all the ingredients to devise an algorithm to estimate the original spectrum. By iteratively enforcing the aforementioned properties, our results show that the procedure converges toward $P_r^Z(\omega')$. This is summarized in Algorithm 4.1 and demonstrated on synthetic data in Figure 4.25. What is interesting to note is that the synthesized version of the spectrum can be fairly different from the original spectrum, contradicting the common belief that the reflected spectrum is equal to the original. Even when this is the case, our algorithm manages to faithfully estimate the underlying original spectrum.

Experiments

To test our reconstruction algorithm and study the color reproduction of Lippmann plates on arbitrary spectra, we created a Lippmann photograph of an X-Rite ColorChecker using air as reflector. Figure 4.26a shows the reference image and Figure 4.26b the colors reflected from the developed plate. The color reproduction is on par with the experiment from Figure 4.19 in the



(a) Original scene (b) Lippmann plate

Figure 4.26: Color reproduction of a Lippmann plate: an X-Rite ColorChecker (a) has been acquired under a xenon light source. The resulting Lippmann plate is shown in (b); aesthetically, we could correct for the blue cast by soaking the plate in a sorbitol solution (see Appendix 4.A.3 for more details).

sense that we observe a shift of the colors towards the blue colors. The red tones are also fainter and harder to properly reproduce.

In addition, we acquired hyperspectral measurements of the original scene and of the image reflected from the plate. We tested our spectrum recovery algorithm using the spectral data of the replayed scene and report the results for a few colors in Figure 4.27. We notice the following three observations. First, our algorithm is fairly efficient at recovering the global shape of the original spectra and the original colors. It is also effective at eliminating the skewing effect, see for instance Figure 4.27a. On the other hand, it only retrieves a lowpass version of the spectra, which is a direct consequence of the finite thickness of the plates. Second, the red colors are almost nonexistent: this is probably due to the choice of dyes in the making of the plates, a parameter that is not under our control. As a consequence, the estimated spectra also lack the red component; this is particularly visible in Figures 4.27d and 4.27f. Third, we observe that sometimes the amplitude of the oscillations decays rather quickly between the red and the blue colors and we somewhat fail to capture this effect. We postulate that this is due to the relative simplicity of our development model. We believe that using a more elaborate model could improve the estimation, perhaps at the expense of not being able to derive an analytic solution for $S_\tau^Z(\omega)$. To summarize, our algorithm offers a noticeable improvement from the reflected spectra and manages to retrieve the global shape of the original spectra relatively well.

4.8 Conclusion

It is legitimate to question the purpose of analyzing a 100 year-old forgotten photography technique. First, the contributions of Gabriel Lippmann [103] are the cornerstone of spectroscopy and holography and they deserve close attention. From that perspective, reading Lippmann's

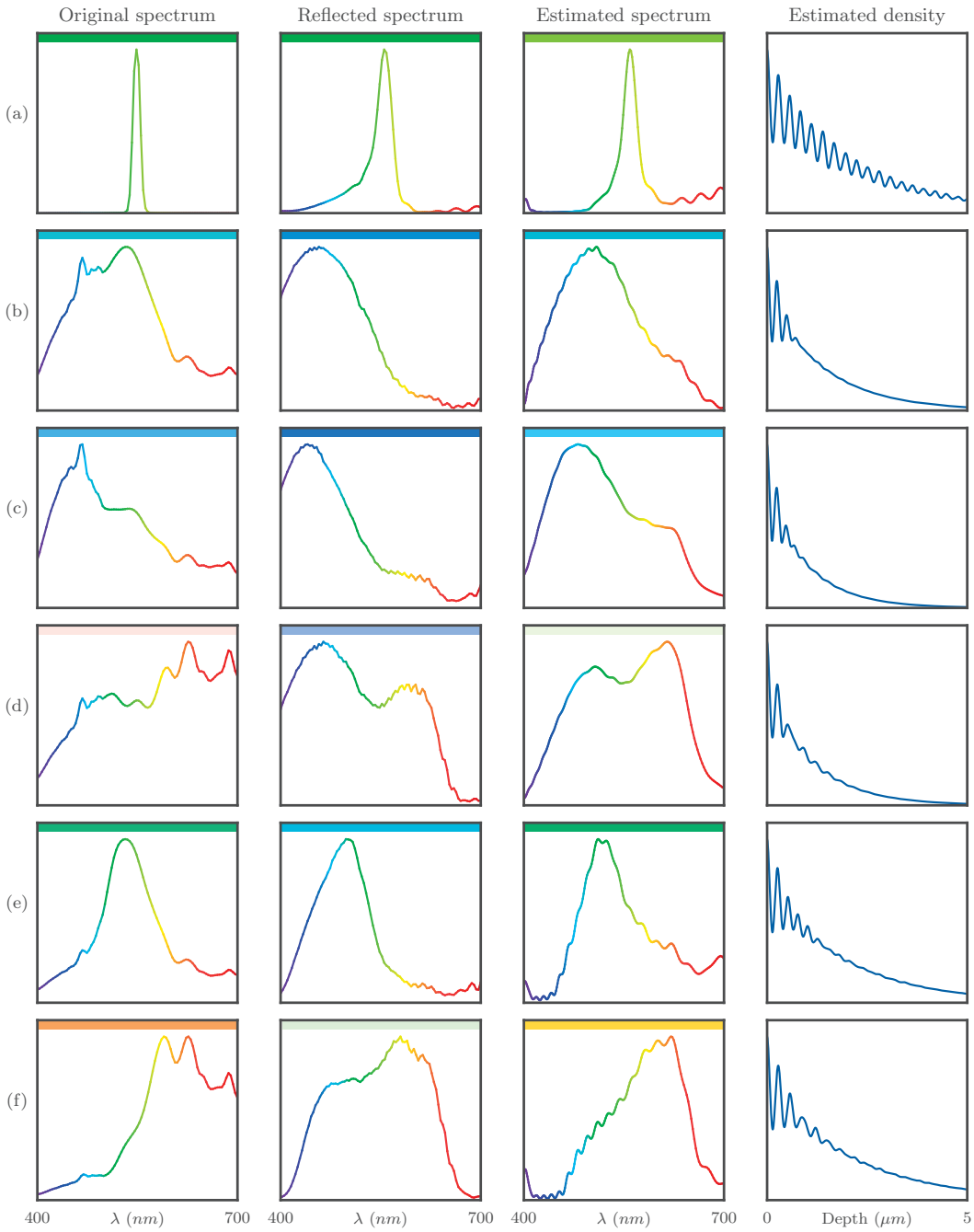


Figure 4.27: Spectrum recovery on real measured data: (a) 550 nm bandpass filter from Figure 4.20, (b)-(f) selected squares from the ColorChecker from Figure 4.26a (indicated with white letters). The first column shows the original spectrum; the second column shows the measured spectrum reflected from the plate; the third column shows our estimation using Algorithm 4.1; and the last column shows the estimated density derived from the estimated spectrum.

manuscripts can be compared to studying Leibnitz publications on differential calculus. Second, a thorough inspection of Lippmann's Nobel prize winning paper [103] enabled us to elucidate a misconception that went unnoticed for more than a century regarding the perfect reconstruction of the synthesized spectrum with infinite plates. Furthermore, our formulation provided a clear explanation of the color reproduction of Lippmann plates. This allowed us to characterize and describe various effects such as the role of lowpass filter played by the thickness of the plate, the skewing effect induced by the choice of reflector, and the presence of oscillations in the synthesized spectra. Third, we proposed an algorithm to recover original spectra from their distorted synthesized versions. This has the potential to enable retrieval of true hyperspectral data from Lippmann plates from the early 1900s and, effectively opening hyperspectral windows to look into the past. Last but not least, technology has evolved and we can develop Lippmann's ideas using modern technology. We will elaborate more on this topic in the final conclusion.

4.A Practical aspects

We summarize here the more practical aspects of Lippmann photography, which remains a challenging technique that requires care and precision to be successful. In hindsight, it is remarkable that the physicists and chemists of the late 19th and early 20th century managed to put all the pieces together to make the technique work. In order to verify our model and findings, we created our own Lippmann photographs and experiments. To achieve this feat, we were lucky to be supported by people from fields as varied as physics, chemistry and material science. We also had the chance to be assisted by Filipe Alves[13], who is one of the few contemporary Lippmann photographers. We provide in this section a brief overview of the chemical processes involved. We also review the different techniques for making and developing Lippmann plates, both from a historical perspective and with a more recent point of view. We stress that photochemists have extensively studied this field in both academia and industry for more than a century. Therefore, our aim is not to provide an exhaustive reference but rather present a high-level overview for people who do not have a strong background in analog photography or holography.

4.A.1 Lippmann plates

The creation of Lippmann plates is very similar to classical analog photography. We review here the main concepts, while stressing the differences between Lippmann and traditional photography plates. For more exhaustive explanations, we refer the reader to [14, 30, 32, 70].

One of the main ingredients to make a photographic plate are silver salts, which have the interesting property that they transform into metallic silver when in contact with light. Of all the silver salts, the family of so-called *sliver halides* was found to be particularly suitable for photography. In the context of Lippmann photography, the key difference between metallic silver and silver salts is that the former reflects much more light than the latter.

Emulsion Now that we have photosensitive materials in silver halides, the next step is to isolate them in tiny grains, which are distributed uniformly at random throughout a medium. This is precisely the role of a *photographic emulsion* (see Figure 4.29a). Several components have been tested to find the perfect emulsion for photography, and some of the most appropriate are gelatin, albumen, polyvinyl acetate [182], or collodion. The size of the silver halide grains has a direct influence on the sensitivity of the photographic emulsion, where larger grains are synonyms of higher sensitivity and shorter exposure times. Translated to the digital camera jargon, the grain size could be broadly equated to signal gain applied to the sensor. In Lippmann photography, it is important to have extremely fine grains—of the order of 10 nm—to be able to capture the small variations of standing waves. An advantage of the use of extremely fine grains is that the resolution of the recording material is phenomenal: in fact, the resolution of Lippmann photographs is only limited by the resolving power of the optics of the camera, but not by the emulsion.

Dyes Once the emulsion has been prepared, it is usually coated on a glass plate. At this point, we have a photographic plate that is sensitive to light. Unfortunately, silver salts react only to blue and violet light, and are almost unresponsive to red light [87]. In order to make the emulsion panchromatic and extend the sensitivity to all visible colors, sensitizing dyes are put in contact

[13] http://www.lippmannphotography.com

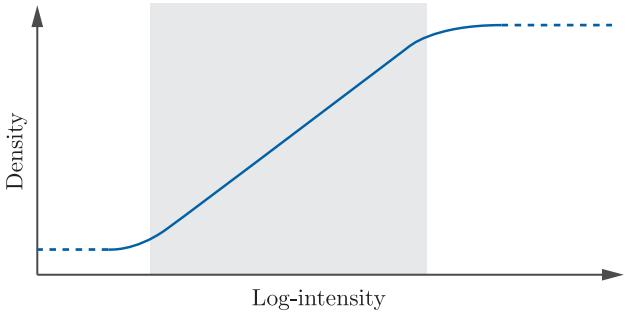


Figure 4.28: A Hurter and Driffield curve.

with the silver halide grains. In addition to dyes, other sensitizers are often added to boost the sensitivity of the silver-halide crystals. The choice of dyes is crucial to determine the spectral response of the plate, the goal being to make it as isochromatic as possible. A large body of work has been historically dedicated to finding the best combination of dyes [87], and the most convincing results in terms of color reproduction have been arguably obtained by Lehmann, who blended together pinacyanol, orthochrome T and acridine orange [85].

4.A.2 Hurter and Driffield curve

In Section 4.2.2, we assumed that the density of metallic silver was linear with respect to the intensity of exposing light. This is obviously not entirely true: there should be some threshold of intensity such that the density saturates and some threshold such that the silver density is at its minimum. By plotting the density of silver against the intensity of light, we obtain the *Hurter and Driffield* [79], or *characteristic curve*, an example of which is shown in Figure 4.28.

In essence, the Hurter and Driffield curve characterizes how the photographic films packs the dynamic range of a scene into the variation of silver densities. One of the effects of the characteristic curve is the clipping of the highlights and shadows of a scene. This is the reason why it is important to properly meter and expose a scene, since an imprecise exposure heavily compresses the shadows or the highlights. A properly exposed image should capture intensities that span the entire gray area in Figure 4.28.

4.A.3 Development and processing

Another essential part of interferential photography is the processing stage, which encompasses development, fixing, bleaching, and drying. Regarding development, the basic concept is identical to classical photography, where the developing bath accelerates the reduction of silver halide and its transformation into metallic silver (see Figure 4.29c). The silver specks that constitute the latent image serve as catalyst for this reaction. A developer contains the following three components: a *developing agent*, an *accelerator*, and a *restrainer*. The role of the developing agent is to convert the silver halide to metallic silver. The restrainer has the opposite effect and helps containing the chemical reaction. More precisely, it prevents unexposed areas from being developed. As its name suggests, the function of the accelerator is to speed up the development procedure.

	Lumière [105]	**Cajal** [202]	**Filipe Alves**
Pyrogallol	1 g/l	0.52 g/l	0.83 g/l
KBr	15 g/l	7 g/l	8.3 g/l
Ammonia	4.5 g/l	1.2-1.8 g/l	1.25 g/l
Temperature	15° C	?	?
Time	1-3 min	?	$\sim$ 1 min

Table 4.1: Different development processes based on the Lumière developer.

A notable difference is that, in classical photography, the development usually creates a *negative*, in which the more exposed silver specks appear darker than the less exposed areas. This negative image needs to be inverted: the inversion process can be done for instance by washing away all the metallic silver and later exposing and redeveloping the remaining silver halides. This step is not needed in Lippmann photography, as a high density of metallic silver is synonym of a larger amount of light reflected. Interestingly, we cannot categorize Lippmann photography as a positive or negative photography technique, since the silver grain density through the depth corresponds to a Fourier-based representation of the image.

At the end of the 19th century, significant efforts have been made to devise developers suitable for Lippmann photography. The brothers Auguste and Louis Lumière led the movement and conceived the *Lumière developer*, which consists of pyrogallol for development, potassium bromide (KBr) as restrainer and ammonia as accelerator [105]. The proportion of each components was found by trial and error and can substantially depend on external parameters such as the temperature of the bath and, mainly, the hardening of the gelatin. Inspired by their recipe, other scientists and artists devised their own variations of the Lumière developer: Table 4.1 summarizes a few development solutions found in the literature or shared by Filipe Alves based on the Lumière recipe. The Lumière developer is known to be more active on the surface and does not penetrate deeply into the plate. As a consequence, it smoothes out the higher frequencies of the Fourier transform of the recorded spectra and damps the oscillations. This is however not necessarily an issue as the most important spectral information is usually located in the first few fringes.

An alternative to the Lumière developer is the so-called GP2 solution, which contains hydroquinone [31]. Ives [81] observed that hydroquinone-based developers generate brighter colors and better defined spectra thanks to a better diffusion inside the depth of the plate. Holographic developers can also be used for the development of Lippmann plates, since holographic plates are strictly identical to Lippmann plates, besides the choice of dyes.

Fixing After development, a plate still contains silver halide crystals that are sensitive to light. If it is exposed to light and redeveloped, the original image will be lost. Even without development, a prolonged exposure to light slowly creates a latent image that can irremediably wash out the original image. To eliminate the residual silver halide as well as the dyes and halide crystals resulting from the development, it is necessary to apply a *fixing* bath (see Figure 4.29d). Historically, fixing baths were made of potassium cyanide [30] but nowadays, common fixing solutions are made with sodium thiosulfate.

A side effect of the fixing process is that it tends to shrink the plates, which results in shifting

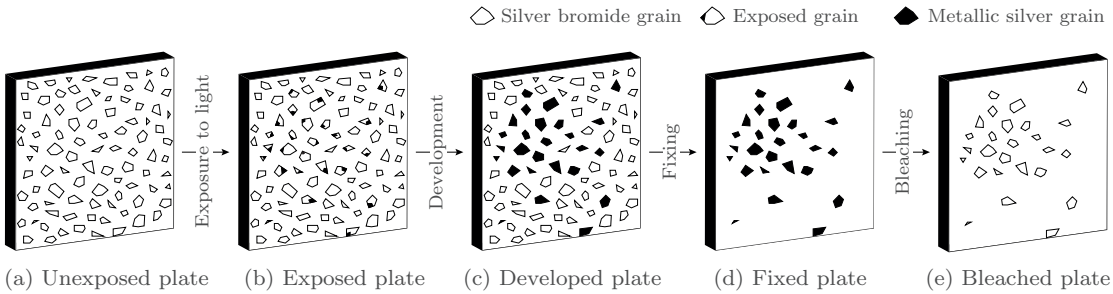


Figure 4.29: Chemistry of photography: (a) the unexposed plate emulsion contains silver halide crystals; (b) during exposure to light, the crystals that are in contact with light start transforming into metallic silver; (c) development accelerates this transformation and fully transforms the crystals into metallic silver; (d) fixing washes out the remaining unexposed silver halides; (e) re-halogenating bleaching.

the spectra toward the blue colors. For that reason, many Lippmann photographers suggests to skip this step. If fixing is applied, the shrinking can be counterbalanced with additional baths. For example, in [30], a *pre-hardening* step (before the development) is suggested to make the gelatin harder and more resilient to shrinkage. Another option is to apply an expanding bath such as an aqueous solution of sorbitol after fixing to compensate for the shrinkage.

Bleaching Another optional step is *bleaching* [32], whose goal is to obtain a transparent plate that contains only silver halides and no more metallic silver. There are two main variations of the bleaching process. The first one is called *rehalogenating* bleach, which basically converts the grains of metallic silver back into silver halide crystals. This bleaching bath is typically applied after fixing such that these newly formed silver halides are not confounded with the grains that have not been exposed to light (see also Figure 4.29e); but in certain holographic development procedures, it can also be applied without fixing, in which case it is called *fixation-free* [32, 173].

The other option is to apply a *reversal* bleach, which makes the metallic silver soluble and washable. What remains after a reversal bleach is the unexposed silver halides, so the final image is essentially the negative of the image obtained with a rehalogenating bath.

Drying

The drying of the plates is also a crucial step as it can potentially modify the silver density patterns if not done properly. More precisely, drying can lead to chirping, in which the silver density is expanded or shrinked at different depths. This can lead to artifacts in the reproduced spectrum; nevertheless, we ignored this effect in this manuscript.

4.A.4 Thickness of plate and choice of development technique

Very often in the literature (see for example [102, 103]), scientists differentiate the monochromatic from the polychromatic case. As we have seen in this chapter, these two cases can be unified under one single theory; however, from a practical point of view, it can be useful to differentiate them. Polychromatic light typically has a short coherence length, which means that the relevant information is confined within the first couple of micrometers from the surface of the plate (see

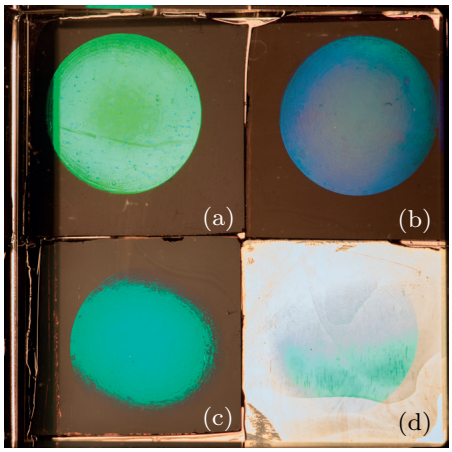


Figure 4.30: The effect of different developers: (a) holographic developer (Ultimate developer[14]) followed by bleaching; (b) holographic developer without bleaching; (c)-(d) Lumière-based developer. The last plate presents an example of *fogging*, which occurs when the surface overdevelops and becomes almost mirror-like; this typically happens when the proportion ammonia/KBr is too high or when the development time is too long.

for instance Figure 4.5). After the coherence length, only the baseline remains, which is not adding any meaningful information to the replayed spectrum.

On the other hand, monochromatic light theoretically has infinite coherence. In this regime, the deeper the plate the finer the spectral resolution of the reproduced spectrum. Hence better results are obtained with thicker plates. This difference also influences the development process. For polychromatic scenes, practitioners tend to favor shallower plates and development techniques, whereas for the monochromatic case, deeper development and plates are preferred. The bleaching step is also often added to the development of monochromatic plates, since silver halides enable more light to pass through than metallic silver, allowing the deeper layers to be reached.

4.A.5 Modern Lippmann photography

Nowadays, no more research is dedicated to the creation of Lippmann plates and their development solutions. However, scientists are still active on the front of holography, which bears several similarities with the Lippmann technique. Holographic plates constitute the perfect medium for recording Lippmann photographs. The only problem is that such plates are usually only made sensitive to one or two (three for color holography) spectral bands. Besides the difference of dyes, holographic plates are strictly identical to Lippmann plates. There are a few companies that produce holographic plates; for our experiments, we chose to order plates made by Yves Gentet [67] from Ultimate Holography[14]. Thanks to holography, appropriate developers have also been elaborated. In our study, we also tried the holographic developer made by Yves Gentet whose composition is not disclosed.

From a more practical point of view, there are also a few contemporary artists currently making Lippmann photographs. To our knowledge, Filipe Alves from Portugal and Darran

[14] http://www.ultimate-holography.com/

Green [31] from England studied the art of Lippmann photography and make their own emulsions to produce beautiful artworks.

Brief comparison of developers The development is one of the most crucial steps in Lippmann photography as it can dramatically change the final result. To illustrate the difficulty, we created four plates with the same exposure—4 seconds under a green laser light having a wavelength of 531 nm—and we only changed the development formula; these plates can be seen in Figure 4.30. We observe that the reproduced colors are severely influenced by the developer. In fact, it took us several attempts to identify the ideal developer for our needs. After many trial-and-error experiments with three different developers (Ultimate[14], Lumière developer and GP2), we obtained our most conclusive results with a variation of the Lumière formula.

Chapter 5

Electromagnetic Theory: Micro-Polarizer Arrays*

From a long view of the history of mankind—seen from, say, ten thousand years from now—there can be little doubt that the most significant event of the 19th century will be judged as Maxwell's discovery of the laws of electrodynamics.

The Feynman Lectures on Physics
Richard Feynman

5.1 Introduction

In Chapter 3, we described algorithms to represent specularities from multiple images. To isolate the specular components, we either ignored the diffuse term or removed an estimation of its shape by fitting a cosine function to it. Unfortunately, this strategy lacks stability when the measured SVBRDF contains more complex shapes such as for example occlusions. In this chapter, we propose to separate the diffuse and specular term using dedicated hardware, namely cameras with polarizing micro-filter arrays.

*The material in this chapter is the result of joint work with Laurent Jospin and Adam Scholefield. Author contributions: GB and AS designed research; LJ, GB and AS performed research; LJ carried out experiments and simulations; and GB wrote the chapter based on [88], which was written by GB, AS and LJ.

The separation of the diffuse and specular components in a photo is useful in many different imaging applications. For example, photometric stereo [201] estimates surface normals from images taken under different illuminations; however, like many algorithms, the process assumes that the scene is Lambertian, i.e. the reflection is purely diffuse. As a second example, consider the problem of estimating spatially-varying bidirectional reflectance distribution functions (SVBRDFs). Typically, one fits a parametric model to samples obtained by capturing images with different viewing and lighting directions. Although for very simple materials the specular and diffuse components of the model can be fit from only the shape of the SVBRDF, often this process can be done much more accurately if the two components are separated prior to model fitting. Other applications that benefit from this separation include tracking, object recognition, image segmentation and the problem of removal of glare.

5.1.1 Related work

There are a number of approaches to perform the separation of the diffuse and specular components, using either single or multiple images [5]. Most single-image approaches are color-based: this process was pioneered by Shafer [164], with the so-called *dichromatic reflection model*. A key observation is that, for dielectrics, the spectrum of the specularity is similar to the spectrum of the light source, whereas the diffuse spectrum also encompasses information about the color of the surface: this results in a T-shaped space [92] whose limbs represent the diffuse and specular components in the dichromatic reflection model. Approaches based on this model [69, 93] assume that the scene has been segmented into different parts with a uniform diffuse distribution, which make them unpractical for highly textured scenes. More recent work alleviates the need for segmentation by integrating spatio-temporal information [110]. Similar approaches with different colorspaces such as HSI [203], rotated RGB [111], or custom spaces [15, 170, 180] have also been proposed. As well as color-based techniques, the so-called *neighborhood analysis* methods [109, 179, 204] leverage the information from neighboring pixels to infer the reflective component at a given point.

Multi-image approaches use strategies as varied as structured illumination [98, 108, 130], different viewing angles [86, 101], or light-field imaging [117, 193]. Another multiple image approach makes use of polarizing filters. With a single filter in front of the lens, Nayar et al. [127] recast the separation problem as a linear system using multiple captures under different polarizations. Building on their work, a number of methods propose to combine color information with polarizing filters to separate the reflective components [90, 108, 128, 185].

Ma et al. [108] and Debevec et al. [49] suggest the use of filters *both* in front of the camera and the light sources, which leads to a slightly different model for the recorded intensity.

Recently, cameras have been manufactured with an array of polarizing micro-filters placed just in front of the image sensor [1, 61, 149, 156]. Analogously to a Bayer filter, the micro-filter of each pixel is oriented in one of a finite number of possible directions (see Figure 5.1). Polarization is of interest here because the diffuse and specular reflections affect polarization differently. Specular reflections are caused by direct reflection of the incoming light at an object's surface. Therefore, this type of reflection preserves the polarization of the incoming light. Diffusion is slightly more complex. The diffusion created by the surface reflections can be thought of as the combination of several specular reflections in all orientations due to the rough nature of most materials at the microscopic level. The polarization of the light reflected by this model is not preserved. On the other hand, some of the light is also scattered inside the medium and the polarization of such light is retained.

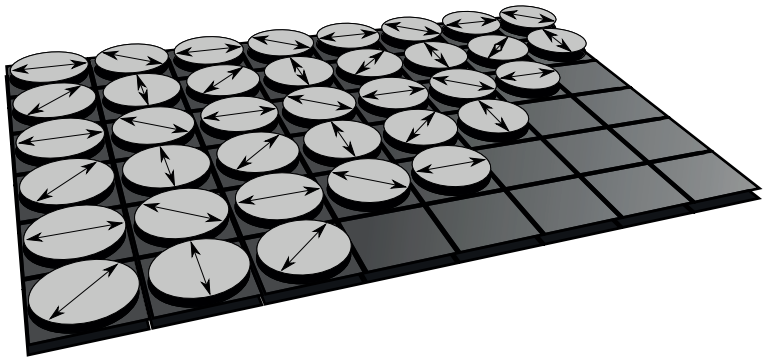


Figure 5.1: Illustration of the concept of a filter array with different polarization orientations. The micro-filters are represented with disks and the arrow indicates the direction of polarization.

We can leverage this observation to separate diffuse and specular reflection. In fact, we separate polarized and unpolarized light, which we can associate to the specular and diffuse components as long as the subsurface scattering is relatively weak. Potentially, micropolarizer cameras can enable the separation of the diffuse and specular components from a single image with the accuracy and robustness of multi-image polarization methods.

5.1.2 Proposed approach

In this chapter, we investigate the feasibility of using these types of cameras for the separation of specularities by proposing an algorithm to demosaic images generated with polarizing micro-filter arrays. Unfortunately, we were unable to obtain a real micro-polarizing camera and we thus simulate such a setup from our own rendered images, real world data from [170] and our own acquisitions. While this is not ideal, it does give us some additional advantages. In particular, the aforementioned camera designs are based on regular patterns with four different orientations; however, we also study random arrangements with varying numbers of orientations and show that these arrangements offer advantages over regular patterns.

In addition to demosaicing, we demonstrate the usefulness of micro-filter arrays for photometric stereo. By placing polarizers with different orientations in front of the light sources of the dome described in Section 3.2.3, we show that the separation of the diffuse and specular components can significantly improve the estimation of the surface normals of objects. This approach is essentially the dual of [46], where a camera with micro-polarizers is combined with a multi-view approach to infer the shape of objects.

5.2 Problem statement

Suppose we have a scene illuminated by a single light source polarized with an angle ϕ; this can be achieved by placing a linear polarizer in front of it. For now, assume that ϕ is known; later, we show that this is unnecessary as it can be easily estimated. Furthermore, suppose we have a digital camera with a filter array of linear polarizers with K different orientations in front of its sensor. We describe the image formation process as follows. Let $\boldsymbol{X}_k \in \mathbb{R}^{N \times N}$, $k = 1, 2, ..., K$, be a collection of K images, each filtered with a distinct orientation θ_k. From Malus' law (1.4),

we can express each individual image as the sum of a constant diffuse term and a specular term that is modulated according to the polarization of the filters:

$$\boldsymbol{X}_k = \frac{1}{2}\boldsymbol{Z}_d + \boldsymbol{Z}_s \cos^2(\phi - \theta_k), \tag{5.1}$$

where $\boldsymbol{Z}_d$ and $\boldsymbol{Z}_s$ are the diffuse and specular components, respectively. Equivalently, we can rewrite (5.1) by flattening $\boldsymbol{X}_k$ into a vector $\boldsymbol{x}_k = \mathrm{vec}(\boldsymbol{X}_k)$[1]:

$$\boldsymbol{x}_k = \underbrace{\begin{bmatrix}\frac{1}{2}\boldsymbol{I} & \cos^2(\phi - \theta_k)\boldsymbol{I}\end{bmatrix}}_{\boldsymbol{C}_k} \begin{bmatrix}\boldsymbol{z}_d \\ \boldsymbol{z}_s\end{bmatrix}.$$

Here, $\boldsymbol{z}_d = \mathrm{vec}(\boldsymbol{Z}_d) \in \mathbb{R}^{N^2}$, $\boldsymbol{z}_s = \mathrm{vec}(\boldsymbol{Z}_s) \in \mathbb{R}^{N^2}$ and $\boldsymbol{C}_k \in \mathbb{R}^{N^2 \times 2N^2}$ is the matrix that modulates the specularity with the correct attenuation and combines the diffuse and specular terms.

To model a camera with a micro-array of polarizing filters, we need to select only one polarization orientation for each pixel. To model this, let $\boldsymbol{y}_k = \boldsymbol{A}_k \boldsymbol{x}_k$, where $\boldsymbol{A}_k \in \mathbb{R}^{N^2 \times N^2}$ is a mask that zeroes the measurements that we do not have access to; it has the form of a diagonal matrix with a 1 where the pixel is selected and a 0 otherwise. Note that, since we have one polarizing filter per pixel, $\sum_{k=1}^{K} \boldsymbol{A}_k = \boldsymbol{I}$. At this stage, we make no assumption about the filter orientations: they can be regularly or irregularly structured over the array. Finally, the measured image is given by

$$\boldsymbol{y} = \sum_{k=1}^{K} \boldsymbol{y}_k = \boldsymbol{A}\boldsymbol{C} \begin{bmatrix}\boldsymbol{z}_d \\ \boldsymbol{z}_s\end{bmatrix},$$

where $\boldsymbol{A} = [\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_K] \in \mathbb{R}^{N^2 \times KN^2}$ and $\boldsymbol{C} = [\boldsymbol{C}_1, \boldsymbol{C}_2, \ldots, \boldsymbol{C}_K]^\top \in \mathbb{R}^{KN^2 \times 2N^2}$.

Using this formulation, our aim is to estimate $\boldsymbol{z}_d$ and $\boldsymbol{z}_s$ given the measurements $\boldsymbol{y}$ and the downsampling matrix $\boldsymbol{S} = \boldsymbol{A}\boldsymbol{C}$. To simplify notation we denote the "superimage" with both diffuse and specular components as $\boldsymbol{z} = [\boldsymbol{z}_d, \boldsymbol{z}_s]^\top$.

5.3 Algorithms

We propose a general approach to separate the diffuse and specular components. Finding both components using only a single image is an under-constrained problem, which we propose to regularize using the TV norm:

$$\hat{\boldsymbol{z}} = \arg\min_{\boldsymbol{z}} \|\boldsymbol{y} - \boldsymbol{S}\boldsymbol{z}\|_2^2 + T(\boldsymbol{z}), \tag{5.2}$$

where $T(\boldsymbol{z}) = \gamma_d \|\boldsymbol{Z}_d\|_{TV} + \gamma_s \|\boldsymbol{Z}_s\|_{TV}$ is the regularization term. Here γ_d, γ_s are regularization parameters and $\|\cdot\|_{TV}$ is the total variation (TV) norm[2]. Typically, the TV norm is defined as the ℓ^1-norm of the first order differential operator, although some authors also consider the ℓ^2-norm of the same differential operator. A compromise between the two approaches is the *Huber norm* or *Huber loss*, defined as $L(\boldsymbol{x}) = \sum L(x_i)$, where

$$L(x) = \begin{cases} \frac{1}{2}x^2 & \text{for } |x| < 1 \\ |x| - \frac{1}{2} & \text{otherwise.} \end{cases}$$

[1] Note that we use bold uppercase letters for matrix notation and the same letter in lowercase for its flattened version. Throughout this chapter, we interchangeably use both notations to denote the same object.

[2] Technically, the TV norm is only a semi-norm.

We show below how to solve (5.2) with these different variations of the TV norm. The advantage of the ℓ^1 approach is that, since it favors solutions that have a sparse first order derivative, it maintains edges in the image. In contrast, the ℓ^2 approach can blur edges but performs well on other parts of the image. The ℓ^2 formulation can also lead to faster algorithms.

5.3.1 Minimizing the ℓ^2 TV norm

Although general optimization packages can be used to solve (5.2), the sampling matrix $\boldsymbol{S}$ is very large and therefore these techniques are limited to relatively small-sized images. In order to overcome this, we observe that (5.2) can be expressed as a sparse linear system. To see this, note that $T(\cdot)$, in the case of the ℓ^2 TV norm, can be written as

$$T(\boldsymbol{z}) = \boldsymbol{z}^\top \boldsymbol{D}^\top \boldsymbol{W} \boldsymbol{D} \boldsymbol{z},$$

where $\boldsymbol{D}$ is the 2D discrete differential operator matrix and $\boldsymbol{W}$ encapsulates the effect of γ_d and γ_s. Finally, setting the derivative of (5.2) to zero yields the following sparse linear system:

$$(\boldsymbol{S}^\top \boldsymbol{S} + \boldsymbol{D}^\top \boldsymbol{W} \boldsymbol{D})\hat{\boldsymbol{z}} = \boldsymbol{S}^\top \boldsymbol{y}.$$

Although this system is very large, the matrix $\boldsymbol{S}^\top \boldsymbol{S} + \boldsymbol{D}^\top \boldsymbol{W} \boldsymbol{D}$ is sparse, symmetric and positive definite, and thus can be efficiently solved using approaches such as the conjugate gradient method. This leads to an algorithm that has linear complexity in the number of pixels.

5.3.2 Minimizing the ℓ^1 TV norm

Unfortunately, unlike the ℓ^2 case, we cannot find a closed form solution with the ℓ^1-norm. Instead, we propose to use the split Bregmann method [72] to solve (5.2) iteratively. In particular, (5.2) can be written as

$$\hat{\boldsymbol{z}} = \arg\min_{\boldsymbol{z},\boldsymbol{d}} \|\boldsymbol{y} - \boldsymbol{S}\boldsymbol{z}\|_2^2 + \gamma_d \|\boldsymbol{d}_{\boldsymbol{d}}\|_1 + \gamma_s \|\boldsymbol{d}_{\boldsymbol{s}}\|_1 \quad \text{s.t.} \quad \begin{bmatrix} \boldsymbol{d}_{\boldsymbol{d}} \\ \boldsymbol{d}_{\boldsymbol{s}} \end{bmatrix} = \boldsymbol{D} \begin{bmatrix} \boldsymbol{z}_{\boldsymbol{d}} \\ \boldsymbol{z}_{\boldsymbol{s}} \end{bmatrix},$$

so that each step of the Bregmann iteration algorithm consists of solving

$$(\boldsymbol{z}^{k+1}, \boldsymbol{d}^{k+1}) = \arg\min_{\boldsymbol{z},\boldsymbol{d}} \|\boldsymbol{y} - \boldsymbol{S}\boldsymbol{z}\|_2^2 + \gamma_d \|\boldsymbol{d}_{\boldsymbol{d}}\|_1 + \gamma_s \|\boldsymbol{d}_{\boldsymbol{s}}\|_1 + \lambda \left\|\boldsymbol{d} - \boldsymbol{D}\boldsymbol{z} - \boldsymbol{b}^k\right\|_2^2, \tag{5.3}$$

where $\boldsymbol{b}^{k+1} = \boldsymbol{b}^k + (\boldsymbol{D}\boldsymbol{z}^{k+1} - \boldsymbol{d}^{k+1})$. To solve (5.3), we can alternate between minimizing over $\boldsymbol{z}$ and $\boldsymbol{d}$:

$$\boldsymbol{z}^{k+1} = \arg\min_{\boldsymbol{z}} \|\boldsymbol{y} - \boldsymbol{S}\boldsymbol{z}\|_2^2 + \lambda \left\|\boldsymbol{d}^k - \boldsymbol{D}\boldsymbol{z} - \boldsymbol{b}^k\right\|_2^2, \tag{5.4}$$

$$\boldsymbol{d}^{k+1} = \arg\min_{\boldsymbol{d}} \gamma_d \|\boldsymbol{d}_{\boldsymbol{d}}\|_1 + \gamma_s \|\boldsymbol{d}_{\boldsymbol{s}}\|_1 + \lambda \left\|\boldsymbol{d} - \boldsymbol{D}\boldsymbol{z}^{k+1} - \boldsymbol{b}^k\right\|_2^2. \tag{5.5}$$

The main advantage of this approach is that (5.5) is now decoupled over space, and thus has a closed form solution. Additionally, (5.4) is now an ℓ^2 minimization, which again has a closed form solution. In practice, only a few iterations of the split Bregmann algorithm are needed to converge so the method is acceptably fast and has linear complexity in the number of pixels in the image.

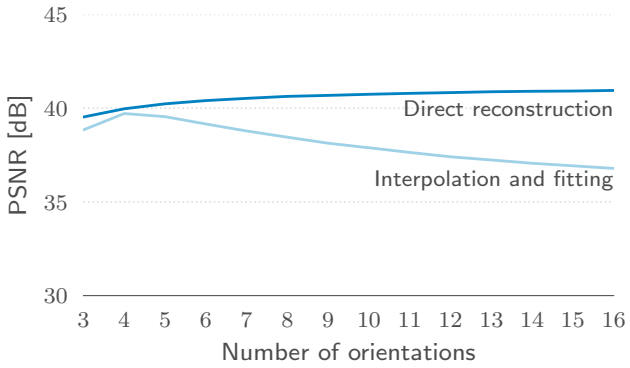


Figure 5.2: Influence of the number of orientations with a random filter design: comparison of our proposed algorithm, which performs the separation in one shot, with a two stage approach, which first interpolates the different channels and then extracts the specularity by fitting a cosine. Both methods use the ℓ^2 TV-norm as regularizer.

5.3.3 Minimizing the Huber TV norm

When using the Huber norm regularization, the minimization problem (5.2) becomes

$$\hat{\boldsymbol{z}} = \arg\min_{\boldsymbol{z},\boldsymbol{d}} \|\boldsymbol{y} - \boldsymbol{S}\boldsymbol{z}\|_2^2 + \gamma_d L\left(\boldsymbol{D_s}\boldsymbol{z_d}\right) + \gamma_s L\left(\boldsymbol{D_s}\boldsymbol{z_s}\right).$$

Setting the derivative to 0, we obtain

$$2\boldsymbol{S}^\top\boldsymbol{S}\boldsymbol{z} - 2\boldsymbol{S}^\top\boldsymbol{y} + \gamma_d\boldsymbol{D_d}^\top L'\left(\boldsymbol{D_s}\boldsymbol{z_d}\right) + \gamma_s\boldsymbol{D_s}^\top L'\left(\boldsymbol{D_s}\boldsymbol{z_s}\right) = \boldsymbol{0}. \tag{5.6}$$

As $L(\boldsymbol{z})$ is twice differentiable, (5.6) is easy to solve using, for example, Newton's method. To initialize the iteration, we propose to use the solution of the ℓ^2 TV norm.

5.3.4 Unknown light polarization

In the above description, we assumed that the polarization angle ϕ of the incoming light was known. Although it can be directly measured in controlled experiments, this is not necessary since it can be accurately estimated using the following procedure. First, for each image $\boldsymbol{x}_k$, we estimate its mean from the available pixels: $\mu_k = \sum_{n=1}^{N^2}[\boldsymbol{y}_k]_n / \sum_{n=1}^{N^2}[\boldsymbol{A}_k]_{nn}$. Using the linearity of the mean, we have

$$\begin{bmatrix}\mu_1\\ \mu_2\\ \vdots\\ \mu_K\end{bmatrix} = \mu_d + \begin{bmatrix}\cos(\phi-\theta_1)^2\\ \cos(\phi-\theta_2)^2\\ \vdots\\ \cos(\phi-\theta_K)^2\end{bmatrix}\mu_s,$$

where μ_d and μ_s are the mean values of $\boldsymbol{z}_d$ and $\boldsymbol{z}_s$. For $K > 3$, this set of non-linear equations is overdetermined, and finding ϕ, μ_d and μ_s corresponds to solving the following minimization

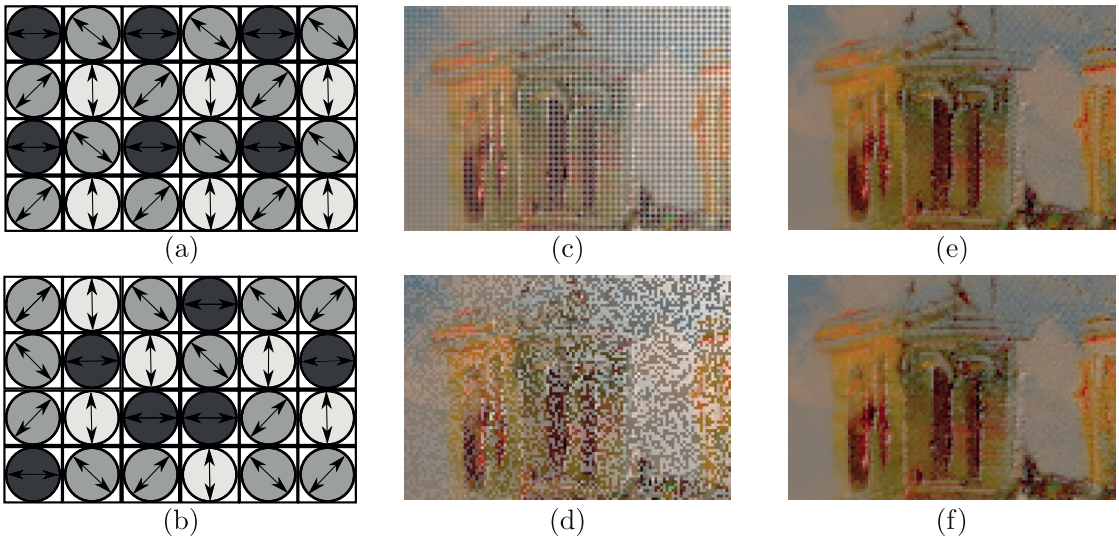


Figure 5.3: Filter designs with 4 different orientations: example of (a) a regular grid vs. (b) a random grid design. Simulated image acquisitions using (c) a regular grid and (d) a random grid filter design. The last column illustrates the estimation of the diffuse component using our algorithm.

problem:

$$\begin{bmatrix} \hat{\phi} \\ \hat{\mu}_d \\ \hat{\mu}_s \end{bmatrix} = \underset{\phi,\mu_d,\mu_s}{\arg\min} \sum_{k=1}^{K} \left(\mu_k - \mu_d - \mu_s \cos(\phi - \theta_k)^2\right)^2 .$$

5.4 Practical considerations

Before evaluating and comparing our proposed approach to existing methods, we discuss below a few algorithmic and design choices.

5.4.1 One-step vs. two-step algorithm

Our proposed algorithm separates the components and interpolates the images in a single step. One might wonder how much we gain by performing this estimation in a single step as opposed to an approach where we first demosaic every individual image $\boldsymbol{X}_k$ and then fit a cosine to the interpolated images to identify the specular and diffuse components.

We tested such a two-step approach on a number of simulated images and reported our results in Figure 5.2, which shows a comparison of our direct reconstruction algorithm with a two-stage approach. Here, the same ℓ^2 TV-norm regularization is used for both the direct approach and the interpolation step of the two-stage approach.

Even though the two approaches are close in performance when the number of orientations is 4, the two-stage approach performance clearly worsens when the number of orientations increases. Also, the two-stage approach peaks at 39.7 dB, whereas the direct reconstruction continues to

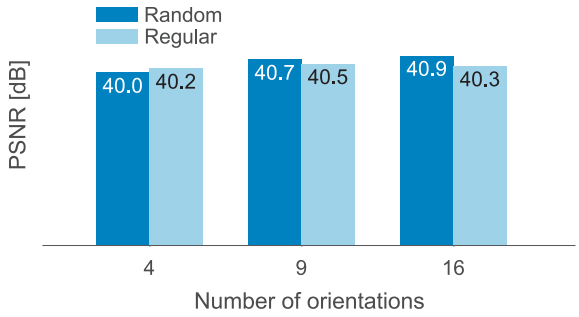


Figure 5.4: Comparison of the reconstruction error on the diffuse component with regular versus random grid designs.

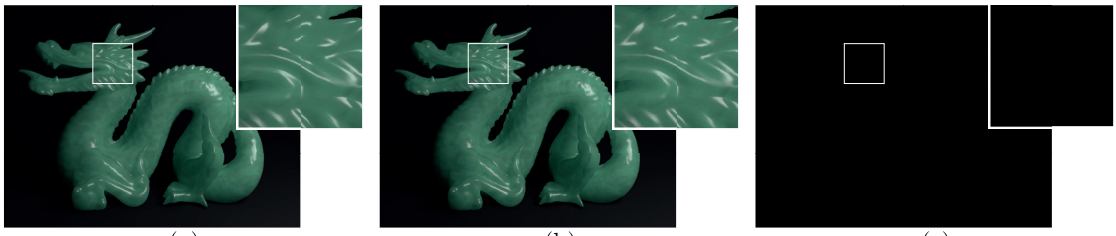
increase to 40.95 dB. While the observed performance gain is not huge, it is significant for many applications. The versatility of our algorithm over the two-stage approach is also an advantage when investigating different filter design patterns. In this regard, we observe that the direct reconstruction saturates at around 8-10 orientations, suggesting that existing camera designs based on 4 orientations are not optimal. In the rest of this chapter, we exclusively use the one-step approach.

5.4.2 Filter design patterns and number of orientations

There are a number of ways to design the micro-filter array, including (pseudo) random patterns and regular grids. Examples of images acquired with these designs are depicted in Figures 5.3c and 5.3d. To obtain these images, we use 3D computer graphics software Blender with the raytracing engine Cycles [33] and create ground truth diffuse and specular images from different render passes. We then simulate the effect of the filters by appropriately weighting and mixing the diffuse and specular components. The corresponding estimated diffuse components are shown in Figures 5.3e and 5.3f. Additionally, in Figure 5.4, we compare the average performance of these two designs for a selection of synthetic images. Overall, we observe that the filter design does not significantly influence the peak signal-to-noise ratio (PSNR). Nevertheless, as shown in Figures 5.3e and 5.3f, random patterns lead to a slightly better qualitative estimation and, more importantly, the reconstruction artifacts induced by random patterns are more pleasing to the eye.

5.4.3 Cost of introducing the filter array

Finally, it is also interesting to investigate how much is lost by using a polarizing micro-filter, compared to a traditional camera, when one does not wish to separate the diffuse and specular components. In Figure 5.5, we see that the sum of our diffuse and specular estimations is very close to the original image, suggesting that, even if the separation fails, the sum is almost indistinguishable from the output of a traditional camera.



(a) (b) (c)

Figure 5.5: Comparison of (a) the ground truth image to (b) the sum of our diffuse and specular estimations (PSNR = 43.8 dB). The difference (c) is almost zero suggesting that very little is lost by introducing a polarizing micro-filter.

5.5 Experiments

We test our ℓ^2 TV norm-based algorithm in three different scenarios: rendered images, real images with simulated polarizing filters, and real images with real polarizers. Simulated data is again generated with Blender Cycles raytracing engine [33]. The second scenario consists of real images captured by Shen and Zheng [170]: they provided the ground truth of the diffuse components along with the original images. We also provide our own captured images, taken under fixed polarized light sources, using a Nikon D810 DSLR camera with a polarizer in front of its lens. Note that even though the polarizing filter is not placed directly in front of the sensor, we neglect the effect that the lens might have on the light polarization. In practice, it may slightly change the polarization phase, but this would be accounted for in the device calibration. To obtain a diffuse ground truth, we capture one image with the polarizing filter oriented orthogonal to the light's polarization. For all setups, we first generate all complete images $\boldsymbol{X}_k$ for $k = 1, 2, \ldots, K$ and then apply the downsampling operator $\boldsymbol{A}$ to simulate the effect of the polarized filter array.

5.5.1 Single image diffuse and specular separation

Given the results from Figure 5.4, we focus on a random filter array design and sixteen orientations for our algorithm. For all experiments, we set $\gamma_d = 0.01$ and $\gamma_s = 0.002$. It is possible to obtain slightly improved results with parameter tuning, but we avoided it.

To provide context, we compare our approach with the single image color-based technique proposed by Shen and Zheng in [170]. We should emphasize that their method takes as input a standard image while ours takes an image captured with the proposed filter array. Rather than being a fair comparison between algorithms, this experiment allows us to quantify the performance improvement offered by the proposed setup.

Performing this comparison raises a number of challenges in terms of color management. In particular, our algorithm operates in a linear color space, whereas [170] uses the sRGB space. To deal with this, we run our algorithm in the linear space and convert our estimated images to sRGB for comparison. Additionally, the images provided by [170] are in sRGB format and we convert them to a linear color space to apply our algorithm.

Figures 5.6 and 5.7 depict the results for two images from each of the three different scenarios previously mentioned. For all images, our custom setup significantly outperforms [170]. This is particularly true for images with complex surfaces such as Figure 5.6b and 5.7b.

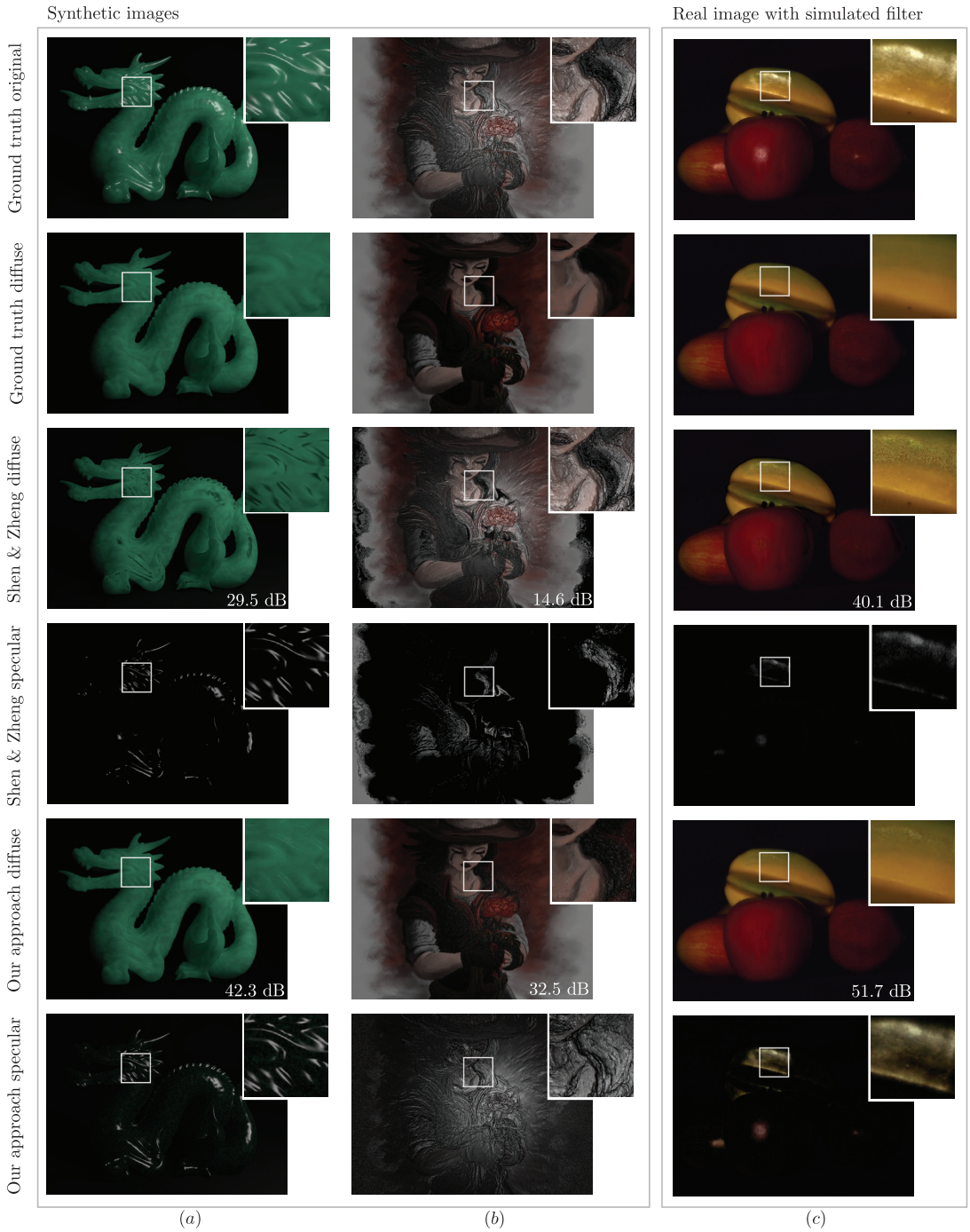


Figure 5.6: Performance evaluation on two synthetic images and one real image with a simulated polarizer: (a) the Stanford dragon [176], (b) a digital painting with a computer-generated normal map, and (c) the fruits image from [170].

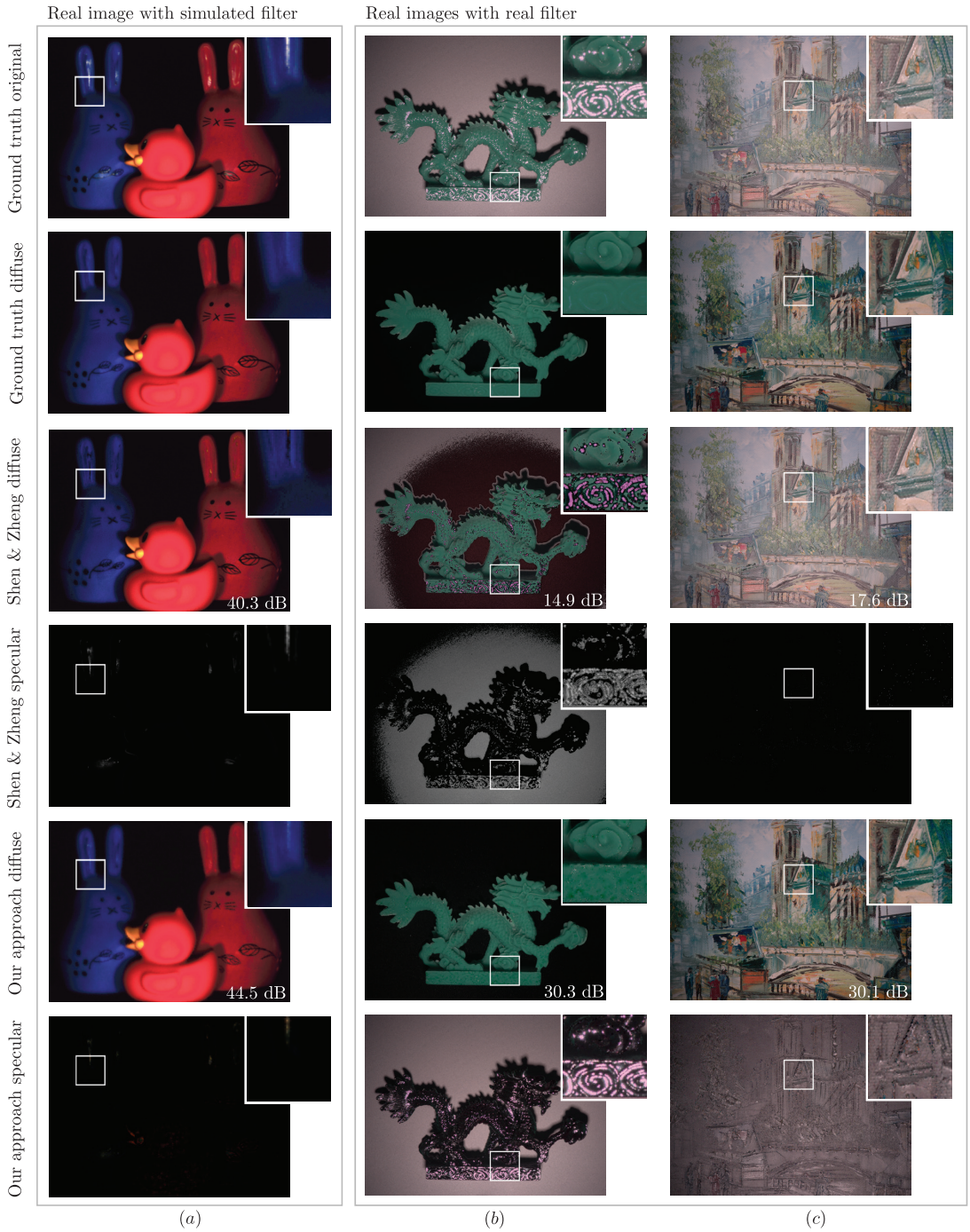


Figure 5.7: Performance evaluation on three real images, one with a simulated polarizer and two with real polarizers: (a) the animals image from [170], (b) a jade dragon, and (c) a painting.

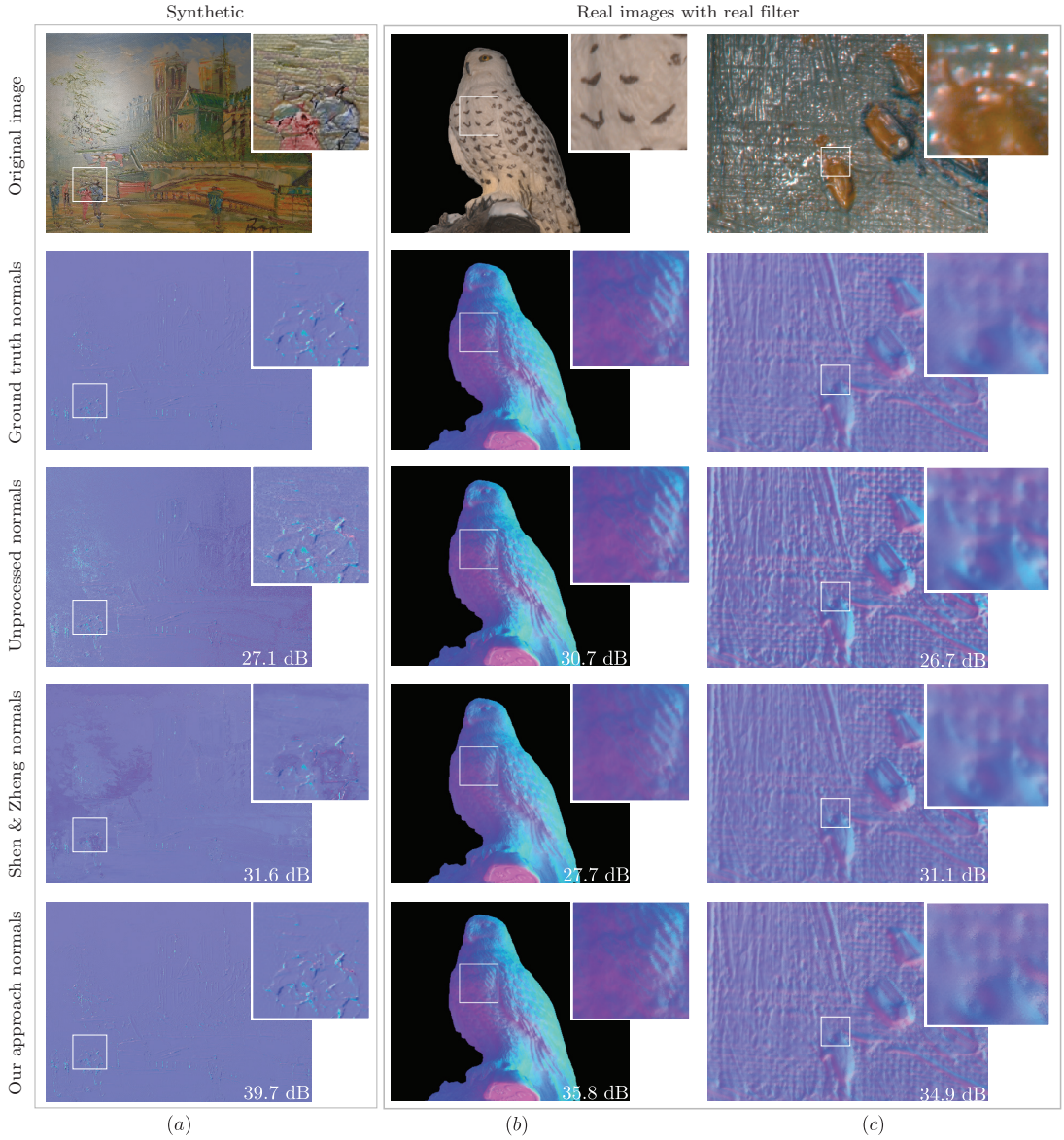


Figure 5.8: Estimation of the normal map of three different scenes: (a) a rendered painting, (b) an owl figurine, and (c) a close-up of a real painting. From top to bottom, we show the original image, the ground truth normals, the normal estimation using the original image, the normal estimation using the proposed diffuse estimate and the normal estimation using the diffuse estimate from [170].

5.5.2 Multiple image photometric stereo

We also evaluate our algorithm in a multiple image situation and demonstrate how it can improve photometric stereo [201]. Photometric stereo infers surface normals from multiple images under different lighting by assuming that the surface of the objects consists of Lambertian materials.

Unfortunately, this assumption is rarely satisfied as scenes often contain specular components and occlusions. The idea is to use our algorithm to first obtain a solid approximation of the diffuse part of a scene and later use it to recover its normals.

In this experiment, we use our *light dome* composed of 58 lamps (see Figure 3.4a) and we install linear polarizing filters with unknown orientation in front of each light source as well as in front of the camera. We capture several images, one under each illumination of the fixed lights for four different polarization orientations of the filter positioned in front of the camera; these four orientations are used to create mosaiced images on which we apply our algorithm to estimate the diffuse component. These diffuse images are then fed to the photometric stereo algorithm [201]. In Figure 5.8, we compare the normal map generated by our algorithm with the normal map obtained from unprocessed images (i.e. with no separation of the diffuse component) as well as from Shen and Zheng's algorithm [170]. Note that the ground truth normal maps are computed using the non-mosaiced images. In all scenes, the gain of using our algorithm is clearly noticeable.

5.6 Conclusion

We studied the benefits of using a camera equipped with polarizing micro-filters for the separation of diffuse and specular terms. We presented a simple algorithm to demosaic images produced by such cameras and extract the diffuse term. Regarding the diffuse extraction, we have shown that our relatively simple algorithm can significantly outperform other single-image based techniques. A more accurate knowledge of the diffuse term can then be leveraged in other imaging applications such as photometric stereo; we demonstrate that our technique improves the estimation of the normal maps on various scenes. For future work, we believe that including more elaborate priors based on, for example some of the existing color-based techniques, will improve the estimation.

Chapter 6

Super-Resolution Sparse Phase Retrieval*

All things are made of atoms—little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another. In that one sentence, you will see, there is an enormous amount of information about the world.

Six Easy Pieces
Richard Feynman

6.1 Introduction

Imagine that instead of hearing a song you can only see the absolute value of its Fourier transform (FT) on a graphic equalizer. Can you recover the song from just this visual information? The general answer is "No" as there exist infinitely many signals that fit the curve displayed by the equalizer. However, if we have additional information (or priors) about the song, we may be able to recover it successfully. The reconstruction process is the subject of this chapter and is generally known as phase retrieval (PR).

*The material in this chapter is the result of joint work with Miranda Kreković, Juri Ranieri, Amina Chebira, Yue M. Lu and Martin Vetterli. Author contributions: JR, AC, YL and MV designed research; JR devised the support recovery algorithm and its performance bound; GB and MK proposed algorithmic improvements and carried out experiments; and GB wrote the chapter based on [8], which was written by GB, MK and JR.

Beside this day-to-day example, PR is of great interest for many real-world scenarios, where it is easier to measure the FT of a signal instead of the signal itself. During the measurement process, it may happen that the phase of the FT is lost or distorted. Phase loss occurs in many scientific disciplines, particularly those involving optics and communications; the reason is that optical waves oscillate extremely fast—around 10^{15} Hz and above. These waves are measured with systems that convert photons to electrons. To date, no detection device can operate in such a high frequency range, and current systems only measure the photon flux[1], which is proportional to the squared magnitude of the electromagnetic wave [169]. A few typical examples involving phase loss are given below.

- X-ray crystallography: we measure the diffraction pattern of a crystallized molecule—that is the magnitude of its FT—and we would like to recover the structure of the molecule itself [119].
- Speckle imaging in astronomy: we measure many images of an astronomical subject and the phase of the images is compromised by the atmospheric distortion. We would like to recover the subject without the resolution downgrade imposed by the atmosphere [94].
- Blind channel estimation of multi-path communication channels: we measure samples of the channel output without knowing the input. We would like to estimate the impulse response of the channel to optimize its capacity [17].

In Chapter 4 on Lippmann photography, we also encountered an instance of PR, since the phase of the resynthesized spectrum is lost when viewing a Lippmann plate. As we saw in Section 4.7.2, the knowledge of the phase is crucial to enable the recovery of the original spectrum.

6.1.1 Previous work

The field of phase retrieval was born along with X-ray crystallography, when the first structures were solved with trial-and-error methods leveraging crystal symmetries. These initial attempts prepared the ground for more systematic approaches, a first example of which was proposed by Patterson in 1935 [144]. The Patterson method is based on locating the peaks of the auto-correlation function of the electron density to determine pairwise differences between the locations of the atoms constituting a molecule.

In the 1950s, a rich family of approaches exploiting the unique relationships between intensities and phases of measured diffraction patterns was developed, e.g. Cochran [42], Sayre [163], Karle [89]. These methods operate in the Fourier space and are known as *direct methods* because they seek to solve the phase problem directly based on the observed intensities.

We would also like to emphasize the relevance of *dual-space* algorithms, where both spatial and Fourier domains play a fundamental role in reconstructing the signal. While the origin of these methods dates back to 1972 with the work of Gerchberg and Saxton [68], a lot of interest was recently sparked by the introduction of *Charge Flipping* [138, 139].

This short literature review of phase retrieval algorithms in X-ray crystallography is focused on *ab initio* methods, which attempt to solve the phase problem with zero or very little prior information about the structure we are trying to infer. Hence, *ab initio* methods are considered very challenging, given the minimal amount of information they have access to. Successful

[1]Note that the phase of optical waves can still be measured and recorded by indirect means; such is the case in holography for instance.

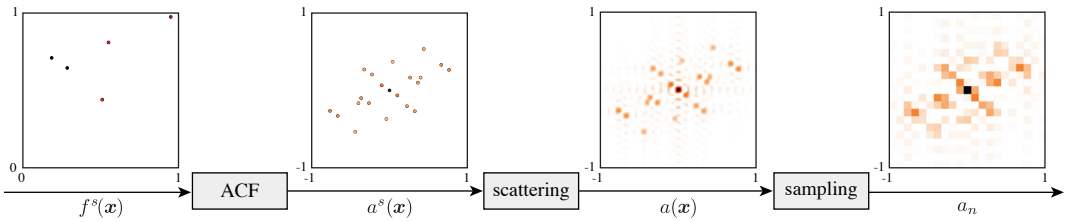


Figure 6.1: Typical PR measurement pipeline: the signal of interest $f^s(\boldsymbol{x})$ generates the autocorrelation function $a^s(\boldsymbol{x})$, which is first filtered by the scattering function $\psi(\boldsymbol{x})$ (here an ideal lowpass filter) to yield $a(\boldsymbol{x})$ and then sampled, resulting in $a_{\boldsymbol{n}}$. Note that the spatial samples $a_{\boldsymbol{n}}$ can be obtained via the inverse discrete FT of the Fourier samples $A_{\boldsymbol{m}}$, when the periodicity in the two domains holds. Darker colors represent higher intensities.

methods hinge on the design of an abstract data structure that reduces the degrees of freedom of the desired signal and simplifies its reconstruction. For example, direct methods exploit statistical relationships between the phases to reduce the number of unknowns, while Charge Flipping considers a discretization of the electron density.

In this chapter, we focus on the PR problem on sparse signals. The sparsity assumption is legitimate and encountered in many applications; for example atoms in crystallography form a sparse structure. We consider the most compact structure one can imagine for a sparse signal: a set of K atoms defined by their locations $\boldsymbol{x}_k$ and their amplitudes c_k,

$$f(\boldsymbol{x}) = \sum_{k=1}^{K} c_k \phi(\boldsymbol{x} - \boldsymbol{x}_k) = f^s(\boldsymbol{x}) * \phi(\boldsymbol{x}), \tag{6.1}$$

where $f^s(\boldsymbol{x}) = \sum_{k=1}^{K} c_k \delta(\boldsymbol{x} - \boldsymbol{x}_k)$ represents the structure, $\boldsymbol{x}$ is a spatial variable defined over $\mathbb{R}^D$ (with D being the dimensionality of the signal), $\phi(\boldsymbol{x})$ is the scattering function induced by one atom and $*$ is the convolution operator.

Even if the advantage of the compact model defined in (6.1) looks appealing, the associated algorithmic challenges are often overwhelming. Computer scientists attempted to design a scalable (i.e. with a computational complexity that is polynomial in the number of atoms K) and stable to noise algorithm that could solve all possible instances of this problem without much success; to date, it is not even clear that such an algorithm would exist [172]. In other words, we encounter a nontrivial trade-off between the compactness of such data structures (i.e. the number of unknown variables) and the ease of solving the PR problem using them. For example, Charge Flipping easily solves many PR problems in X-ray crystallography, but it is based on a discrete spatial structure, which is not the most compact representation.

Recently, we observed the emergence of new PR algorithms leveraging the notion of sparsity while assuming a discrete spatial domain. Two notable examples are *GrEedy Sparse PhAse Retrieval* (GESPAR) [168], based on the *2-opt* algorithm [45], and *Two-stage Sparse Phase Retrieval* (TSPR) [84], where the support is recovered by solving the discrete *turnpike problem* [47, 171]. Both algorithms differ from our approach in that their models are discrete and the locations are bound to a discrete grid. Even though it was not designed with continuous setups in mind, TSPR can theoretically recover locations on a continuous domain. However, while it handles noise on the measured coefficients, it does not tolerate noise in the support, which makes it impractical for continuous setups.

The major benefit of having a continuous parametric model is that it enables the estimation of the locations and amplitudes avoiding any discretization. In such a case, the achievable resolution is theoretically infinite and only limited by the noise corrupting the measurements. This is what we call *super resolution* phase retrieval.

6.1.2 Main contributions and outline

We propose a three-stage framework that precisely determines a sparse signal from the absolute value of its FT. In Section 6.2, we formalize the problem and describe the typical PR measurement pipeline. In Section 6.3, we give a high-level overview of our modular approach, discuss the main challenges and introduce a few relevant properties. Algorithms to solve these different modules are proposed in Section 6.4. We then describe the details of the proposed method to recover the support, which constitutes the critical element of the pipeline: its complexity analysis can be found in Section 6.5, together with a method to reduce its computational cost, while Section 6.6 identifies a theoretical bound (confirmed by numerical simulations) to successfully recover the signal support in a noisy regime. Then in Section 6.7, we propose a few improvements and variations of the algorithm to make it more robust to noise. In Section 6.8, we discuss the influence of the support configuration on the resulting reconstruction. Finally, Section 6.9 compares our PR pipeline with the state-of-the-art.

6.2 Problem statement

We consider the FT of the signal defined in (6.1),

$$F(\boldsymbol{\omega}) = \sum_{k=1}^{K} c_k e^{-j\boldsymbol{\omega}^\top \boldsymbol{x}_k} \Phi(\boldsymbol{\omega}), \tag{6.2}$$

where $\boldsymbol{\omega}$ is the frequency variable and $\Phi(\boldsymbol{\omega})$ is the FT of the known kernel $\phi(\boldsymbol{x})$.

In practice, it is impossible to measure the whole FT (6.2), hence we sample it. Furthermore, due to limitations of the measurement setup, we are usually only able to measure the absolute values of such samples, that we denote $|F_{\boldsymbol{m}}|$, where $F_{\boldsymbol{m}} = F(\boldsymbol{m}\Omega)$, $\boldsymbol{m} = \mathbb{Z}^D$ and Ω is the sampling frequency. As previously mentioned, the PR problem has infinite solutions without a priori knowledge of the signal $f(\boldsymbol{x})$, since we can assign any phase to the measurements and obtain a plausible reconstruction. The role of structures, such as (6.1), is to constrain the PR problem to a correct and possibly unique solution. Under the sparsity assumption, retrieving the phase is equivalent to retrieving the locations and amplitudes of $f(\boldsymbol{x})$.

The auto-correlation function (ACF) $a(\boldsymbol{x})$ of $f(\boldsymbol{x})$ is given by the inverse FT of $|F(\boldsymbol{\omega})|^2$:

$$a(\boldsymbol{x}) = f(\boldsymbol{x}) * f(-\boldsymbol{x}) = \mathcal{F}^{-1}\left[|F(\boldsymbol{\omega})|^2\right],$$

where $\mathcal{F}^{-1}$ is the inverse FT operator [190]. Interestingly, the ACF structure is completely

inherited from the signal (6.1):

$$\begin{aligned} a(\boldsymbol{x}) &= \sum_{k=1}^{K}\sum_{\ell=1}^{K} c_k c_\ell \psi(\boldsymbol{x} - (\boldsymbol{x}_k - \boldsymbol{x}_\ell)) \\ &= \left[\sum_{k=1}^{K}\sum_{\ell=1}^{K} c_k c_\ell \delta(\boldsymbol{x} - (\boldsymbol{x}_k - \boldsymbol{x}_\ell))\right] * \psi(\boldsymbol{x}) \\ &= a^s(\boldsymbol{x}) * \psi(\boldsymbol{x}), \end{aligned} \tag{6.3}$$

where the kernel $\psi(\boldsymbol{x})$ is the ACF of $\phi(\boldsymbol{x})$ and $a^s(\boldsymbol{x})$ is the ACF of the sparse structure of the train of Diracs $f^s(\boldsymbol{x})$. Equivalently, in the Fourier domain, we have

$$A(\boldsymbol{\omega}) = \sum_{k=1}^{K}\sum_{\ell=1}^{K} c_k c_\ell e^{-j\boldsymbol{\omega}^\top(\boldsymbol{x}_k - \boldsymbol{x}_\ell)} |\Phi(\boldsymbol{\omega})|^2. \tag{6.4}$$

The PR acquisition pipeline can be summarized as the filtering of the ACF $a^s(\boldsymbol{x})$ followed by sampling, where the filtering represents the scattering operation, as illustrated in Figure 6.1. We now have all the ingredients to state the core problem of this chapter.

PROBLEM 6.1 Given Fourier samples $A_{\boldsymbol{m}} = A(\boldsymbol{m}\Omega)$ of the sparse ACF defined in (6.3), recover the support $\mathcal{X} = \{\boldsymbol{x}_k\}_{k=1}^K$ and amplitudes $\{c_k\}_{k=1}^K$ determining the signal $f(\boldsymbol{x})$.

Note that the information we are interested in is hidden behind two walls: the convolution with the kernel $\psi(\boldsymbol{x})$ that spatially blurs the sparse structure of the ACF and the phase loss of the original sparse signal, $f^s(\boldsymbol{x})$, that usually characterizes any PR problem.

6.3 A three-stage approach

We propose to solve Problem 6.1 in three distinct stages: i) reconstruct the continuous ACF $a(\boldsymbol{x})$ from a set of its discrete Fourier coefficients, ii) estimate the support $\mathcal{X}$ of $f(\boldsymbol{x})$ given $a(\boldsymbol{x})$, and iii) estimate its amplitudes $\{c_k\}_{k=1}^K$.

The first step is a classical sampling problem where we would like to fully characterize a continuous signal from a set of discrete measurements.

PROBLEM 6.1.A (SPARSE ACF SUPER RESOLUTION) Given samples $A_{\boldsymbol{m}}$ of the sparse ACF as defined in (6.3), recover its continuous version $A(\boldsymbol{\omega})$.

The most well-known sampling result is due to Nyquist-Shannon-Kotelnikov and guarantees perfect recovery for signals that lie in the subspace of bandlimited functions, provided that the sampling rate is high enough.

In our case, $f(\boldsymbol{x})$ and $a(\boldsymbol{x})$ are obviously not bandlimited, but we assume that such signals are sparse, as in (6.1). Sparsity has two antagonistic effects on PR: it makes the problem combinatorial and hence hard to solve, but at the same time enables a divide-and-conquer approach,

in which we first recover the support $\mathcal{X}$ and then the amplitudes of $f(\boldsymbol{x})$. We argue that the support contains more information than the amplitudes, hence we choose to estimate it first. As an example, if all the atoms have the same amplitude, then only the support is useful to recover the original signal. On the other hand, if all the atoms have the same location, the problem is trivially solvable.

PROBLEM 6.1.B (SUPPORT RECOVERY) Assume we are given the complete set of unlabeled differences $\mathcal{D} = \{\boldsymbol{d}_{k,\ell}\}_{k,\ell} = \{\boldsymbol{x}_k - \boldsymbol{x}_\ell\}_{k,\ell}$, recover the support $\mathcal{X}$ of the sparse signal $f(\boldsymbol{x})$.

In most real-world scenarios, the unlabeled differences of Problem 6.1.B are corrupted by noise. Hence, we assume additive Gaussian noise affecting $\boldsymbol{d}_{k,\ell}$:

$$\widetilde{\boldsymbol{d}}_{k,\ell} = \boldsymbol{d}_{k,\ell} + \boldsymbol{\nu}_{k,\ell}, \tag{6.5}$$

where $\boldsymbol{\nu}_{k,\ell} \sim \mathcal{N}(\boldsymbol{0}, \sigma \boldsymbol{I})$. Furthermore, we denote the set of measured differences as $\widetilde{\mathcal{D}} = \{\widetilde{\boldsymbol{d}}_{k,\ell}\}_{k,\ell}$. For simplicity of notation, we convert the pairs of indices (k,ℓ) to $n \in \{1,\dots,N\}$, where $N = K^2 - K + 1$, and order them such that $\|\widetilde{\boldsymbol{d}}_1\| \leq \|\widetilde{\boldsymbol{d}}_2\| \leq \ldots \leq \|\widetilde{\boldsymbol{d}}_N\|$. We do not assume any ordering on the elements of $\mathcal{X}$.

In what follows, we state a few interesting observations related to Problem 6.1.B. First, when we measure a set of differences, some information is inevitably lost.

OBSERVATION 6.2 A set of points can be reconstructed from their pairwise differences, even when labeled, only up to shifts and reflections.

To show that, we first translate and reflect the set of points $\mathcal{X}$ as $\mathcal{X}' = -\mathcal{X} + \bar{\boldsymbol{x}}$, where we overload the arithmetic operators on sets to transform each point as $\boldsymbol{x}'_k = -\boldsymbol{x}_k + \bar{\boldsymbol{x}}$. Then, the set of differences of the transformed points is equivalent to the original one,

$$\boldsymbol{d}'_{k,\ell} = \boldsymbol{x}'_k - \boldsymbol{x}'_\ell = -\boldsymbol{x}_k + \bar{\boldsymbol{x}} + \boldsymbol{x}_\ell - \bar{\boldsymbol{x}} = \boldsymbol{x}_\ell - \boldsymbol{x}_k = -\boldsymbol{d}_{k,\ell},$$

where the natural symmetry of $\mathcal{D}$ compensates for the negative sign.

Second, while excluding shifts and reflections does not lead to a unique solution in general, we can still prove uniqueness under certain assumptions.

OBSERVATION 6.3 Assume that the points $\boldsymbol{x}_k$ are drawn independently at random from a sufficiently smooth distribution, then the solution is unique [155].

Third, we briefly discuss the occurrence of *collisions* in the ACF. We say that there is a collision in the ACF when two different pairs of distinct points from $\mathcal{X}$ map to the same difference in $\mathcal{D}$. Since we consider a continuous domain for the support, it inherently prevents the appearance of collisions.[2]

[2]The support recovery algorithm proposed in this chapter can in fact also handle collisions and could be used on a discretized space as well. However, assuming no collisions simplifies the recovery of the amplitudes and enables a few improvements to make the algorithm more resilient to noise.

OBSERVATION 6.4 If the locations of the points are independently drawn uniformly from a finite interval, then collisions in the ACF occur with probability zero.

Last, we note that the set of differences $\mathcal{D}$ contains many valid solutions. In particular, we can construct two solutions from every element of $\mathcal{X}$; this is a direct consequence of Observation 6.2.

OBSERVATION 6.5 The set of differences $\mathcal{D}$ is a superset of $2K$ valid solutions $\widehat{\mathcal{X}}$ to Problem 6.1.B and such solutions always contain the point zero, that is $\mathbf{0} \in \widehat{\mathcal{X}}$.

To verify this, we pick an element of the support, e.g. $\boldsymbol{x}_\ell$, and build the following tentative solution:

$$\widehat{\mathcal{X}} = \{\boldsymbol{x}_k - \boldsymbol{x}_\ell \mid k = 1, \dots, K\}. \tag{6.6}$$

Then, we notice that i) $\widehat{\mathcal{X}}$ is a valid solution with the shift fixed as $-\boldsymbol{x}_\ell$, ii) $\widehat{\mathcal{X}} \in \mathcal{D}$ and iii) we have a solution for every element of $\mathcal{X}$. Moreover, we can exploit the symmetry of the ACF to reach the aforementioned $2K$ solutions. Such an observation can be extended to the noisy case, assuming we allow the solution to be noisy as well. This property is essential to the greedy algorithm for support recovery proposed in the next section.

Once the support $\widehat{\mathcal{X}}$ of the solution has been retrieved, it remains to find the amplitudes $\{c_k\}_{k=1}^K$ of the signal $f(\boldsymbol{x})$.

PROBLEM 6.1.C (AMPLITUDE RECOVERY) Given an ACF $a(\boldsymbol{x})$ as defined in (6.3) together with the estimated support $\widehat{\mathcal{X}}$ of $f(\boldsymbol{x})$, find the amplitudes $\{c_k\}_{k=1}^K$.

6.4 Algorithms

In this section, we lay down our solutions to Problems 6.1.A, 6.1.B and 6.1.C, effectively providing an end-to-end framework to solve the sparse PR problem.

6.4.1 ACF super resolution

When we look at (6.3), we notice that $a(\boldsymbol{x})$ is completely defined by the locations $\boldsymbol{x}_k - \boldsymbol{x}_\ell$ and the amplitudes $c_k c_\ell$. Hence, we can recast Problem 6.1.A as a parameter estimation problem given the measured samples $A_{\boldsymbol{m}}$ of the FT of the ACF. As we have seen in Chapter 2, an effective approach to do so is finite rate of innovation (FRI) sampling. Recall that the essential ingredient in FRI is to represent the signal of interest as a weighted sum of complex exponentials in the following form:

$$b_m = \sum_{n=1}^{N} \alpha_n u_n^m. \tag{6.7}$$

Note that in this subsection, we restrict ourselves to the 1-dimensional case for clarity, even though our implementation is generalized to higher dimensions (see [113, 141] for more details). The formulation (6.7) has several similarities with (6.4); to see this, we define $t_n = x_k - x_\ell$, substitute $\alpha_n = c_k c_\ell$ and $u_n = \exp(-j\Omega t_n)$ and rewrite the sampled ACF A_m as follows,

$$A_m = \sum_{n=1}^{N} \alpha_n u_n^m |\Phi(m\Omega)|^2. \tag{6.8}$$

We remark that $|\Phi(m\Omega)|^2$ does not allow us to express (6.8) as a sum of complex exponentials yet. However, if we assume that the kernel function $\phi(x)$ is an ideal low-pass filter[3], i.e. a sinc function, its FT becomes a box function. Thus, we can ignore such a kernel for some neighborhood of m around zero, since $\Phi(m\Omega) = 1$ for $|m\Omega|$ smaller than the bandwidth of the signal.

Then, from (6.8), we can fully recover the locations $\{d_n\}_{n=1}^N$ by identifying the filter annihilating the signal A_m (see Section 2.2.3 for more details). Our situation differs from usual FRI applications in the sense that the locations of the ACF describe a symmetric structure. As a consequence, all roots u_n come in conjugate pairs (except for the one corresponding to the zero location).

Once the locations are known, the amplitudes $\alpha_n = c_k c_\ell$ are found by injecting u_n in (6.8) and solving a linear system of equations.

6.4.2 Support recovery

For the recovery of the support, we propose a novel greedy algorithm that is initialized with a partial solution $\widehat{\mathcal{X}}_2$, which contains two locations. At a given iteration k, we generate a partial solution $\widehat{\mathcal{X}}_{k+1}$ composed of $k+1$ locations, hence the algorithm has a total of $K-2$ iterations indexed from 2 to $K-1$.

Initialization

From Observation 6.5, we know that the solution set $\widehat{\mathcal{X}}$ is contained in $\widetilde{\mathcal{D}}$ and $\mathbf{0} \in \mathcal{X}$; this gives us the first point of the solution, that is $\widehat{\boldsymbol{x}}_1 = \mathbf{0}$. Next, we identify the element $\widetilde{\boldsymbol{d}}_N$ in $\widetilde{\mathcal{D}}$ with the largest norm, so that we maximize the noise resilience of our algorithm. Indeed, assuming that the locations are corrupted by identically distributed noise, picking the largest norm ensures the maximal SNR of our initial solution. The value $\widetilde{\boldsymbol{d}}_N$ is the noisy difference between two unknown locations of $f(\boldsymbol{x})$; without loss of generality, we call them $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. The elements $\widehat{\boldsymbol{x}}_1 = \mathbf{0}$ and $\widehat{\boldsymbol{x}}_2 = \widetilde{\boldsymbol{d}}_N$ are nothing but $\boldsymbol{x}_1$ and $\boldsymbol{x}_2 + \boldsymbol{\nu}_{2,1}$ translated by $-\boldsymbol{x}_1$. Therefore, we are always guaranteed that the initialized solution $\widehat{\mathcal{X}}_2 = \{\mathbf{0}, \widetilde{\boldsymbol{d}}_N\}$ is a (noisy) subset of the set of locations $\mathcal{X} - \boldsymbol{x}_1$.

Referring again to Observation 6.5, we know that the set of differences $\widetilde{\mathcal{D}}$ contains the rest of the points $\{\boldsymbol{x}_k - \boldsymbol{x}_1 + \boldsymbol{\nu}_{k,1}\}_{k=3}^K$, that should belong to the final solution $\widehat{\mathcal{X}} = \widehat{\mathcal{X}}_K$. Furthermore, since we do not want to duplicate points in $\widehat{\mathcal{X}}_k$, we initialize a set of possible elements of the solution $\mathcal{P}_2 = \widetilde{\mathcal{D}} \setminus \{\widetilde{\boldsymbol{d}}_1, \widetilde{\boldsymbol{d}}_N\}$. Due to noise, the vector $\mathbf{0}$ is not in $\widetilde{\mathcal{D}}$, so we remove the closest element $\widetilde{\boldsymbol{d}}_1$.

[3] As mentioned in Section 2.3.4, the FRI theory has also been generalized to a wide range of kernels such as combinations of B-splines and E-splines [55, 186, 188] or even arbitrary sampling kernels [189], where a linear operation enables us to obtain the desired form (6.7) from (6.8).

Algorithm 6.1 Support recovery

Input: A set of $N = K^2 - K + 1$ differences $\widetilde{\mathcal{D}} = \{\widetilde{\boldsymbol{d}}_n\}_{n=1}^N$ ordered by their norms
Output: A set of K points $\widehat{\mathcal{X}}$ such that their pairwise differences generate $\widetilde{\mathcal{D}}$

$\widehat{\mathcal{X}}_2 = \{\mathbf{0}, \widetilde{\boldsymbol{d}}_N\}$
$\mathcal{P}_2 = \widetilde{\mathcal{D}} \setminus \{\widetilde{\boldsymbol{d}}_1, \widetilde{\boldsymbol{d}}_N\}$
for $k = 2, \dots, K-1$ **do**
  $\widehat{\boldsymbol{x}}_{k+1} = \arg\min_{\boldsymbol{p} \in \mathcal{P}_k} \sum_{\widehat{\boldsymbol{x}} \in \widehat{\mathcal{X}}_k} \min_{\widetilde{\boldsymbol{d}} \in \widetilde{\mathcal{D}}} \left\| \boldsymbol{p} - \widehat{\boldsymbol{x}} - \widetilde{\boldsymbol{d}} \right\|^2$
  $\widehat{\mathcal{X}}_{k+1} = \widehat{\mathcal{X}}_k \cup \widehat{\boldsymbol{x}}_{k+1}$
  $\mathcal{P}_{k+1} = \mathcal{P}_k \setminus \widehat{\boldsymbol{x}}_{k+1}$
end for
return $\widehat{\mathcal{X}}_K$

Main algorithm

At each step k, we identify the element in $\mathcal{P}_k$ that, when added to the partial solution $\widehat{\mathcal{X}}_k$, minimizes the error with respect to the measured set of differences $\widetilde{\mathcal{D}}$. More precisely, at every iteration k we solve the following optimization problem,

$$\widehat{\boldsymbol{x}}_{k+1} = \underset{\boldsymbol{p} \in \mathcal{P}_k}{\arg\min} \sum_{\widehat{\boldsymbol{x}} \in \widehat{\mathcal{X}}_k} \min_{\widetilde{\boldsymbol{d}} \in \widetilde{\mathcal{D}}} \left\| \boldsymbol{p} - \widehat{\boldsymbol{x}} - \widetilde{\boldsymbol{d}} \right\|^2. \tag{6.9}$$

Intuitively speaking, we would like to identify the element $\widehat{\boldsymbol{x}}_{k+1} \in \mathcal{P}_k$ such that the set of pairwise differences of the points in $\widehat{\mathcal{X}}_{k+1} = \widehat{\mathcal{X}}_k \cup \widehat{\boldsymbol{x}}_{k+1}$ is the closest to a subset of the measured $\widetilde{\mathcal{D}}$. The main challenge is that we do not know the correct labeling between these two sets. In the noiseless case, we are looking for a set $\widehat{\mathcal{X}}_{k+1}$, whose pairwise differences form exactly a subset of $\widetilde{\mathcal{D}}$. Hence, we can solve the labeling by matching identical elements. In the noisy case, we cannot leverage the definition of a subset. Therefore, we loosen the equality between elements and determine the labeling by searching for the differences in $\widetilde{\mathcal{D}}$ that are closest in ℓ^2-norm to the pairwise differences of the elements in $\widehat{\mathcal{X}}_{k+1}$. This procedure is summarized in Algorithm 6.1 and its application on the ACF $a^s(\boldsymbol{x})$ from Figure 6.1 is illustrated in Figure 6.2.

6.4.3 Amplitude recovery

If we assume that collisions can occur, recovering the amplitudes with a given ACF and support is equivalent to solving a system of quadratic equations. However, if there are no collisions, we suggest a simple but efficient algebraic solution to Problem 6.1.C, inspired from [155]. Our new approach is different in that it avoids a matrix inversion step and hence, it is both faster and more robust to noise.

Let $\boldsymbol{c} = [c_1, c_2, \dots, c_K]^\top$ be a vector made of the amplitudes to be recovered. If we define a matrix $\boldsymbol{C} = \boldsymbol{c}\boldsymbol{c}^\top$, all the elements outside of the diagonal of such a matrix are the amplitudes of the measured ACF, that is $C_{i,j} = c_i c_j$. Notice that we cannot observe the diagonal entries $C_{i,i} = c_{i,i}^2$ as we just have access to their sum $a_{\mathbf{0}}^s = \sum_i c_{i,i}^2$, which is the value of the ACF at $\mathbf{0}$. This is unfortunate since they are precisely the values we are interested in, up to a squaring operator.

We recast Problem 6.1.C as a matrix completion problem, where we would like to estimate the diagonal entries $C_{i,i}$ under the constraint of $\boldsymbol{C}$ being a rank-one matrix. The first step of

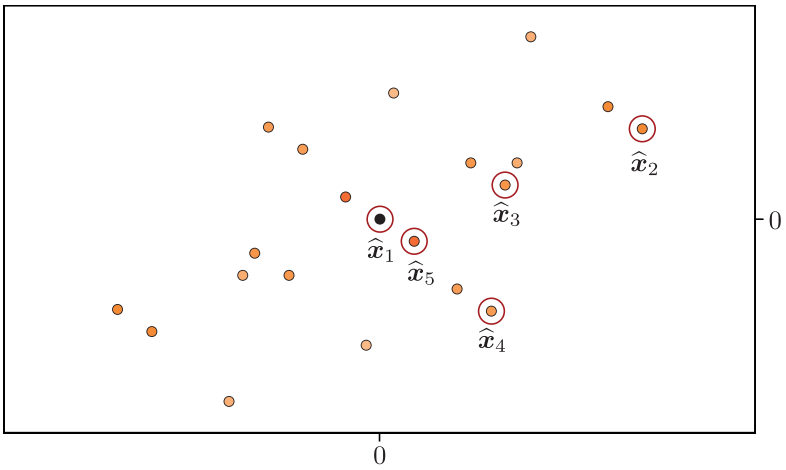


Figure 6.2: Instance of Algorithm 6.1 on the ACF $a^s(\boldsymbol{x})$ from Figure 6.1. First, we identify $\widehat{\boldsymbol{x}}_1$, the element with the smallest norm and set it to zero. Second, we find $\widehat{\boldsymbol{x}}_2$, the point with the largest norm. Then, points $\widehat{\boldsymbol{x}}_3$ to $\widehat{\boldsymbol{x}}_5$ are selected in a greedy way according to (6.9). The solution coincides with the initial signal $f^s(\boldsymbol{x})$ displayed in Figure 6.1.

our proposed method is to introduce a matrix $\boldsymbol{L}$ such that

$$L_{i,j} = \begin{cases} \log(C_{i,j}) = \ell_i + \ell_j & \text{for } i \neq j \\ 0 & \text{otherwise,} \end{cases}$$

where $\ell_i = \log(c_i)$. The sum of the ith row of $\boldsymbol{L}$ is given by

$$\sum_{j=1}^{K} L_{i,j} = (K-1)\ell_i + \sum_{j=1}^{K} \ell_j - \ell_i = (K-2)\ell_i + \sum_{j=1}^{K} \ell_j,$$

where the term $\sum_j \ell_j$ does not vary between rows. Hence, its value can be obtained from summing all the entries in $\boldsymbol{L}$,

$$s = \sum_{i=1}^{K}\sum_{j=1}^{K} L_{i,j} = (K-2)\sum_{i=1}^{K} \ell_i + K\sum_{j=1}^{K} \ell_j = 2(K-1)\sum_{j=1}^{K} \ell_j.$$

Then, we recover the vector $\boldsymbol{\ell} = [\ell_1, \ell_2, \ldots, \ell_K]^\top$ for $K > 2$ as

$$\boldsymbol{\ell} = \frac{1}{K-2}\left(\sum_{j=1}^{K} L_{i,j} - \frac{s}{2(K-1)}\mathbb{1}\right),$$

where $\mathbb{1}$ is the all-ones vector.[4] Finally, it suffices to compute $c_i = \exp(\ell_i)$ to retrieve the amplitudes.

Note that this solution assumes that $\boldsymbol{C}$ is symmetric; this might not be the case in a noisy setup, but we enforce it by replacing $\boldsymbol{C}$ with $\frac{1}{2}(\boldsymbol{C} + \boldsymbol{C}^\top)$. This strategy yields the symmetric

[4]When $K = 2$, the entries ℓ_1, ℓ_2 can be recovered by solving a system of two equations.

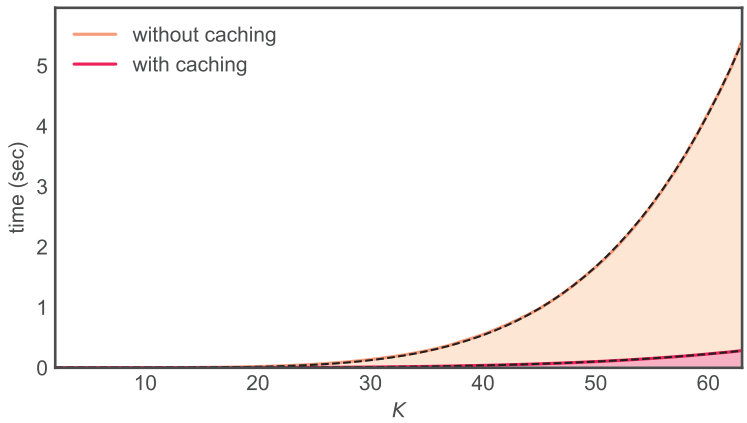


Figure 6.3: Comparison of the average run time of the original algorithm and its cached version. The times reported are the average of 100 runs of the algorithm. The dashed lines represent curves of the form CK^α that are fitted to the data. For the method without caching, we have $C = 4.25 \cdot 10^{-6}$ and $\alpha = 5.06$, while for the method with caching we have $C = 3.88 \cdot 10^{-6}$ and $\alpha = 4.37$. Remark how the caching is reducing the polynomial degree of the computational cost by approximately one.

matrix that is the closest to $\boldsymbol{C}$ with respect to the Frobenius norm. In case of collisions, the problem does not have an algebraic solution and a possible convex relaxation is provided in [84]. In practice, this is often not a concern due to Observation 6.4.

In what follows, we study and propose improvements to the performance of our PR algorithm, focusing our attention on the support recovery step, i.e. Algorithm 6.1. In fact, the first step—the super-resolution with FRI—is well represented in literature, where theoretical analyses, extensive simulations in noisy scenarios and efficient denoising schemes have been proposed [54, 113, 142]; Chapter 2 also provides a more thorough analysis of the method. On the other hand, the amplitude recovery, while being novel, only consists of simple algebraic manipulations that are not computationally costly.

6.5 Complexity analysis

Algorithm 6.1 has K rounds. In each of these rounds, we go through all points in the existing solution set $\widehat{\mathcal{X}}_k$, and for each point we compute the difference with all the values in $\widetilde{\mathcal{D}}$. Since there are $\mathcal{O}(K)$ points in $\widehat{\mathcal{X}}_k$ and $\mathcal{O}(K^2)$ elements in $\widetilde{\mathcal{D}}$, this is done in $\mathcal{O}(K^3)$ operations. Furthermore, for each of these computed differences, we need to find the closest element in $\widetilde{\mathcal{D}}$, which requires additional $\mathcal{O}(K^2)$ comparisons. In total, the complexity of our algorithm is $\mathcal{O}(K^6)$. Even though this is high and limits the field of application to reasonable sizes, it compares favorably to an exhaustive search strategy, which grows exponentially with K.

It is possible to trade time complexity for storage complexity. Indeed, we observe that we

compute at each round the following values

$$\widetilde{\boldsymbol{d}}_{i,j} = \arg\min_{\widetilde{\boldsymbol{d}} \in \widetilde{\mathcal{D}}} \|\boldsymbol{p}_j - \widehat{\boldsymbol{x}}_i - \widetilde{\boldsymbol{d}}\|^2, \tag{6.10}$$

for every point $\widehat{\boldsymbol{x}}_i \in \widehat{\mathcal{X}}_k$ and candidate $\boldsymbol{p}_j \in \mathcal{P}_k$. However, since we are just moving one element from $\mathcal{P}_k$ to $\widehat{\mathcal{X}}_{k+1}$ at each iteration, we propose to cache the values (6.10) in a lookup table to reduce the total computational cost. By doing so, we only need to update each $\widetilde{\boldsymbol{d}}_{i,j}$ when the corresponding candidate $\boldsymbol{p}_j$ is removed from $\mathcal{P}_k$ to be added to $\widehat{\mathcal{X}}_{k+1}$.

The theoretical complexity when caching $\widetilde{\boldsymbol{d}}_{i,j}$ is not trivial to analyze, but in practice we notice a significant improvement, as illustrated in Figure 6.3.

6.6 Performance analysis

In what follows, we study the expected performance of Algorithm 6.1 in the presence of noise. More precisely, we estimate the expected mean squared error (MSE) of the support recovery algorithm when the correct solution is obtained in Section 6.6.1. Then, we approximate the probability of the algorithm to find such a correct solution in Section 6.6.2. We consider the problem with $D = 1$ dimension to lighten notation and simplify the discussion. However, all the results can be easily generalized to the multidimensional setup introduced in Problem 6.1.B.

6.6.1 Expected performance

After $K - 2$ iterations and if the algorithm successfully finds any correct solution as defined in (6.6), then this solution will be noisy as it is constructed by selecting noisy elements from $\widetilde{\mathcal{D}}$, see (6.5). If we assume Gaussian noise affecting the support, then the MSE of the support recovery solution can be computed as

$$\text{MSE} = \|\mathcal{X} - \widehat{\mathcal{X}}\|_2^2 = \sum_{k=2}^{K} \|\boldsymbol{\nu}_{k,1}\|^2 = \sigma^2 Q_{K-1},$$

where $Q_K \sim \chi_K^2$ follows a chi-squared distribution with K degrees of freedom. Therefore, the expected value of the MSE of any correct solution is

$$\mathbb{E}[\text{MSE}] = (K-1)\sigma^2.$$

6.6.2 Probability of success

We model the probability that Algorithm 6.1 finds the correct solution as a function of the noise variance σ^2 and the number of elements K to characterize its performance. Such a probability can be factored as $K - 2$ iterations

$$P(\sigma, K) = \prod_{k=2}^{K-1} P_k(\sigma, K),$$

where $P(\sigma, K)$ is the probability of success of the support recovery algorithm and $P_k(\sigma, K)$ is the conditional probability of success at iteration k, given that the algorithm was correct until iteration $k - 1$.

We focus our attention on what happens at iteration k, i.e. we study the probability $P_k(\sigma, K)$. First, we split the set of possible elements of the solutions $\mathcal{P}_k$ as the union of two disjoint sets: $\mathcal{C}_k$ containing the elements that when picked by the algorithm generate a correct partial solution at iteration k, and $\mathcal{W}$ containing the elements that when picked corrupt the partial solution. Second, we generalize the cost function used in the main optimization problem (6.9) to a generic set of 1D elements $\mathcal{A}$ as,

$$g(\mathcal{A}, \widehat{\mathcal{X}}_k) = \min_{p \in \mathcal{A}} \sum_{\widehat{x} \in \widehat{\mathcal{X}}_k} \min_{\widetilde{d} \in \widetilde{\mathcal{D}}} \left(p - \widehat{x} - \widetilde{d} \right)^2 . \tag{6.11}$$

Below, we use $g(\mathcal{A}, \widehat{\mathcal{X}}_k)$ with both sets and single elements as arguments: in other words, the expression $g(a, \widehat{\mathcal{X}}_k)$ is interpreted as $g(\{a\}, \widehat{\mathcal{X}}_k)$.

Then, we compute the probability that the support recovery algorithm picks an element from $\mathcal{C}_k$ instead of an element from $\mathcal{W}$, when searching for the solution of (6.9). This happens if the cost of $\mathcal{C}_k$ is smaller than the one of $\mathcal{W}$ measured via (6.11),

$$\begin{aligned} P_k(\sigma, K) &= \mathrm{P}(g(\mathcal{C}_k, \widehat{\mathcal{X}}_k) < g(\mathcal{W}, \widehat{\mathcal{X}}_k)) \\ &= \mathrm{P}(\exists c \in \mathcal{C}_k | g(c, \widehat{\mathcal{X}}_k) < g(\mathcal{W}, \widehat{\mathcal{X}}_k)) \\ &= 1 - \mathrm{P}(\nexists c \in \mathcal{C}_k \mid g(c, \widehat{\mathcal{X}}_k) < g(\mathcal{W}, \widehat{\mathcal{X}}_k)) \\ &= 1 - \mathrm{P}(\forall c \in \mathcal{C}_k \mid g(c, \widehat{\mathcal{X}}_k) \geq g(\mathcal{W}, \widehat{\mathcal{X}}_k)). \end{aligned} \tag{6.12}$$

We assume that the events $g(c, \widehat{\mathcal{X}}_k) \geq g(\mathcal{W}, \widehat{\mathcal{X}}_k)$ are independent for all $c \in \mathcal{C}_k$ and obtain

$$P_k(\sigma, K) = 1 - \prod_{c \in \mathcal{C}_k} \mathrm{P}\left(\frac{g(c, \widehat{\mathcal{X}}_k)}{g(\mathcal{W}, \widehat{\mathcal{X}}_k)} \geq 1 \right). \tag{6.13}$$

This is a crude simplification, but it enables us to compute an approximation of $P_k(\sigma, K)$ that will not impair the quality of the end result, as we will demonstrate later. With a similar development as (6.12), we can write

$$\mathrm{P}\left(\frac{g(c, \widehat{\mathcal{X}}_k)}{g(\mathcal{W}, \widehat{\mathcal{X}}_k)} \geq 1 \right) = 1 - \mathrm{P}\left(\forall w \in \mathcal{W} \;\middle|\; \frac{g(c, \widehat{\mathcal{X}}_k)}{g(w, \widehat{\mathcal{X}}_k)} < 1 \right).$$

Again, we approximate $P_k(\sigma, K)$ assuming the independence of the events $g(w, \widehat{\mathcal{X}}_k)$ as

$$P_k(\sigma, K) = 1 - \prod_{c \in \mathcal{C}_k} \left(1 - \prod_{w \in \mathcal{W}} \mathrm{P}\left(\frac{g(c, \widehat{\mathcal{X}}_k)}{g(w, \widehat{\mathcal{X}}_k)} < 1 \right) \right). \tag{6.14}$$

Then, we focus our attention on the term $\mathrm{P}\left(\frac{g(c, \widehat{\mathcal{X}}_k)}{g(w, \widehat{\mathcal{X}}_k)} < 1 \right)$. First, we compute the cost of adding an element c from $\mathcal{C}_k$ to $\widehat{\mathcal{X}}_{k+1}$,

$$\begin{aligned} g(c, \widehat{\mathcal{X}}_k) &= \sum_{\widehat{x} \in \widehat{\mathcal{X}}_k} \min_{\widetilde{d} \in \widetilde{\mathcal{D}}} \left(c - \widehat{x} - \widetilde{d} \right)^2 \\ &= \sum_{\ell=1}^{k} \min_{\widetilde{d} \in \widetilde{\mathcal{D}}} \left(c - (x_\ell - x_1 + \nu_{\ell,1}) - \widetilde{d} \right)^2 , \end{aligned} \tag{6.15}$$

where each $\widehat{x} \in \widehat{\mathcal{X}}_k$ is a shifted noisy version of an element of $\mathcal{X}$. Following a similar reasoning, the newly added element c can be expressed as $c = x_{k+1} - x_1 + \nu_{k+1,1}$. By substituting this expression into (6.15), we further obtain

$$\begin{aligned} g(c, \widehat{\mathcal{X}}_k) &= \sum_{\ell=1}^{k} \min_{\widetilde{d} \in \widetilde{\mathcal{D}}} \left(x_{k+1} + \nu_{k+1,1} - x_\ell - \nu_{\ell,1} - \widetilde{d} \right)^2 \\ &\overset{(a)}{\approx} \sum_{\ell=1}^{k} \left(\nu_{k+1,1} - \nu_{\ell,1} - \nu_{k+1,\ell} \right)^2 \\ &= 3\sigma^2 Q_k^{(1)}, \end{aligned} \tag{6.16}$$

where $Q_k^{(1)} \sim \chi_k^2$, and in (a) we approximate $g(c, \widehat{\mathcal{X}}_k)$ by selecting the difference $\widetilde{d} = x_{k+1} - x_\ell + \nu_{k+1,\ell}$. We select this specific $\widetilde{d}$ as it is likely to be picked, provided that the noise variance σ^2 is small compared to the values x_i. This also significantly simplifies (6.16) by dropping the random variables x_1, x_{k+1}, and x_ℓ.

Second, we analyze the cost of making an error $g(w, \widehat{\mathcal{X}}_k)$ at iteration k—that is selecting any element $w \in \mathcal{W}$ given $\widehat{\mathcal{X}}_k$:

$$g(w, \widehat{\mathcal{X}}_k) = \sum_{\widehat{x} \in \widehat{\mathcal{X}}_k} \min_{\widetilde{d} \in \widetilde{\mathcal{D}}} \left(w - \widehat{x} - \widetilde{d} \right)^2. \tag{6.17}$$

We express the minimum in (6.17) as an exhaustive check of all the possible selections of k differences from $\widetilde{\mathcal{D}}$. To do so, we define $\mathcal{M}_k$ as the set containing all the k-permutations of $\widetilde{\mathcal{D}}$, and rewrite the probability of selecting a correct location c instead of a wrong one w for any given c and w from (6.14) as follows,

$$\mathrm{P}\left(\frac{g(c, \widehat{\mathcal{X}}_k)}{g(w, \widehat{\mathcal{X}}_k)} < 1 \right) = \mathrm{P}\left(\frac{g(c, \widehat{\mathcal{X}}_k)}{e(w, \boldsymbol{\pi}, \widehat{\mathcal{X}}_k)} < 1, \ \forall \boldsymbol{\pi} \in \mathcal{M}_k \right).$$

Here, we introduced $e(w, \boldsymbol{\pi}, \widehat{\mathcal{X}}_k)$ as the cost for a given permutation $\boldsymbol{\pi}$,

$$e(w, \boldsymbol{\pi}, \widehat{\mathcal{X}}_k) = \sum_{i=1}^{k} \left(w - \widehat{x}_i - \pi_i \right)^2, \tag{6.18}$$

where the elements in $\boldsymbol{\pi}$ and $\widehat{\mathcal{X}}_k$ are indexed with i.

Once more, we assume that all these selections are independent to obtain

$$\mathrm{P}\left(\frac{g(c, \widehat{\mathcal{X}}_k)}{g(w, \widehat{\mathcal{X}}_k)} < 1 \right) = \prod_{\boldsymbol{\pi} \in \mathcal{M}_k} \mathrm{P}\left(\frac{g(c, \widehat{\mathcal{X}}_k)}{e(w, \boldsymbol{\pi}, \widehat{\mathcal{X}}_k)} < 1 \right). \tag{6.19}$$

Finally, we discuss the probabilistic aspects of (6.18). The terms ω, $\widehat{x}_i$ and π_i are each made of the difference between two points plus a noise value. Indeed, they have the form

$$p = x_i - x_j + \nu_{i,j},$$

for some specific indices i and j. Assuming that the points in $\mathcal{X}$ are uniformly distributed between -0.5 and 0.5, and the noise is Gaussian with zero mean and variance σ^2, we can

approximate (6.18) as

$$\begin{aligned} e(w, \boldsymbol{\pi}, \widehat{\mathcal{X}}_k) &\overset{(a)}{\approx} \sum_{\ell=1}^{k} \left(\sum_{i=1}^{6} Y_i + \sum_{j=1}^{3} Z_j \right)^2 \\ &\overset{(b)}{\approx} \sum_{\ell=1}^{k} \left(W + \sum_{j=1}^{3} Z_j \right)^2 \\ &= \left(3\sigma^2 + \frac{1}{2} \right) Q_k^{(2)}, \end{aligned}$$

where $Q_k^{(2)} \sim \chi_k^2$, $Y_i \sim U[-0.5, 0.5]$, $Z_j \sim \mathcal{N}(0, \sigma^2)$ and $W \sim \mathcal{N}(0, \frac{1}{2})$. In (a), we approximate the sum by assuming independence between all the random variables and in (b) we approximate the sum of six random variables uniformly distributed on $[-0.5, 0.5]$ with a normal random variable with variance $\sigma^2 = \frac{1}{2}$.

We now have all the ingredients to compute the probability of success at iteration k (6.14), as

$$\begin{aligned} P_k(\sigma, K) &= 1 - \prod_{c \in \mathcal{C}_k} \left(1 - \prod_{w \in \mathcal{W}} \mathrm{P} \left(\frac{g(c, \widehat{\mathcal{X}}_k)}{g(w, \widehat{\mathcal{X}}_k)} < 1 \right) \right) \\ &\approx 1 - \prod_{c \in \mathcal{C}_k} \left(1 - \prod_{w \in \mathcal{W}} \prod_{\mathcal{S} \in \mathcal{M}_k} \mathrm{P} \left(\frac{Q_k^{(1)}}{Q_k^{(2)}} < \frac{3\sigma^2 + \frac{1}{2}}{3\sigma^2} \right) \right) \\ &= 1 - \left(1 - \mathrm{P} \left(\frac{Q_k^{(1)}}{Q_k^{(2)}} < \frac{3\sigma^2 + \frac{1}{2}}{3\sigma^2} \right)^{|\mathcal{M}_k||\mathcal{W}|} \right)^{|\mathcal{C}_k|} \\ &= 1 - \left(1 - \mathrm{F} \left(\frac{3\sigma^2 + \frac{1}{2}}{3\sigma^2}, k, k \right)^{|\mathcal{M}_k||\mathcal{W}|} \right)^{|\mathcal{C}_k|}, \end{aligned} \tag{6.20}$$

where $\mathrm{F}(x, k_1, k_2)$ is the cumulative distribution function of an F-distribution with parameters k_1 and k_2; it can be calculated using regularized incomplete beta functions. Last, we determine the size of the sets as

$$\begin{aligned} |\mathcal{C}_k| &= K - k, \\ |\mathcal{W}| &= N - K = K^2 - 2K + 1, \\ |\mathcal{M}_k| &= N^k. \end{aligned}$$

Note that, as the number of points K increases, these exponents grow faster and push any probability that is not 1 to 0; hence, we expect a steep phase transition.

Along the path of our analysis, we made a few rough assumptions that we cannot theoretically justify regarding the independence of events, e.g. in (6.13), (6.14) and (6.19). While we would like to be more rigorous, we provide below numerical evidence that such assumptions hold in practice as the algorithm's performance exhibits a phase transition matching closely the derived theoretical bound (6.20).

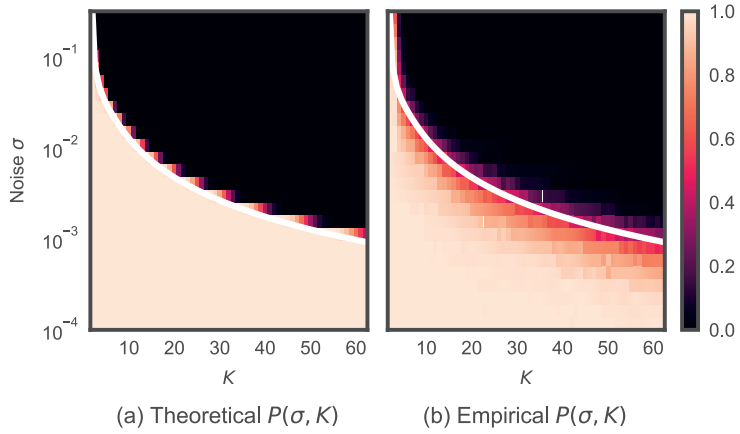


Figure 6.4: Comparison of the (a) theoretical and (b) empirical probability of success for Algorithm 6.1 in 2 dimensions with respect to the size of the problem K and the noise σ affecting the set of differences. In both plots, the white line represents $P(\sigma, K) = 0.5$.

6.6.3 Numerical simulations

We define the *index-based* error as a binary metric that is equal to 0 if the solution set $\widehat{\mathcal{X}}$ is of the form (6.6), and 1 otherwise. This error can be used to empirically measure the probability of success of Algorithm 6.1: we approximate it by running several experiments with different levels of noise σ and numbers of points K. In Figure 6.4, we report the results of such an experiment and compare it with our theoretical result obtained in (6.20). We confirm that $P(\sigma, K)$ exhibits a sharp phase transition—we can identify pairs (K, σ) for which the algorithm always succeeds and pairs for which it always fails. However, the empirical phase transition is less sharp than the theoretical one and this is probably due to our approximations regarding the independence of events. Nonetheless, the two phase transitions are closely aligned for each value of K.

These experimental results are rather surprising given the shortcuts we took in Section 6.6.2, and we would like to alleviate the astonishment by developing some intuition that may explain why these approximations appear to be so good. We claim that, even though not *all* events are pairwise independent, *most of them* are. As an example, when we look at

$$g(p, \widehat{\mathcal{X}}_k) = \sum_{\widehat{x} \in \widehat{\mathcal{X}}_k} \min_{\widetilde{d} \in \widetilde{\mathcal{D}}} \left(p - \widehat{x} - \widetilde{d} \right)^2 ,$$

for two different values p_1 and p_2 of p, the respective cost functions $g(p_1, \widehat{\mathcal{X}}_k)$ and $g(p_2, \widehat{\mathcal{X}}_k)$ probably share a few common differences $\widetilde{d}$. However, we observe that at round k, only k out of $K^2 - 2K + 1$ differences are selected, one for every $\widehat{x} \in \widehat{\mathcal{X}}_k$. Then, assuming that most pairs $\left(g(p_1, \widehat{\mathcal{X}}_k), g(p_2, \widehat{\mathcal{X}}_k) \right)$ are independent is practically equivalent to assuming that we select the differences $\widetilde{d}$ uniformly at random within the minimization. Moreover, we believe that the few dependent events ignored by such assumptions are one of the likely causes of the different steepness exhibited by the theoretical and observed phase transition.

6.7 Improving noise resilience

We now discuss strategies and variations of our support recovery algorithm aiming at improving the quality of the solution in noisy settings. We chose not to include them in the analysis as they make it more intricate.

6.7.1 Deleting solutions from the set of differences

When a new point $\widehat{\boldsymbol{x}}_{k+1}$ is added to $\widehat{\mathcal{X}}_k$, Algorithm 6.1 ignores some useful information. Assuming that there are no collisions and no noise, we know that the values $\widehat{\mathcal{X}}_k - \widehat{\boldsymbol{x}}_{k+1}$ and $\widehat{\boldsymbol{x}}_{k+1} - \widehat{\mathcal{X}}_k$ in $\mathcal{D}$ cannot belong to the solution $\widehat{\mathcal{X}}$ as they are part of $\mathcal{W}$. Thus, as soon as $\widehat{\boldsymbol{x}}_{k+1}$ is added to the solution set, we can remove all values of the form $\widehat{\mathcal{X}}_k - \widehat{\boldsymbol{x}}_{k+1}$ and $\widehat{\boldsymbol{x}}_{k+1} - \widehat{\mathcal{X}}_k$ from $\mathcal{D}$.

The same reasoning applies to the noisy case, but we pick the closest values in $\widetilde{\mathcal{D}}$ as we do not have exact differences. More formally, when we add a new point $\widehat{\boldsymbol{x}}_{k+1}$ to the solution $\widehat{\mathcal{X}}_k$, we dispose of the following $2k$ elements of $\widetilde{\mathcal{D}}$,

$$\widetilde{\boldsymbol{d}}^* = \operatorname*{arg\,min}_{\widetilde{\boldsymbol{d}} \in \widetilde{\mathcal{D}}} \| \pm \widehat{\boldsymbol{x}} \mp \widehat{\boldsymbol{x}}_{k+1} - \widetilde{\boldsymbol{d}} \|^2, \quad \forall \widehat{\boldsymbol{x}} \in \widehat{\mathcal{X}}_k.$$

This approach results in two opposing effects. On one hand, we introduce the risk of erroneously discarding a point $\widetilde{\boldsymbol{d}}^*$ that belongs to the solution. On the other hand, we are pruning many elements out of $\widetilde{\mathcal{D}}$ and naturally reduce the risk of picking an erroneous candidate later on in the recovery process. As we will show in Section 6.7.4, the benefits outweight the risks.

6.7.2 Symmetric cost function

Next, we replace the cost function (6.9) with a symmetric one to leverage the natural symmetry of the ACF.

In Algorithm 6.1, we search for the vectors in $\widetilde{\mathcal{D}}$ closest to the computed differences $\boldsymbol{p} - \widehat{\mathcal{X}}_k$ for each candidate $\boldsymbol{p}$. We strengthen its noise resilience by jointly searching for the vectors closest to $\mp\widehat{\mathcal{X}}_k \pm \boldsymbol{p}$ and choosing the candidate $\boldsymbol{p}$ that minimizes the sum of both errors. Specifically, we rewrite the cost function (6.9) as

$$\widehat{\boldsymbol{x}}_{k+1} = \operatorname*{arg\,min}_{\boldsymbol{p} \in \mathcal{P}_k} \sum_{\widehat{\boldsymbol{x}} \in \widehat{\mathcal{X}}_k} \min_{\widetilde{\boldsymbol{d}}, \widetilde{\boldsymbol{d}}' \in \widetilde{\mathcal{D}}} \left\| \boldsymbol{p} - \widehat{\boldsymbol{x}} - \widetilde{\boldsymbol{d}} \right\|^2 + \left\| \widehat{\boldsymbol{x}} - \boldsymbol{p} - \widetilde{\boldsymbol{d}}' \right\|^2.$$

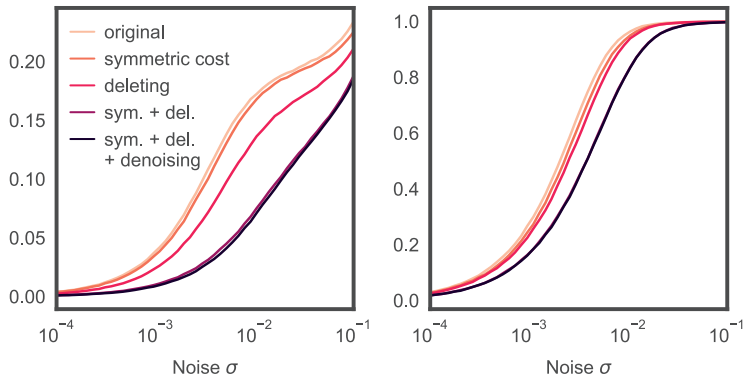
We stress that this improvement is compatible with the idea of caching introduced in Section 6.5. We can in fact cache the following pairs

$$(\widetilde{\boldsymbol{d}}, \widetilde{\boldsymbol{d}}')_{i,j} = \operatorname*{arg\,min}_{\widetilde{\boldsymbol{d}}, \widetilde{\boldsymbol{d}}' \in \widetilde{\mathcal{D}}} \|\boldsymbol{p}_j - \widehat{\boldsymbol{x}}_i - \widetilde{\boldsymbol{d}}\|^2 + \|\widehat{\boldsymbol{x}}_i - \boldsymbol{p}_j - \widetilde{\boldsymbol{d}}'\|^2,$$

for each $\widehat{\boldsymbol{x}}_i \in \widehat{\mathcal{X}}_k$ and $\boldsymbol{p}_j \in \mathcal{P}_k$ and recompute them when $\boldsymbol{p}_j$ gets added to the solution $\widehat{\mathcal{X}}_{k+1}$.

6.7.3 Denoising of partial solutions

At each iteration k of Algorithm 6.1, we have a partial solution $\widehat{\mathcal{X}}_{k+1}$ and, from (6.9), we identify, for each pair $\widehat{\boldsymbol{x}}_i, \widehat{\boldsymbol{x}}_j \in \widehat{\mathcal{X}}_{k+1}$, a difference $\widehat{\boldsymbol{d}}_{i,j}$ that is the closest to $\widehat{\boldsymbol{x}}_i - \widehat{\boldsymbol{x}}_j$. In other words, we



(a) ℓ_2 error (b) Index-based error

Figure 6.5: Average error for the different combinations of improvements of the algorithm. We create $\mathcal{X}$ from $K = 6$ 1D points chosen uniformly at random from the interval $[0, 1]$, create $\mathcal{D}$ accordingly and add Gaussian noise $\mathcal{N}(0, \sigma^2)$ to its elements. The ℓ^2 and the index-based errors are computed for different levels of noise σ and different improvements of the original algorithm.

are simultaneously labeling the differences $\widehat{\boldsymbol{d}}_{i,j}$ using our current partial solution; such a labeling is completed as k reaches the final iteration $K - 1$.

This partial labeling can be exploited to *denoise* the set $\widehat{\mathcal{X}}_{k+1}$ as it provides unused additional constraints and mitigates the error propagation between the iterations. More precisely, we propose to find a set of points $\{\widehat{\boldsymbol{x}}_i\}_{i=1}^{k+1}$ that minimizes the following cost function

$$J\left(\{\widehat{\boldsymbol{x}}_i\}_{i=1}^{k+1}\right) = \sum_{i,j} \|\widehat{\boldsymbol{d}}_{i,j} - (\widehat{\boldsymbol{x}}_i - \widehat{\boldsymbol{x}}_j)\|^2. \tag{6.21}$$

The solution to (6.21) is derived in closed-form by setting its first derivative to 0. As it is based on a least-square-error criterion, it is then optimal when the differences are corrupted by additive Gaussian noise.

This leads to a simple but effective strategy: refine the estimate of the solution set at each step by taking the average of the differences related to each point $\widehat{\boldsymbol{x}}_i \in \widehat{\mathcal{X}}_{k+1}$ as

$$\widehat{\boldsymbol{x}}_i = \frac{1}{k+1} \sum_{j=1}^{k+1} \widehat{\boldsymbol{d}}_{i,j},$$

where we recompute all $\widehat{\boldsymbol{x}}_i$ as they are used in the $k + 1$ iteration. To see why this works, we separate the sum as

$$\frac{1}{k+1} \sum_{j=1}^{k+1} \widehat{\boldsymbol{d}}_{i,j} = \boldsymbol{x}_i - \frac{1}{k+1} \sum_{j=1}^{k+1} \boldsymbol{x}_j + \frac{1}{k+1} \sum_{j=1}^{k+1} \boldsymbol{\nu}_{i,j}.$$

We observe that $-\frac{1}{k+1} \sum_{j=1}^{k+1} \boldsymbol{x}_j$ is the same translation for all points $\widehat{\boldsymbol{x}}_i$. The consequence of this approach is that the total noise is reduced as we average its different realizations over $k + 1$ values. Note that since Algorithm 6.1 assumes that $\widehat{\boldsymbol{x}}_1 = \boldsymbol{0}$ in $\widehat{\mathcal{X}}_k$, we also translate back all the points by $-\widehat{\boldsymbol{x}}_1$ after each update.

Unfortunately, the idea of caching the differences introduced in Section 6.5 is not compatible with the denoising of the partial solutions. Since at each step we modify the partial solution set $\widehat{\mathcal{X}}_k$, the differences between $\widehat{\mathcal{X}}_k$ and $\widetilde{\mathcal{D}}$ change accordingly, which makes it impossible to cache them. Hence, there exists a hard trade-off between quality and complexity, and we should pick the right strategy depending on the requirements of each specific practical scenario.

6.7.4 Comparison of improvement strategies

Last, we evaluate the significance of our proposed improvements on Algorithm 6.1. We quantify the results using the index-based error introduced in Section 6.6, as well as the ℓ^2 error, which we define as the ℓ^2-norm of the difference between the underlying points $\mathcal{X}$ and their estimation $\widehat{\mathcal{X}}$.[5]

The comparison of the different improvement strategies is illustrated in Figure 6.5. In this experiment, we draw $K = 6$ one-dimensional points uniformly at random from the interval $[0, 1]$ and add Gaussian noise $\mathcal{N}(0, \sigma^2)$ on their pairwise differences. We run Algorithm 6.1 and the proposed improvements for different noise levels σ. It is clear that all the proposed strategies enhance the original algorithm, with respect to both the index-based error and the ℓ^2 error.

Moreover, we also observe that different strategies combine constructively: for example, the symmetric cost function decreases the ℓ^2 error by 5% on average, while deleting solutions from the set of differences improves the results by 27% on average. When combined together, the average error decreases by 59%. Including the denoising further enhances the algorithm, as the average error decreases by 62%. Similarly, for the index-based error there is an evident shift between the phase transitions of the original algorithm with and without improvements.

6.8 Influence of the point locations

The algorithm performance around the phase transition in Figure 6.4 also seems to indicate that some configurations of points are easier to recover than others. In this section, we run a small experiment to visualize the challenges posed by certain configurations.

We consider a low-complexity setup ($K = 4$, and $D = 1$), fix the support boundaries, that is $x_1 = 0$ and $x_2 = 1$, and study the reconstruction error for various pairs $(x_3, x_4) \in [0, 1]^2$. We generate several instances of this problem and perturb the differences in $\mathcal{D}$ with additive Gaussian noise with zero mean and $\sigma = 0.01$. We measure the performance of Algorithm 6.1 (with all the improvements introduced in Section 6.7) using both the index-based and the ℓ^2 error. The average errors are then shown in Figure 6.6, where we observe that there exist some combinations of points that lead to a significantly higher error.

We now develop intuition about a few interesting cases that emerged from the previous experiment. For the sake of simplicity, we consider a noiseless setting where collisions in the ACF or non-uniqueness of the solution are the only causes of challenging configurations.

1. *Collision between a difference and a point.* When a difference and a point collide, it can happen that the difference is mistaken for the point. This does not influence the ℓ^2 error, but causes an index-based error. An example of such a case is when $x_3 = x_4$ (the main diagonal in Figure 6.6): both the difference $x_4 - x_3$ and x_1 have value 0. As a consequence,

[5] This requires to first align the two sets of points $\mathcal{X}$ and $\widehat{\mathcal{X}}$ by minimizing the ℓ^2-norm between their elements, subject to any shift and/or reflection.

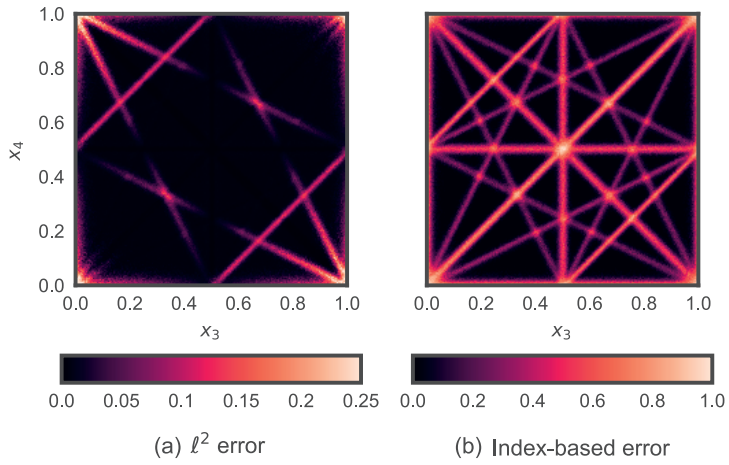


Figure 6.6: Influence of the points' locations on the estimation errors. We solve a 1D instance of the problem with $K = 4$, $x_1 = 0$, and $x_2 = 1$. The locations x_3 and x_4 vary along the x- and y-axis.

the sets $\mathcal{X}' = \{x_1, x_2, x_3, x_4\}$ and $\mathcal{X}'' = \{x_4 - x_3, x_2, x_3, x_4\}$ are both equal to $\mathcal{X} = \{0, 1, x_3, x_3\}$, but the latter is not of the form (6.6).

2. *Constant difference 0.5.* When $x_4 = x_3 \pm 0.5$, we can actually find more than one set of 4 points that map to a subset of the given differences. In the case $x_4 = x_3 + 0.5$, the differences are $\mathcal{D} = \pm\{0, 1, x_3, x_3 + 0.5, 1 - x_3, 0.5 - x_3, 0.5\}$; thus, $\mathcal{D}$ contains all pairwise difference from both $\mathcal{X}' = \{0, 1, x_3, x_3 + 0.5\}$ and $\mathcal{X}'' = \{0, 1, 0.5, x_3\}$. However, $\mathcal{X}''$ does not lead to a zero ℓ^2 error.

3. *Collision of differences when adding a new point to the solution set.* This is for example the case of $x_4 = 1 - 2x_3$ with $\mathcal{D} = \pm\{0, 1, x_3, 1 - 2x_3, 1 - x_3, 2x_3, 1 - 3x_3\}$. The differences 0 and 1 are always selected in the first and the second step. In the third step, we could potentially add $2x_3$ to $\mathcal{X}_2 = \{0, 1\}$ and reduce the set of differences to $\mathcal{D} = \pm\{x_3, 1 - x_3, 1 - 3x_3\}$. Next, we select x_3 as a new point. We can verify that the differences of x_3 and the values in $\mathcal{X}_3 = \{0, 1, 2x_3\}$ exist in $\mathcal{D}$. However, in this verification we use the value x_3 in $\mathcal{D}$ twice: once as the difference between x_3 and 0, and once as the difference between x_3 and $2x_3$. The set of pairwise differences of $\mathcal{X}_4 = \{0, 1, 2x_3, x_3\}$ is indeed contained in the original $\mathcal{D}$, but neither its ℓ^2 error nor its index-based error is zero. Notice that if we swap the third and the fourth step, this confusion would be avoided as x_3 would be removed from the set of differences in the third step.

These three cases explain all the segments visible in Figure 6.6. Such an analysis also applies to noisy regimes; the main difference is that we move from very localized configurations to blurrier areas where the solution is ambiguous. In fact, we introduced some noise into the experiment in Figure 6.6 to enable the visualization of the *lines* identifying challenging patterns—a noiseless setting would have just led to infinitesimally thin lines. Such patterns become blurrier and wider as noise increases. These areas where reconstruction is harder also explain the not-so-sharp phase transition in Figure 6.4: when drawing supports of K elements at random, the probability of

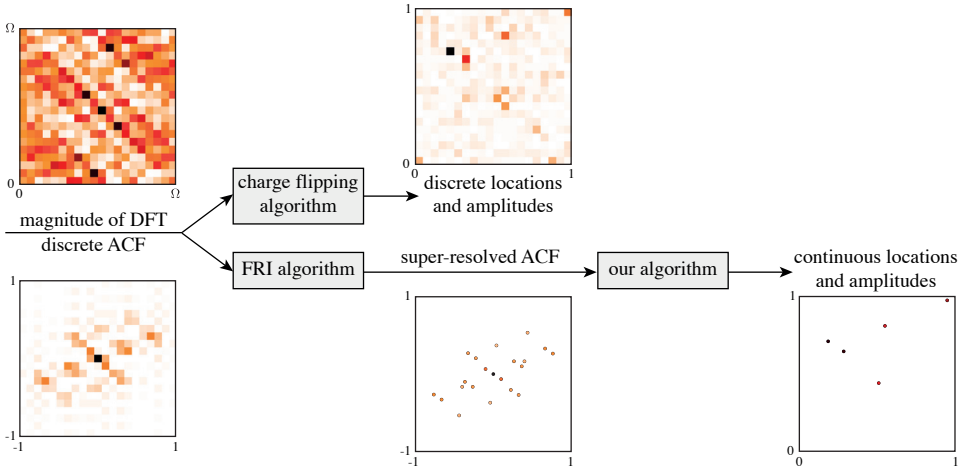


Figure 6.7: PR pipeline for Charge Flipping and our algorithm. First the signal $a(\boldsymbol{x})$ is sampled and we observe the magnitude of its DFT, $A_{\boldsymbol{m}}$, which also corresponds to a discrete version of its ACF. These DFT coefficients are directly used by Charge Flipping to recover a discretized support of $f(\boldsymbol{x})$. Our approach proceeds in two stages: first, using FRI we compute a super-resolved version of the ACF, and then by applying the proposed algorithm, we recover the continuous version of $f(\boldsymbol{x})$.

hitting a challenging pattern significantly grows with the noise. To the limit, these blurred lines cover the entire domain and the probability of success is null.

6.9 Comparison with Charge Flipping

In this section, we evaluate the performance of the proposed PR algorithm in comparison with other state-of-the-art methods. Recall that our algorithm is, to the best of our knowledge, the first to operate in a continuous-support setup, whereas other algorithms assume discrete signals. Indeed, the vast majority of PR methods are simply not designed to work with continuous supports; examples are PhaseLift [38], which recasts the PR problem as a semi-definite program, and GESPAR [168], which linearizes the PR problem with the damped Gauss-Newton method. In general, these approaches assume that the signals of interest are sparse vectors. As seen in Figure 6.1, when the locations are not aligned with the sampling grid, the discretized signal contains very few—if any—nonzero entries as the scattering function spreads the sharp continuous locations.

A few algorithms can be adapted to work with continuous supports, but they fall short when the measured support $\widetilde{\mathcal{D}}$ is noisy. This is the case of TSPR [84], which relies on the triangle equality between locations to recover the support; as soon as the locations are corrupted with noise, these equalities do not hold anymore.

The closest point of comparison to our method is the *Charge Flipping* algorithm [138, 139]; even though it operates in a discrete domain, our experiments have shown that it is resilient to some noise on the ACF support.

6.9.1 Charge Flipping

Charge Flipping is one of the reference algorithms in crystallography. It belongs to the class of dual-space algorithms as it alternatively acts on the spatial and Fourier domains, designated *real* and *reciprocal* space in crystallography. After randomly assigning a phase to the observed magnitudes of the discrete Fourier transform (DFT) coefficients, it iteratively performs the following two operations:

1. In the real space, it flips the sign of the values that are below some fixed threshold δ.
2. In the reciprocal space, it enforces that the magnitude of the signal corresponds to the measured magnitude.

Charge Flipping directly takes as input the DFT coefficients of the ACF, while our support recovery algorithm operates on a continuous version of the ACF. This is a significant advantage of our algorithm over Charge Flipping as we do not assume that the support of the points is aligned with a grid. To have an adequate comparison between the two, we need to consider the entire pipeline, combining the three algorithms exposed in Section 6.4; this is illustrated in Figure 6.7.

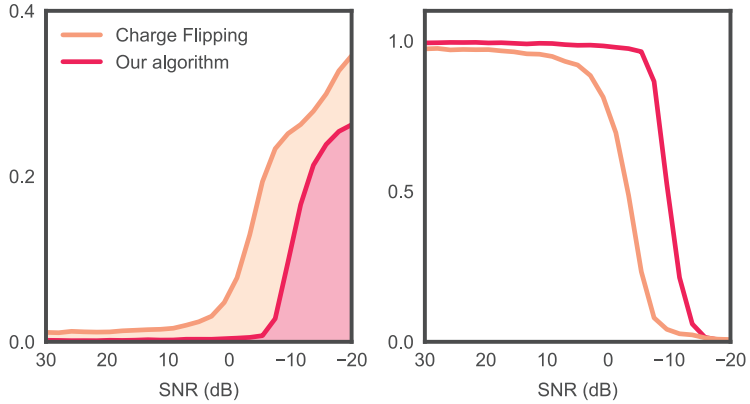
6.9.2 Experimental setup and results

We consider sparse signals of the form (6.1), where $\phi(\boldsymbol{x})$ is an ideal lowpass filter. Under this form, $f(\boldsymbol{x})$ corresponds to a stream of Diracs, which is the typical FRI signal; remark that other sampling kernels or different pulse shapes—such as the VPW pulses described in Chapter 2—could also be used. We then generate the corresponding DFT coefficients, discard their phase information, and corrupt them with zero-mean Gaussian noise. Notice that in Sections 6.6 and 6.7, we assume noise on the support of the points; here, we are dealing with noise that is applied to the DFT coefficients instead. Obviously, these noisy DFT coefficients also lead to a noisy support of the super-resolved ACF, but it is not Gaussian anymore. In fact, as FRI algorithms rely on nonlinear methods, the noise of its output is difficult to characterize.

The discretization of the Fourier domain is equivalent to a periodization of the spatial domain. As a consequence, the squared magnitude of the DFT coefficients corresponds to a *circular* ACF. While it is certainly possible to adapt Algorithm 6.1 to handle circular ACFs by testing all the possible 2^D quadrants for every observed $\widetilde{\boldsymbol{d}} \in \widetilde{\mathcal{D}}$, we chose to zero-pad the support of $f(\boldsymbol{x})$ until its ACF is not circular anymore.

Regarding Charge Flipping, we notice that its performance highly depends on the initial solution as well as the choice of the threshold δ. To avoid giving an unfair advantage to our algorithm, we run Charge Flipping 10 times for each experiment and pick the best solution; practical experiments show that the performance gain is marginal when going above such a number of repetitions. Furthermore, best practice [139] suggests to pick $\delta = b\theta$, where b is a constant around 1-1.2 and θ is the standard deviation of the measured signal. Our experiments showed that progressively decreasing the value of δ also improves the performance of Charge Flipping. This mimics the behavior of the simulated annealing algorithm, where the temperature is steadily decreased until convergence.

Then, given noisy DFT coefficients as input, we compare the ℓ^2 error on the support of the points for both algorithms, as well as a probability of successfully recovering the support. To define the latter, we say that an algorithm fails when the ℓ^2 error is higher than a specific threshold. Figure 6.8 shows that our FRI super-resolution algorithm surpasses Charge Flipping



(a) Average error (b) Probability of success

Figure 6.8: Comparison of our algorithm with Charge Flipping. The performance is evaluated for $K = 5$ 1D points chosen uniformly at random from $[0, 1]$. The number of DFT coefficients is 200. Figure (a) shows the ℓ^2 reconstruction error on the locations for different values of the input SNR. Figure (b) reports the percentage of success: we consider that the algorithms fail when the resulting ℓ^2 error is larger than some threshold 0.04.

in terms of both metrics. It is not surprising that our algorithm exhibits a superior performance in a low noise regime—it even achieves exact reconstruction in the absence of noise—since it is not bound to a grid. On the other hand, Charge Flipping always suffers from approximation errors due to the implicit discretization: in the noiseless case and for a grid of size 200, we calculate that the expected ℓ^2 error on the support of $K = 5$ points is about 0.0056, which is in adequacy with the baseline observed in Figure 6.8a. Simulations also indicate that our algorithm outperforms Charge Flipping in high noise environments. Indeed, the reconstruction error is consistently lower and its phase transition compares favorably as well.

6.10 Conclusion

We presented a novel approach to solve the phase retrieval problem for sparse signals. While conventional algorithms operate in discretized space and recover the support of the points on a grid, the power of FRI sampling combined with the sparsity assumption on the signal model enables us to recover the support of the points in continuous space. We provided a mathematical expression that approximates the probability of success of our support recovery algorithm and confirmed our result via numerical experiments. We observed that, while our algorithm runs in polynomial time with respect to the sparsity number of the signal, it remains relatively costly. To alleviate the computational costs without impacting the quality of the reconstruction, we introduced a caching layer to avoid repeating calculations. Furthermore, we proposed several improvements that contribute to enhance the quality of estimation in the presence of noise. Finally, we showed that our super resolution PR algorithm outperforms Charge Flipping, one of the state-of-the-art algorithms, both in terms of average reconstruction error and success rate.

Conclusion

Words can be meaningless. If they are used in such a way that no sharp conclusions can be drawn.

The Meaning of It All
Richard Feynman

We are reaching the end of our journey in the land of light, sampling and sparse signals. In this dissertation, we conceived new algorithms and theories and proposed applications revolving around the corpuscular, wave, and electromagnetic nature of light.

In Chapter 2, we presented variable pulse width-finite rate of innovation (VPW-FRI), a new parametric model to represent sparse signals having variable pulse width. While the classical FRI result is only applicable for streams of Diracs and pulses with fixed shapes, it has to be noted that most signals in nature consist of pulses with variable widths; such is the case of ECG signals, which are the ideal candidate for our pulse model. Along with the VPW model, we introduced sampling theorems building on the finite rate of innovation paradigm. These sampling theorems guarantee perfect reconstruction from only a limited number of samples in 1D, 2D and in the spherical domain.

In Chapter 3, our FRI-based sampling theorems were used to acquire BRDFs and in particular, specularities, which can present several challenges when using classical sampling theory. The sparse structure of specularities indeed fits perfectly within the finite rate of innovation framework introduced in the previous chapter. In addition to developing a sampling theory for BRDFs, we built light domes, which are specialized devices designed to acquire the reflectance function

of objects. These light domes are very versatile in the sense that they allow the acquisition of BRDFs with arbitrary resolutions. Furthermore, they enable new sampling schemes that were not discussed in this thesis: two examples that come to mind are sampling along trajectories and adaptive sampling.

In Chapter 4, we revived Lippmann photography and, with a signal processing perspective, presented a new model to explain and analyze the process. This allowed us to discover and describe new phenomena regarding the color reproduction of Lippmann photochromes; specifically, we showed that the spectra synthesized by a Lippmann plate are only a distorted version of the original spectra. We precisely characterized the distortions introduced in the process and, more importantly, proposed a way to eliminate these distortions and recover the original spectrum. We verified our theoretical results and algorithm with extensive experiments and through the creation of our own Lippmann plates.

Specularities were again the central theme in Chapter 5, where we exploited the polarization state of light to separate the diffuse and specular components. We investigated cameras with micro-filter arrays having different polarization orientations and proposed demosaicing schemes to retrieve the diffuse and specular components of a scene from the output of such cameras. We demonstrated that our approach outperformed single-image based techniques at that task. The separation of the diffuse and specular components is crucial to many applications; in this thesis, we showed that our method can significantly improve the estimation of the surface normals of a scene based on our reliable estimation of the diffuse component.

Finally, in Chapter 6, we presented a new algorithm to solve the phase retrieval problem under a sparsity assumption. Our approach diverges from traditional applications in the sense that it does not discretize the position of the atoms to be estimated; this provides a virtually unlimited spatial resolution. We proposed a theoretical analysis of our algorithm as well as numerous strategies to improve its resilience to noise. Lastly, we showed through numerical experiments that our algorithm outperforms charge flipping, the state-of-the-art in crystallography, in setups with varying amounts of noise.

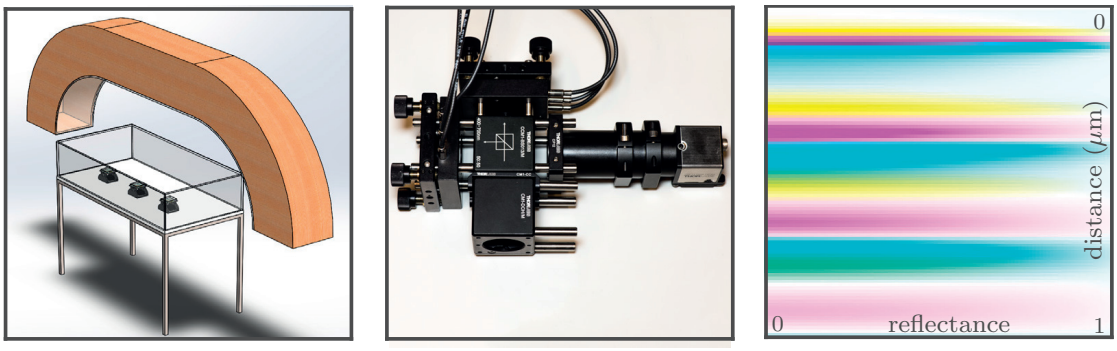
Outlook

After this brief summary, we now turn to the future and describe a few areas of ongoing work as well as propose topics of research to be further investigated. In particular, we chose to highlight a number of ideas related to Lippmann photography.

The historical Lippmann process

Our complete study of the historical Lippmann process enabled us to acquire a better understanding of these artworks. This knowledge can be applied to better showcase and underline the beauty of Lippmann plates. We distinguish three ways to promote and exhibit this photography technique: i) conceive display devices to enhance the visualization of existing plates, ii) create replicas of Lippmann plates to reach a larger audience, and iii) provide a digital experience of Lippmann photochromes through live rendering.

Visualization of historical Lippmann plates and exhibitions As mentioned in Chapter 4, our work on Lippmann was born from our interactions with the Elysée museum, which owns the largest collection of historical Lippmann plates in the world. Our collaboration was very fruitful;



(a) Lippmann display device (b) Interference digital camera (c) Interference display

Figure 6.9: Ideas for future work: (a) a Lippmann display device that consists of a moving platform to hold the photographic plates as well as a lighting equipment surrounding the platform; (b) prototype of an interference-based digital camera based on a Michelson interferometer; (c) interference-based display using two partially reflective mirrors in a Fabry-Pérot design which shows the colors we can reproduce by changing the distance between mirrors and their reflectance.

for instance, in winter 2017, we provided our expertise to design the lighting conditions to illuminate three Lippmann plates, displayed as part of the *Sans limite* exhibition[6]. More recently, we designed a display device in collaboration with Kjell Sverre Høyvik Bergum and Tarjei Nygaard Skulstad, as illustrated in Figure 6.9a. Such a device allows the public to visualize Lippmann plates under proper lighting and a moving platform enables us to witness the color changes when the plate is viewed under different angles. These devices will be used in a large exhibition[7] in 2021 to communicate with the general public about the science behind Lippmann photography. In the meantime, the equipment will be operated in our laboratory to showcase our research.

Making replicas of Lippmann plates Another area of research worth investigating relates to copying and printing Lippmann plates. Due to the extremely small size of the silver halide grains, it is generally believed that Lippmann plates are impossible to copy [30]. We argue that it is feasible. Exploiting our spectrum recovery algorithm from Chapter 4, we can precisely reconstruct the spectra of scenes captured in Lippmann plates. From these spectra, it is actually possible to create replicas of Lippmann plates; the idea is as follows. First, we acquire the full spectrum of a Lippmann plate using modern hyperspectral imaging techniques. Second, we retrieve the original spectra with Algorithm 4.1. Third, we need to somehow engrave or print diffraction patterns corresponding to the recovered spectrum. To that end, we propose the design of Figure 6.10 to reproduce a plate with the recovered spectrum, which is inspired from Newton's experiment in Figure 1.5. In this setup, we use a first prism to decompose white light into its spectrum, which is then collimated using a plano-convex lens. This combination of prism and lens can be interpreted as an analysis operation on the input light. The light spectrum can then be recombined into a light ray using a second pair of plano-convex lens and prism. This second arrangement is essentially a symmetry of the first arrangement and acts as a synthesis operator on the spectrum. As such, this design just acts as an identity operator on the light.

[6] http://www.elysee.ch/expositions-et-evenements/expositions/sans-limite/

[7] http://www.elysee.ch/en/exhibitions-and-events/events/detail/news/rising-to-the-challenge-digital-innovation-in-museums/

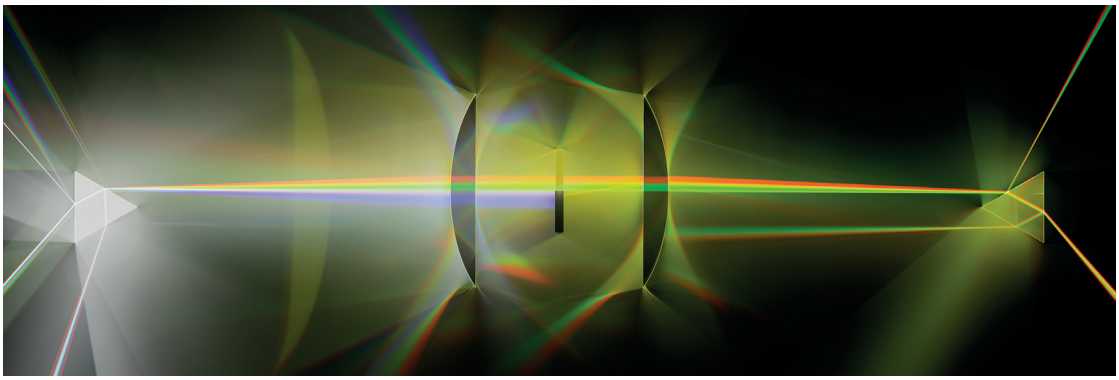


Figure 6.10: Replicating Lippmann plates: the first pair of prism and plano-convex lens acts as an analysis operator on the ray of white light. The decomposed spectrum can then be modulated using the SLM in the middle of the image. In this specific example, the SLM cuts off blue colors and preserves reds and greens. Finally, the modulated spectrum is recombined with a second pair of plano-convex lens and prism that plays the role of a synthesis operator. This enables us to create a ray of light with the spectral shape of our choice.

Things become more interesting when we place a spatial light modulator (SLM) [57] between the two plano-convex lenses. As their name suggests, SLMs spatially modulate the light to specific amplitudes. In our context, we can use the SLM to change the spectrum of a light ray to any arbitrary shape. This light ray can then be projected onto a photosensitive plate to create a 'pixel' with the desired color; a spatial array of such pixels combined together then recreates an image. Finally, the exposed plate is developed using standard Lippmann techniques. For a better efficiency, it is also possible to replace the light ray with a plane of light and use the second dimension of the SML to spatially modulate it. In this case, each plane of light corresponds to a full row and the combination of several rows produces the final image. An advantage of our proposed approach is that it is non-invasive in the sense that it does not require us to destroy the original plate to infer its interference pattern or equivalently the original spectrum; this makes it useable on historical Lippmann plates.

Rendering Another topic that we did not cover in this dissertation is that of digital rendering. Our focus was mainly on the acquisition pipeline, but the synthesis of images using the acquired data is also very relevant. As mentioned in Chapter 3, our laboratory has been the cradle of the startup Artmyn, which specializes in both the acquisition of BRDFs and rendering of paintings and artworks under dynamic lighting: the knowledge of the reflectance function of an artwork combined with live rendering enables the 3D exploration and re-illumination of digitized artworks. Live rendering was tackled using lightweight BRDF models, essentially trading off accuracy for speed.

We are particularly interested in providing a similar experience with Lippmann plates. We already experimented with our mathematical model in an off-line rendering framework and the results are convincing[8]; the next step is to streamline the process and provide instantaneous rendering for the visualization of digitized Lippmann plates.

[8]See for example https://go.epfl.ch/lippmann-rendering

Bringing Lippmann to the XXI century

Our study on Lippmann photography also raises the question of how to transfer this ancient photography technique to the XXI century. Nowadays, we have tools and technologies that are more modern and elaborated than those available to Lippmann when he devised his interference photography technique. Equipped with this additional knowledge, we propose displays and optical designs that share the main ingredients as those found in Lippmann photography.

New camera designs Lippmann photography suggests the comparison with existing hyperspectral camera designs. The goal of hyperspectral imaging is to capture the hyperspectral cube, made of 2 spatial dimension and 1 spectral dimensions. What is remarkable is that Lippmann photographs actually capture the 3 dimensions of the hyperspectral cube in a single take, a feat that is not achieved with modern hyperspectral cameras. Indeed, due to the planar nature of photosenstive sensors, modern cameras can only acquire 2 dimensions in a single take; the third dimension being captured by taking a series of shots. There are two main ways to obtain a hyperspectral cube: *spectral scanning*—where each image contain the entire spatial information but only one spectral band—and *spatial scanning*, or *push-broom*—where the individual images are made of the spectral dimension and only 1 spatial dimension. Spectral scanning is usually performed using (tunable) bandpass filters [66]. Inspired by Lippmann photography, we propose an interference-based spectral scanning camera design. In our laboratory, we recently constructed a prototype of such a camera based on a Michelson interferometer (see Figure 6.9b). The advantages of such a design is that it provides a direct representation of the two spatial dimensions of the scanned objects. Moreover, the number of bands is easily adjustable.

We can also think of a device capturing both polarization and spectral information along with 2 spatial dimensions, essentially borrowing ideas from Chapters 4 and 5.

Lippmann and interference-based displays The design described in Figure 6.10 could also be exploited to create hyperspectral displays; just like an electron gun in cathode ray tube (CRT) displays, we could use it to create 2D images with arbitrary spectra, one row at a time. This is essentially the dual of push-broom hyperspectral cameras.

Another approach to create interference-based displays is the following. Inspired by the design of Fabry-Pérot resonators [59], we propose to build a display where each pixel is made of two tiny stacked partially reflective mirrors. The distance between the two mirrors could be adjusted, driven by piezoelectric motors. Another parameter we can play with is the *reflectance* of the mirrors. In Figure 6.9c, numerical simulations show the range of colors we could reproduce by varying these two values. Perhaps not surprisingly, these colors closely match the ones one can observe in soap bubbles (see Figure 1.4b), in which the interface between the thin layer of soap and air plays the role of partially reflective mirrors. Even though the range of colors might seem limited at first glance, one could imagine fast moving mirrors coupled with temporal or spatial averaging to extend the range of reproducible colors, just like RGB-based displays. This display design has two notable advantages over conventional displays: a much lower energy consumption due to the absence of a backlight and a higher brightness in outdoor environments since it uses ambient light as its illumination source. We note that a very similar design is already in use in displays relying on the interferometric modulator (IMOD) technology [77, 118, 161]. Inspired by the structure of the wings of blue Morpho butterflies, IMOD displays also generate interferences with Fabry-Pérot resonators. The main difference with our proposal is that the distance between the two mirrors is fixed in such displays, essentially restricting every pixel to a single color—red,

green and blue—as in a traditional display. Spatial dithering is then exploited to create other colors by combining these three channels.

Printing interference patterns Finally, modern means can also be used to recreate colors from tailored interference gratings. In the previous section and Figure 6.10, we already proposed a way to shape light to a desired spectrum and imprint it into a plate. We now see other ways of creating interference gratings without the need for chemistry.

A known way of generating colors from interference is via diffraction grating dots [205]. Loosely speaking, diffraction grating dots consist of tiny bubbles which are printed inside a glass plate and contain regular patterns whose spacing is of the order of the wavelength of visible light. These regular patterns can be tuned to a specific frequency and reproduce monochromatic light. By combining several dots together, it is possible to reproduce a wide range of colors. One limitation of the technique is that it is restricted to regular patterns and thus, it cannot reproduce arbitrary spectra.

More recently, the emerging field of femtosecond laser engraving [41] can provide an alternative to diffraction grating dots. The principle behind femtosecond lasers is the concentration of a very large energy over a very short period of time. Usually, laser light passes through glass without modifying it, but if its energy is large enough, it triggers permanent modifications in the refractive index of the glass. These non-linear effects can be applied locally with a subwavelength precision and enable to print custom patterns and reproduce arbitrary spectra. Such a technique makes it possible to modulate the refractive index of a piece of glass in areas as small as 100-200 nm. We stress that we are talking about changing the refractive index, not the local reflectance of the material; this phenomenon can be studied using the model described in Section 4.6.

Closing remark This concludes this dissertation. Hopefully our modest contributions have stimulated your interest and will spark further connections between the fascinating worlds of light and signal processing. As Gabriel Lippmann declared in his Nobel Prize lecture: "*Life is short and progress is slow*" [104].

Bibliography

[1] 4D Technologies, "Polarization camera for image enhancement," https://www.4dtechnology.com/products/polarimeters/polarcam/, 2018.

[2] M. Aittala, T. Weyrich, and J. Lehtinen, "Practical SVBRDF capture in the frequency domain," *ACM Transactions on Graphics*, vol. 32, no. 4, pp. 110:1–110:12, 2013.

[3] F. Andersson, M. Carlsson, and M. V. de Hoop, "Sparse approximation of functions using sums of exponentials and AAK theory," *Journal of Approximation Theory*, vol. 163, no. 2, pp. 213 – 248, 2011.

[4] ——, "Frequency extrapolation through sparse sums of Lorentzians," *Journal of Earth Science*, vol. 25, no. 1, pp. 117–125, 2014.

[5] A. Artusi, F. Banterle, and D. Chetverikov, "A survey of specularity removal methods," in Proc. *Computer Graphics Forum*, vol. 30, no. 8, pp. 2208–2230. Wiley Online Library, 2011.

[6] L. A. Baboulaz, G. G. Soudan, and M. Vetterli, "Device for determining a bidirectional reflectance distribution function of a subject," U.S. Patent US14 559 726, 2014.

[7] G. Baechler, I. Dokmanić, L. Baboulaz, and M. Vetterli, "Accurate recovery of a specularity from a few samples of the reflectance function," in Proc. *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1596–1600, March 2016.

[8] G. Baechler, M. Kreković, J. Ranieri, A. Chebira, Y. M. Lu, and M. Vetterli, "Super resolution phase retrieval for sparse signals," *IEEE Transactions on Signal Processing*, 2018, manuscript submitted for publication.

[9] G. Baechler, A. Latty, M. Pacholska, M. Vetterli, and A. J. Scholefield, "Lippmann photography: A signal processing perspective," *IEEE Transactions on Signal Processing*, 2018, manuscript to be submitted for publication.

[10] G. Baechler, A. J. Scholefield, L. Baboulaz, and M. Vetterli, "Sampling and exact reconstruction of pulses with variable width," *IEEE Transactions on Signal Processing*, vol. 65, no. 10, pp. 2629–2644, May 2017.

[11] G. Baechler, "Sensing ECG signals with variable pulse width finite rate of innovation," Master's thesis, EPFL, 2012.

[12] G. Baechler, A. Latty, M. Pacholska, M. Vetterli, and A. J. Scholefield, "Shedding light on Lippmann photography," 2018, manuscript to be submitted for publication.

[13] R. Bagnold, *The Physics of Blown Sand and Desert Dunes*, Dover Earth Science. Dover Publications, 2005.

[14] H. Baines and E. Bomback, *The science of photography*. Fountain Press, 1967.

[15] R. Bajcsy, S. W. Lee, and A. Leonardis, "Detection of diffuse and specular interface reflections and inter-reflections by color image segmentation," *International Journal of Computer Vision*, vol. 17, no. 3, pp. 241–272, 1996.

[16] Y. Barbotin, "Parametric estimation of sparse channels theory and applications," Ph.D. dissertation, EPFL, Lausanne, 2014.

[17] Y. Barbotin and M. Vetterli, "Fast and robust parametric estimation of jointly sparse channels," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, no. 3, pp. 402–412, 2012.

[18] R. F. Bass and K. Gröchenig, "Random sampling of multivariate trigonometric polynomials," *SIAM Journal on Mathematical Analysis*, vol. 36, no. 3, pp. 773–795, 2005.

[19] E. Becquerel, "De l'image photographique colorée du spectre solaire," *Comptes Rendus Hebdomadaires à l'Académie Des Sciences*, vol. 26, pp. 181–184, 1848.

[20] ——, "De l'image photochromatique du spectre solaire," *Annales de chimie et de physique*, pp. 447–474, 1849.

[21] R. Beinert and G. Plonka, "Enforcing uniqueness in one-dimensional phase retrieval by additional signal information in time domain," *Applied and Computational Harmonic Analysis*, 2017.

[22] T. Bendory, R. Beinert, and Y. C. Eldar, "Fourier phase retrieval: Uniqueness and algorithms," in *Compressed Sensing and its Applications*. Springer, 2017, pp. 55–91.

[23] J. Berent, P. L. Dragotti, and T. Blu, "Sampling piecewise sinusoidal signals with finite rate of innovation methods," *IEEE Transactions on Signal Processing*, vol. 58, no. 2, pp. 613–625, 2010.

[24] G. Beylkin and L. Monzón, "Nonlinear inversion of a band-limited Fourier transform," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 351–366, 2009.

[25] B. N. Bhaskar, G. Tang, and B. Recht, "Atomic norm denoising with applications to line spectral estimation," *IEEE Transactions on Signal Processing*, vol. 23, no. 61, pp. 5987–5999, 2013.

[26] M. S. Billah, T. B. Mahmud, F. S. Snigdha, and M. A. Arafat, "A novel method to model ECG beats using Gaussian functions," in Proc. *4th International Conference on Biomedical Engineering and Informatics*, vol. 2, pp. 612–616, 2011.

[27] H. I. Bjelkhagen, "Lippman photographs recorded in DuPont color photopolymer material," in Proc. *Practical Holography XI and Holographic Materials III*, vol. 3011, pp. 358–367. International Society for Optics and Photonics, 1997.

[28] ——, "Super-realistic-looking images based on colour holography and Lippmann photography," *The International Society for Optical Engineering*, 2002.

[29] ——, "Lippmann photography: its history and recent development," *The PhotoHistorian*, pp. 11–19, 2003.

[30] H. I. Bjelkhagen and D. Brotherton-Ratcliffe, *Ultra-Realistic Imaging: Advanced Techniques in Analogue and Digital Colour Holography*. CRC Press, 2013.

[31] H. I. Bjelkhagen and D. P. Green, "The true colour of photography," unpublished work.

[32] H. Bjelkhagen, *Silver-halide Recording Materials: For Holography and Their Processing*, Springer series in optical sciences. Springer-Verlag, 1993.

[33] Blender Online Community, "Blender—A 3D modelling and rendering package," http://www.blender.org, Blender Foundation, Blender Institute, Amsterdam, 2018.

[34] T. Blu and M. Unser, "Quantitative Fourier analysis of approximation techniques. I. Interpolators and projectors," *IEEE Transactions on Signal Processing*, vol. 47, no. 10, pp. 2783–2795, Oct 1999.

[35] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.

[36] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, no. 5, pp. 1081–1089, 1986.

[37] J. A. Cadzow, "Signal enhancement—A composite property mapping algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 1, pp. 49–62, 1988.

[38] E. J. Candes, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," *SIAM review*, vol. 57, no. 2, pp. 225–251, 2015.

[39] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Communications on Pure and Applied Mathematics*, vol. 67, no. 6, pp. 906–956, 2014.

[40] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The convex geometry of linear inverse problems." *Foundations of Computational Mathematics*, vol. 12, no. 6, pp. 805–849, 2012.

[41] F. Chen and J. R. V. de Aldana, "Optical waveguides in crystalline dielectric materials produced by femtosecond-laser micromachining," *Laser & Photonics Reviews*, vol. 8, no. 2, pp. 251–275, 2014.

[42] W. Cochran, "Relations between the phases of structure factors," *Acta Crystallographica*, vol. 8, no. 8, pp. 473–478, Aug. 1955.

[43] P. Connes, "Silver salts and standing waves: the history of interference colour photography," *Journal of Optics*, vol. 18, no. 4, p. 147, 1987.

[44] R. L. Cook and K. E. Torrance, "A reflectance model for computer graphics," *ACM SIGGRAPH Computer Graphics*, vol. 15, no. 3, pp. 307–316, 1981.

[45] G. A. Croes, "A method for solving traveling-salesman problems," *Operations research*, vol. 6, no. 6, pp. 791–812, 1958.

[46] Z. Cui, J. Gu, B. Shi, P. Tan, and J. Kautz, "Polarimetric multi-view stereo," in Proc. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1558–1567, 2017.

[47] T. Dakić, "On the turnpike problem," Ph.D. dissertation, Simon Fraser University BC, Canada, 2000.

[48] P. Debevec, "The Light Stages and Their Applications to Photoreal Digital Actors," in Proc. *SIGGRAPH Asia*, Singapore, Nov 2012.

[49] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar, "Acquiring the reflectance field of a human face," *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 145–156, 2000.

[50] R. Descartes, *Discours de la Méthode: La Dioptrique.* Imprimerie Ian Maire, 1637.

[51] S. Deslauriers-Gauthier and P. Marziliano, "Sampling signals with a finite rate of innovation on the sphere," *IEEE Transactions on Signal Processing*, vol. 61, no. 18, pp. 4552–4561, 2013.

[52] I. Dokmanić and Y. M. Lu, "Sampling sparse signals on the sphere: Algorithms and applications," *IEEE Transactions on Signal Processing*, vol. 64, no. 1, pp. 189–202, Jan 2016.

[53] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[54] P. L. Dragotti and F. Homann, "Sampling signals with finite rate of innovation in the presence of noise," in Proc. *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 2941–2944, April 2009.

[55] P. L. Dragotti, M. Vetterli, and T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 1741–1757, 2007.

[56] J. R. Driscoll and D. M. Healy, "Computing Fourier transforms and convolutions on the 2-sphere," *Advances in applied mathematics*, vol. 15, no. 2, pp. 202–250, 1994.

[57] U. Efron, *Spatial light modulator technology: materials, devices, and applications.* CRC press, 1994, vol. 47.

[58] A. Einstein, "Concerning an heuristic point of view toward the emission and transformation of light," *Annalen der Physik*, vol. 17, pp. 132–148, 1905.

[59] C. Fabry and A. Pérot, "On a new form of interferometer," *The Astrophysical Journal*, vol. 13, p. 265, 1901.

[60] H. G. Feichtinger and K. Gröchenig, "Theory and practice of irregular sampling," *Wavelets: mathematics and applications*, pp. 305–363, 1994.

[61] FluxData, "Polarization imaging camera FD-1665P," http://www.fluxdata.com/products/fd-1665p, 2018.

[62] D. A. Forsyth and J. Ponce, *Computer vision: a modern approach*. Prentice Hall, 2003.

[63] J.-M. R. Fournier, "Investigation on lippmann photographs: materials, processes, and color rendition," in Proc. *SPIE 2176, Practical Holography VIII*, pp. 144–152, 1994.

[64] J.-M. R. Fournier, R. Benjamin, R. Alexander, L. Paul, L. Burnett, E. Sarah, and E. Stamper, "Recent developments in Lippmann photography," in Proc. *SPIE 3358, Sixth International Symposium on Display Holography*, pp. 95–102, 1998.

[65] A. Gardner, C. Tchou, T. Hawkins, and P. Debevec, "Linear light source reflectometry," *ACM Transactions on Graphics*, vol. 22, no. 3, pp. 749–758, 2003.

[66] N. Gat, "Imaging spectroscopy using tunable filters: a review," in Proc. *Wavelet Applications VII*, vol. 4056, pp. 50–65. International Society for Optics and Photonics, 2000.

[67] Y. Gentet and P. Gentet, "Ultimate emulsion and its applications: a laboratory-made silver halide emulsion of optimized quality for monochromatic pulsed and full-color holography," *Proceedings of SPIE*, vol. 4149, pp. 56–62, 2000.

[68] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of the phase from image and diffraction plane pictures," *Optik*, vol. 35, p. 237, 1972.

[69] R. Gershon, *The Use of Color in Computational Vision*, Technical report. Department of Computer Science, University of Toronto, 1987.

[70] P. Glafkidès, "Chimie et physique photographiques," 1976.

[71] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[72] T. Goldstein and S. Osher, "The split Bregman method for L1-regularized problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, apr 2009.

[73] R. A. Gonsalves, "Perspectives on phase retrieval and phase diversity in astronomy," in Proc. *Adaptive Optics Systems IV*, vol. 9148, pp. 1–10. International Society for Optics and Photonics, 2014.

[74] J. H. Halton, "On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals," *Numerische Mathematik*, vol. 2, no. 1, pp. 84–90, 1960.

[75] P. Hariharan and P. Hariharan, *Optical Holography: Principles, techniques and applications*. Cambridge University Press, 1996.

[76] J. F. W. Herschel, "On the chemical action of rays of the solar spectrum on preparations of silver and other substances, both metallic and non-metallic, and on some photographic processes," *Philosophical Transactions of the Royal Society of London*, vol. 130, pp. 1–59, 1840.

[77] J. Hong, E. Chan, T. Chang, T.-C. Fung, B. Hong, C. Kim, J. Ma, Y. Pan, R. Van Lier, S.-g. Wang, *et al.*, "Continuous color reflective displays using interferometric absorption," *Optica*, vol. 2, no. 7, pp. 589–597, 2015.

[78] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 5, pp. 814–824, 1990.

[79] F. Hurter and V. C. Driffield, "Photochemical investigations and a new method of determination of the sensitiveness of photographic plates," *Journal of the Society of the Chemical Industry*, vol. 9, p. 76, 1890.

[80] C. Huygens, *Traité de la lumiere*. Pieter van der Aa, 1690.

[81] H. Ives, "An experimental study of the Lippmann color photograph," Ph.D. dissertation, Johns Hopkins University, 1908.

[82] H. E. Ives, "Lippmann color photographs as sources of monochromatic light in photometry and optical pyrometry," *Journal of the Optical Society of America*, vol. 1, no. 2, pp. 49–63, Mar 1917.

[83] K. Jaganathan, Y. C. Eldar, and B. Hassibi, "Phase retrieval: An overview of recent developments," *arXiv preprint arXiv:1510.07713*, 2015.

[84] K. Jaganathan, S. Oymak, and B. Hassibi, "Sparse phase retrieval: Uniqueness guarantees and recovery algorithms," *IEEE Transactions on Signal Processing*, vol. 65, no. 9, pp. 2402–2410, may 2017.

[85] R. Jahr, "Über die Herstellung sogenannter "kornloser" Platten für des Interferenz-Farben-Verfahren von Professor Lippmann nach Dr. Hans Lehmann," *Photographische Industrie*, vol. 23, pp. 1013–1014, 1925.

[86] A. Jaklič and F. Solina, "Separating diffuse and specular component of image irradiance by translating a camera," in Proc. *International Conference on Computer Analysis of Images and Patterns*, pp. 428–435. Springer, 1993.

[87] T. H. James and C. E. K. Mees, *The theory of the photographic process*. Macmillan New York, 1966, vol. 977.

[88] L. V. Jospin, G. Baechler, and A. J. Scholefield, "Embedded polarizing filter design to separate diffuse and specular reflection," in Proc. *Asian Conference on Computer Vision*, 2018, to appear.

[89] J. Karle and H. Hauptman, "The phases and magnitudes of the structure factors," *Acta Crystallographica*, vol. 3, pp. 181–187, July 1950.

[90] D. W. Kim, S. Lin, K.-S. Hong, and H.-Y. Shum, "Variational specular separation using color and polarization." in Proc. *MVA*, pp. 176–179, 2002.

[91] W. Kim and M. H. Hayes, "The phase retrieval problem in X-ray crystallography," in Proc. *International Conference on Acoustics, Speech, and Signal Processing*, pp. 1765–1768 vol.3, Apr 1991.

[92] G. J. Klinker, S. A. Shafer, and T. Kanade, "The measurement of highlights in color images," *International Journal of Computer Vision*, vol. 2, no. 1, pp. 7–32, 1988.

[93] ——, "A physical approach to color image understanding," *International Journal of Computer Vision*, vol. 4, no. 1, pp. 7–38, 1990.

[94] K. T. Knox, "Image retrieval from astronomical speckle patterns," *Journal of the Optical Society of America*, vol. 66, no. 11, pp. 1236–1239, 1976.

[95] G. Kopp and J. L. Lean, "A new, lower value of total solar irradiance: Evidence and climate significance," *Geophysical Research Letters*, vol. 38, no. 1, 2011.

[96] V. A. Kotelnikov, "On the carrying capacity of the ether and wire in telecommunications," *Material for the First All-Union Conference on Questions of Communication, Izd. Red. Upr. Svyazi RKKA*, vol. 35, p. 181–194, 1933.

[97] E. E. Kummer, "De integralibus quibusdam definitis et seriebus infinitis." *Journal für die reine und angewandte Mathematik*, vol. 17, pp. 228–242, 1837.

[98] B. Lamond, P. Peers, A. Ghosh, and P. Debevec, "Image-based separation of diffuse and specular reflections using environmental structured illumination," in Proc. *IEEE Computational Photography*, 2009.

[99] C. L. Lawson and R. J. Hanson, *Solving least squares problems*. SIAM, 1995, vol. 15.

[100] H. Lehmann, *Beiträge zur Theorie und Praxis der direkten Farbenphotographie mittels stehender Lichtwellen nach Lippmanns Methode*. Trömer, 1906.

[101] S. Lin, Y. Li, S. B. Kang, X. Tong, and H.-Y. Shum, "Diffuse-specular separation and depth recovery from image sequences," in Proc. *European Conference on Computer Vision*, pp. 210–224. Springer, 2002.

[102] G. Lippmann, "La photographie des couleurs," *Comptes Rendus Hebdomadaires à l'Académie Des Sciences*, vol. 112, pp. 274–275, 1891.

[103] ——, "Sur la théorie de la photographie des couleurs simples et composées par la méthode interférentielle," *Journal de Physique Théorique Appliquée*, vol. 3, no. 1, pp. 97–107, 1894.

[104] ——, "Nobel lecture: Colour photography, december 14, 1908," *Nobel Lectures, Physics 1901–1921*, 1967.

[105] A. Lumière and L. Lumière, "Note sur la photographie des couleurs," *Revue suisse de photographie*, vol. 5, pp. 150–153, 1893.

[106] ——, "Sur les procédés pour la photographie des couleurs d'après la méthode de M. Lippmann," *Bulletin de la Société française de Photographie*, vol. 9, pp. 249–251, 1893.

[107] LuxCoreRender project, "Luxcorerender—An open source physically based renderer," https://luxcorerender.org, 2018.

[108] W. Ma, T. Hawkins, and P. Peers, "Rapid acquisition of specular and diffuse normal maps from polarized spherical gradient illumination," *Proceedings of the 18th Eurographics conference on Rendering Techniques*, pp. 183–194, 2007.

[109] S. P. Mallick, T. Zickler, P. Belhumeur, and D. Kriegman, "Dichromatic separation: specularity removal and editing," in Proc. *ACM SIGGRAPH 2006 Sketches*, p. 166, 2006.

[110] S. P. Mallick, T. Zickler, P. N. Belhumeur, and D. J. Kriegman, "Specularity removal in images and videos: A PDE approach," in Proc. *European Conference on Computer Vision*, pp. 550–563. Springer, 2006.

[111] S. P. Mallick, T. E. Zickler, D. J. Kriegman, and P. N. Belhumeur, "Beyond Lambert: Reconstructing specular surfaces using color," in Proc. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 619–626, 2005.

[112] T. Malzbender, D. Gelb, and H. Wolters, "Polynomial texture maps," in Proc. *28th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 519–528. ACM, 2001.

[113] I. Maravić and M. Vetterli, "Exact sampling results for some classes of parametric nonbandlimited 2-D signals," *IEEE Transactions on Signal Processing*, vol. 52, no. 1, pp. 175–189, 2004.

[114] A. Marraud and J.-M. Fournier, "Formation d'image accompagnée d'une analyse spectrale en champ complet: de J. Fourier à G. Lippmann," *Microscopy Microanalysis Microstructures*, vol. 8, no. 1, pp. 37–39, 1997.

[115] W. Matusik, H. Pfister, M. Brand, and L. McMillan, "A data-driven reflectance model," *ACM Transactions on Graphics*, vol. 22, p. 759, 2003.

[116] J. C. Maxwell, "A dynamical theory of the electromagnetic field," *Philosophical Transactions of the Royal Society of London*, vol. 26, pp. 459–512, 1865.

[117] L. Meng, L. Lu, N. Bedard, and K. Berkner, "Single-shot specular surface reconstruction with gonio-plenoptic imaging," in Proc. *IEEE International Conference on Computer Vision*, pp. 3433–3441, 2015.

[118] M. W. Miles, "A new reflective FPD technology using interferometric modulation," *Journal of the Society for Information Display*, vol. 5, no. 4, pp. 379–382, 1997.

[119] R. P. Millane, "Phase retrieval in crystallography and optics," *Journal of the Optical Society of America A*, vol. 7, no. 3, pp. 394–411, Mar 1990.

[120] D. J. Mitchell, "Reflecting nature: chemistry and comprehensibility in Gabriel Lippmann's 'physical' method of photographing colours," *Notes and Records of the Royal Society*, vol. 64, no. 4, pp. 319–337, 2010.

[121] E. Moiseev, A. Prudnikov, and U. Skórnik, "On the generalized Fourier sine- and cosine-transforms," *Integral Transforms and Special Functions*, vol. 7, no. 1-2, pp. 163–166, 1998.

[122] A. Nair and P. Marziliano, "Noisy channel detection using the common annihilator with an application to electrocardiograms," in Proc. *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 5972–5976, 2015.

[123] A. Nair, P. Marziliano, R. F. Quick, R. E. Crochiere, and G. Baechler, "Multichannel ECG analysis using VPW-FRI," in Proc. *International Conference on Sampling Theory and Applications*, pp. 125–128, 2013.

[124] V. T. Namy, "Computer-controlled light dome," 2014, semester project report.

[125] H. Nareid, "A review of the Lippmann colour process," *The Journal of Photographic Science*, vol. 36, no. 4, pp. 140–147, 1988.

[126] H. Nareid and H. M. Pedersen, "Modeling of the Lippmann color process," *Journal of the Optical Society of America A*, vol. 8, no. 2, pp. 257–265, 1991.

[127] S. K. Nayar, X.-S. Fang, and T. Boult, "Removal of specularities using color and polarization," in Proc. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 583–590, 1993.

[128] ——, "Separation of reflection components using color and polarization," *International Journal of Computer Vision*, vol. 21, no. 3, pp. 163–186, 1997.

[129] S. K. Nayar, K. Ikeuchi, and T. Kanade, "Determining shape and reflectance of hybrid surfaces by photometric sampling," *IEEE Transactions on Robotics and Automation*, vol. 6, no. 4, pp. 418–431, 1990.

[130] S. K. Nayar, G. Krishnan, M. D. Grossberg, and R. Raskar, "Fast separation of direct and global components of a scene using high frequency illumination," in Proc. *ACM Transactions on Graphics*, vol. 25, no. 3, pp. 935–944, 2006.

[131] R. Neuhauss, *Die Farbenphotographie nach Lippmann's Verfahren: neue Untersuchungen und Ergebnisse*. Knapp, 1898.

[132] I. Newton, *Opticks or A Treatise of the Reflections, Refractions, Inflections and Colours of Light*. Royal Society, London, 1703.

[133] ——, "A letter of Mr. Isaac Newton, professor of the mathematicks in the University of Cambridge; containing his new theory about light and colors: Sent by the Author to the Publisher from Cambridge, Febr. 6. 1671/72; In Order to be Communicated to the Royal Society," *Philosophical Transactions*, vol. 6, no. 69-80, pp. 3075–3087, 1972.

[134] A. Ngan, F. Durand, and W. Matusik, "Experimental analysis of BRDF models," *Proceedings of the Eurographics Symposium on Rendering*, pp. 117–126, 2005.

[135] F. E. Nicodemus, "Directional reflectance and emissivity of an opaque surface," *Applied Optics*, vol. 4, no. 7, pp. 767–775, Jul 1965.

[136] H. Nyquist, "Certain topics in telegraph transmission theory," *Transactions of the American Institute of Electrical Engineers*, vol. 47, no. 2, pp. 617–644, April 1928.

[137] J. Oñativia, S. R. Schultz, and P. L. Dragotti, "A finite rate of innovation algorithm for fast and accurate spike detection from two-photon calcium imaging," *Journal of Neural Engineering*, vol. 10, no. 4, pp. 1–30, 2013.

[138] G. Oszlányi and A. Süto, "Ab initio structure solution by charge flipping," *Acta Crystallographica Section A: Foundations of Crystallography*, vol. 60, no. 2, pp. 134–141, mar 2004.

[139] ——, "The charge flipping algorithm," *Acta Crystallographica Section A: Foundations of Crystallography*, vol. 64, no. 1, pp. 123–134, 2008.

[140] M. Ozisik, *Boundary Value Problems of Heat Conduction*, Dover Books on Engineering. Dover Publications, 2013.

[141] H. Pan, T. Blu, and M. Vetterli, "Efficient multidimensional Diracs estimation with linear sample complexity," *IEEE Transactions on Signal Processing*, vol. 66, no. 17, pp. 4642–4656, Sept 2018.

[142] ——, "Towards generalized FRI sampling with an application to source resolution in radioastronomy," *IEEE Transactions on Signal Processing*, vol. 65, no. 4, pp. 821–835, 2017.

[143] M.-A. Parseval, "Mémoire sur les séries et sur l'intégration complète d'une équation aux différences partielles linéaires du second ordre, à coefficients constants," *Mémoires présentés à l'Institut des Sciences, Lettres et Arts, par divers savants, et lus dans ses assemblées. Sciences, mathématiques et physiques. (Savants étrangers.)*, vol. 1, pp. 638–648, 1806.

[144] A. L. Patterson, "A direct method for the determination of the components of interatomic distances in crystals," *Zeitschrift für Kristallographie-Crystalline Materials*, vol. 90, no. 1-6, pp. 517–542, 1935.

[145] J. Peatross and M. Ware, *Physics of Light and Optics*. Optical Society of America, 2015.

[146] M. Pharr, W. Jakob, and G. Humphreys, *Physically Based Rendering: From Theory to Implementation*. Morgan Kaufmann, 2016.

[147] N. J. Phillips, H. Heyworth, and T. Hare, "On Lippmann's photography," *Journal of Photographic Science*, vol. 32, no. 5, pp. 158–169, 1984.

[148] N. Phillips and R. V. der Werf, "The creation of effective Lippmann layers in ultra-fine grain silver halide materials using non-laser sources," *Journal of Photographic Science*, vol. 33, no. 5, pp. 22–28, 1985.

[149] Photonic Lattice, "Polarization imaging camera PI-110," https://www.photonic-lattice.com/en/products/polarization_camera/pi-110/, 2018.

[150] V. F. Pisarenko, "The retrieval of harmonies from a covariance function," *Geophysical Journal of the Royal astrophysical Society*, vol. 33, pp. 347–366, 1973.

[151] M. Ponsot, "Photographie interférentielle; variation de l'incidence; lumière polarisée," *Comptes Rendus Hebdomadaires à l'Académie Des Sciences*, vol. 142, pp. 1506–1511, 1906.

[152] R. F. Quick, R. E. Crochiere, J. H. Hong, A. Hormati, and G. Baechler, "Extension of FRI for modeling of electrocardiogram signals," in Proc. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 2909–2912, 2012.

[153] B. Rafaely, *Fundamentals of Spherical Array Processing*. Springer-Verlag Berlin Heidelberg, 2015.

[154] S. Ramón y Cajal, "Chromomicrophotographie par la méthode interférentielle," *La Photographie des Couleurs*, vol. 2, no. 7, pp. 97–101, 1907.

[155] J. Ranieri, A. Chebira, Y. M. Lu, and M. Vetterli, "Phase retrieval for sparse signals: Uniqueness conditions," *arXiv preprint arXiv:1308.3058*, 2013.

[156] Ricoh Imaging Company Ltd, "Polarization camera," https://www.ricoh.com/technology/tech/051_polarization.html, 2018.

[157] E. Rothé, "Photographies en couleurs obtenues par la méthode interférentielle sans miroir de mercure," *Comptes Rendus Hebdomadaires à l'Académie Des Sciences*, vol. 139, pp. 565–567, 1904.

[158] R. Roy and T. Kailath, "ESPRIT—Estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984–995, 1989.

[159] S. M. Rusinkiewicz, "A new change of variables for efficient BRDF representation," *Rendering Techniques, Proceedings of Eurographics Rendering Workshop*, pp. 11–22, 1998.

[160] B. E. A. Saleh and M. C. Teich, *Fundamentals of photonics*. Wiley New York, 1991, vol. 22.

[161] J. B. Sampsell, "MEMS-based display technology drives next-generation FPDs for mobile applications," *Information Display*, vol. 22, no. 6, p. 24, 2006.

[162] G. Saxby and S. Zacharovas, *Practical Holography, Fourth Edition*. CRC Press, 2015.

[163] D. Sayre, "The squaring method: a new method for phase determination," *Acta Crystallographica*, vol. 5, no. 1, pp. 60–65, Jan. 1952.

[164] S. A. Shafer, "Using color to separate reflection components," *Color Research & Application*, vol. 10, no. 4, pp. 210–218, 1985.

[165] C. E. Shannon, "Communication in the presence of noise," *Proc. Institute of Radio Engineers*, vol. 37, no. 1, pp. 10–21, 1949.

[166] ——, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 623–656, 1948.

[167] K. K. Sharma, *Optics: principles and applications*. Academic Press, 2006.

[168] Y. Shechtman, A. Beck, and Y. C. Eldar, "GESPAR: efficient phase retrieval of sparse signals," *IEEE Transactions on Signal Processing*, vol. 62, no. 4, pp. 928–938, 2014.

[169] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: a contemporary overview," *IEEE signal processing magazine*, vol. 32, no. 3, pp. 87–109, 2015.

[170] H.-L. Shen and Z.-H. Zheng, "Real-time highlight removal using intensity ratio," *Applied optics*, vol. 52, no. 19, pp. 4483–4493, 2013.

[171] S. S. Skiena, W. D. Smith, and P. Lemke, "Reconstructing sets from interpoint distances," in Proc. *Sixth annual symposium on Computational geometry*, pp. 332–339. ACM, 1990.

[172] S. S. Skiena and G. Sundaram, "A partial digest approach to restriction site mapping," *Bulletin of Mathematical Biology*, vol. 56, no. 2, pp. 275–294, 1994.

[173] H. M. Smith, *Holographic recording materials*. Springer Science & Business Media, 2006, vol. 20.

[174] I. M. Sobol, "The distribution of points in a cube and the approximate evaluation of integrals, zh. vychisl. mat. i mat. fiz. 7 (1967), 784–802," *USSR Computational Mathematics and Mathematical Physics*, vol. 7, no. 4, pp. 86–112, 1967.

[175] G. Soudan, "Design and building of an illumination prototype," 2014, semester project report.

[176] "The Stanford 3D Scanning Repository," http://graphics.stanford.edu/data/3Dscanrep/.

[177] P. Stoica and R. Moses, *Spectral Analysis of Signals*. Pearson Prentice Hall, 2005.

[178] G. Strang and G. Fix, "A Fourier analysis of the finite element variational method," *Constructive Aspects of Functional Analysis*, pp. 795–840, 1973.

[179] R. T. Tan and K. Ikeuchi, "Separating reflection components of textured surfaces using a single image," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 178–193, 2005.

[180] R. T. Tan, K. Nishino, and K. Ikeuchi, "Separating reflection components based on chromaticity and noise analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 10, pp. 1373–1379, 2004.

[181] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7465–7490, Nov 2013.

[182] S. Toxqui-López, A. Olivares-Pérez, B. Pinto-Iguanero, A. Aguilar-Mora, and I. Fuentes-Tapia, "PVA glue as a recording holographic medium," in Proc. *Practical Holography XXVI: Materials and Applications*, vol. 8281, p. 828102. International Society for Optics and Photonics, 2012.

[183] B. Tunwattanapong, G. Fyffe, P. Graham, J. Busch, X. Yu, A. Ghosh, and P. Debevec, "Acquiring reflectance and shape from continuous spherical harmonic illumination," *ACM Transactions on Graphics*, vol. 32, no. 4, July 2013.

[184] R. Tur, Y. C. Eldar, and Z. Friedman, "Innovation rate sampling of pulse streams with application to ultrasound imaging," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1827–1842, 2011.

[185] S. Umeyama and G. Godin, "Separation of diffuse and specular components of surface reflection by use of polarization and statistical analysis of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 5, pp. 639–647, 2004.

[186] M. Unser, "Splines: a perfect fit for signal and image processing," *IEEE Signal Processing Magazine*, vol. 16, no. 6, pp. 22–38, Nov 1999.

[187] ——, "Sampling—50 years after shannon," *Proceedings of the IEEE*, vol. 88, no. 4, pp. 569–587, April 2000.

[188] M. Unser and T. Blu, "Cardinal exponential splines: Part I—Theory and filtering algorithms," *IEEE Transactions on Signal Processing*, vol. 53, no. 4, pp. 1425–1438, 2005.

[189] J. A. Urigüen, T. Blu, and P. L. Dragotti, "FRI sampling with arbitrary kernels," *IEEE Transactions on Signal Processing*, vol. 61, no. 21, pp. 5310–5323, 2013.

[190] M. Vetterli, J. Kovacevic, and V. K. Goyal, *Foundations of Signal Processing*. Cambridge University Press, 2014.

[191] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, 2002.

[192] B. Walter, S. Marschner, H. Li, and K. Torrance, "Microfacet models for refraction through rough surfaces," *Eurographics*, pp. 195–206, 2007.

[193] H. Wang, C. Xu, X. Wang, Y. Zhang, and B. Peng, "Light field imaging based accurate image specular highlight removal," *PLOS ONE*, vol. 11, no. 6, p. e0156173, 2016.

[194] G. J. Ward, "Measuring and modeling anisotropic reflection," *ACM SIGGRAPH Computer Graphics*, vol. 26, no. July, pp. 265–272, 1992.

[195] E. T. Whittaker, "On the functions which are represented by the expansions of the interpolation-theory," *Proceedings of the Royal Society of Edinburgh*, vol. 35, pp. 181–194, 1915.

[196] J. M. Whittaker, "Interpolatory function theory," *Bulletin of the American Mathematical Society*, vol. 42, no. 5, pp. 305–306, 1936.

[197] O. Wiener, "Stehende Lichtwellen und die Schwingungsrichtung polarisirten Lichtes," *Annalen der Physik und Chemie*, vol. 40, pp. 203–243, 1890.

[198] ——, "Ursache und Beseitigung eines Fehlers bei der Lippmann'schen Farbenphotographie, zugleich ein Beitrag zu ihrer Theorie," *Annalen der Physik*, vol. 305, no. 10, pp. 488–530, 1899.

[199] D. R. Williams, "Sun fact sheet—sun/earth comparison," https://nssdc.gsfc.nasa.gov/planetary/factsheet/sunfact.html, 2018.

[200] E. Wolf, "Three-dimensional structure determination of semi-transparent objects from holographic data," *Optics Communications*, vol. 1, no. 4, pp. 153–156, 1969.

[201] R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Optical engineering*, vol. 19, no. 1, pp. 139–144, 1980.

[202] S. y Cajal, "Estructura de las imágenes fotocrómicas de G. Lippmann," *Revista de la real Academia de Ciencias Exactas, Físicas y Naturales de Madrid*, vol. 4, pp. 386–428, 1906.

[203] J. Yang, L. Liu, and S. Z. Li, "Separating specular and diffuse reflection components in the HSI color space," in Proc. *IEEE International Conference on Computer Vision Workshops*, pp. 891–898, Dec 2013.

[204] K.-J. Yoon, Y. Choi, and I. S. Kweon, "Fast separation of reflection components using a specularity-invariant image representation," in Proc. *IEEE International Conference on Image Processing*, pp. 973–976, 2006.

[205] B. Zarkov, D. Grujić, and D. Pantelić, "High-resolution dot-matrix hologram generation," *Physica Scripta*, no. T149, pp. 1–3, 2012.

[206] W. Zenker, *Lehrbuch der Photochromie*. W. Zenker, 1868.

Photographic Credits

All photographs and figures in this thesis have been captured or made by the author (and in a few cases by his collaborators). The photographs that serve as title for each chapter often involved several hours of hiking and overnight stays in remote places to catch the golden hour. Below is a short description of them as well as their connection to the topic of their respective chapters.

Acknowledgments

The Troll of Skógafoss Waterfall, South Iceland. 14 June 2015. We can distinguish one or two faces in the moss on the left side, as this passage is dedicated to people. The rainbow is also an allusion to the theme of the thesis.

Chapter 1

Finsteraarhorn and Schreckhorn from Oberaarhorn at dawn, Bernese Alps, Switzerland. 28 July 2016. The first light of the day announces the start of our journey in the land of light.

Chapter 2

Kirkjufell mountain, Snæfellsnes peninsula, Iceland. 5 June 2015. The shape of the mountain closely resembles a VPW pulse.

Chapter 3

Lacs de Fenêtre, Swiss Alps. 26 July 2015. The reflection of light is showcased here by the inverted view of the mountains on the surface of the lake.

Chapter 4

Hvítserkur rock formation, North-West Iceland. 7 June 2015. The waves of the ocean leave patterns in the sand, just like light waves imprint interference patterns in Lippmann plates. Note that the patterns from the photograph are not created from interfering waves, but are instead due to a phenomenon called *saltation* [13].

Chapter 5

Reflection of Wetterhorn, Schreckhorn and Finsteraarhorn in Bachalpsee, Bernese Alps, Switzerland. 18 July 2018. The connection with Chapter 5 is that the photograph has been captured with a circular polarizing filter, giving this blue color to the sky and cutting some of the glare on the lake.

Chapter 6

Sunset on Resurrection Bay, Kenai peninsula, Alaska. 5 September 2012. The wooden poles are a direct reference to streams of Diracs and sparse signals.

Conclusion

Ibex on the slopes of Vanil Noir, Fribourg, Switzerland. 21 August 2013. The ibex is reaching the top of the mountain and contemplates the next challenges ahead. Interestingly, EPFL is located in the mist in the top right corner of the photograph.

Gilles **BAECHLER**

Rte de Vallaire 4A
1025 St-Sulpice (VD), Switzerland
Tel: +41 79 759 15 93
gilles.baechler@gmail.com

Swiss citizen
Born February 9th 1986
http://www.gbaechler.com
github: @gbaechler

RESEARCH INTERESTS

Mathematical signal processing, continuous-domain sparse recovery, sampling theory, optics, phase retrieval, computational photography, machine learning, computer vision, computer graphics

EDUCATION

2013 – 2018 **Ph.D. in Computer and Communication Sciences**
Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
Advisors: Prof. Martin Vetterli and Dr. Adam Scholefield

2009 – 2012 **M.Sc. in Communication Systems**
Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

2008 – 2009 **Exchange in Electrical and Computer Engineering**
Carnegie Mellon University (CMU), United States

2006 – 2009 **B.Sc. in Communication Systems**
Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

TEACHING AND SUPERVISION

Teaching assistant for

COM-514	Mathematical Foundations of Signal Processing
COM-415	Audio Signal Processing and Virtual Accoustics
COM-303	Signal Processing for Communications
MATH-111	Linear Algebra
CS-110	Information, Computation, Communication

Co-advisor of undergraduate and M.Sc. students

INDUSTRY

Starting Nov 2018 **Google Inc. – Zürich**, Switzerland
Software Engineer
Google Sense team.

Sep – Dec 2017 **Google Inc. – Zürich**, Switzerland
Software Engineering Intern
Implementation of a backend system for the automated intrusion detection software used by the security team. Several features and bugs resolved on the system.

Sep 2013 – Jul 2014	**Morphean SA – Fribourg**, Switzerland *Research and Development Engineer* Research and development on machine learning and video analysis. In particular, study on the feasibility of a fall detection algorithm based on video streams and implementation in a video surveillance product.
Sep 2011 – Apr 2012	**Qualcomm Inc. – San Diego CA**, United States *Research Intern* Research on ECG signals and finite rate of innovation sampling. Introduction and investigation of a new sampling scheme for ECG signals, called variable pulse width finite rate of innovation.
Jul – Aug 2011	**LCAV, EPFL – Lausanne**, Switzerland *Summer Intern* Projects on finite rate of innovation sampling to get familiar with the tools and theories developed at the laboratory.
Feb – Jul 2010	**Alcatel-Lucent Bell Labs – Murray Hill NJ**, United States *Research Intern* Research on cryptography: secure multi-party computing, study of state-of-the-art algorithms and proposition of speed improvements.
Jun – Aug 2009	**Infoteam SA – Fribourg**, Switzerland *Summer Intern* Development of a news management system for the administration of the canton of Fribourg and demo apps in Android.

SKILLS

Programming	Python, Go, Java, C/C++, Scala, Perl, Shell Script, .NET and C#
Mobile dev.	Google Android, iOS
Databases	Google Spanner, MySQL, SQL Server
Math	Numpy/Scipy, Matlab, Mathematica, Sage Math, Pari/GP, LaTeX
Languages	French, English, German (basic)

PROFESSIONAL ACTIVITIES

Reviewer for IEEE Transactions on Signal Processing
Member of the IEEE Signal Processing Society

PUBLICATIONS

Journal Articles

[17] G. Baechler, A. Latty, M. Pacholska, M. Vetterli, and A. J. Scholefield, "Shedding light on Lippmann photography", 2018, Manuscript to be submitted for publication.

[16] G. Baechler, A. Latty, M. Pacholska, M. Vetterli, and A. J. Scholefield, "Lippmann photography: A signal processing perspective", *IEEE Transactions on Signal Processing*, 2018, Manuscript to be submitted for publication.

[15] G. Baechler, F. Dümbgen, G. Elhami, M. Kreković, and M. Vetterli, "Coordinate difference matrices", *SIAM Journal on Matrix Analysis and Applications*, 2018, Manuscript to be submitted for publication.

[14] G. Baechler, M. Kreković, J. Ranieri, A. Chebira, Y. M. Lu, and M. Vetterli, “Super resolution phase retrieval for sparse signals”, *IEEE Transactions on Signal Processing*, 2018, Manuscript submitted for publication.

[13] G. Baechler, A. J. Scholefield, L. Baboulaz, and M. Vetterli, “Sampling and exact reconstruction of pulses with variable width”, *IEEE Transactions on Signal Processing*, vol. 65, no. 10, pp. 2629–2644, 2017.

Conference Papers

[12] L. V. Jospin, G. Baechler, and A. J. Scholefield, “Embedded polarizing filter design to separate diffuse and specular reflection”, in *Asian Conference on Computer Vision*, To appear, 2018.

[11] G. Baechler, F. Dümbgen, G. Elhami, M. Kreković, R. Scheibler, A. J. Scholefield, and M. Vetterli, “Combining range and direction for improved localization”, in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2018, pp. 3484–3488.

[10] M. Kreković, G. Baechler, I. Dokmanić, and M. Vetterli, “Structure from sound with incomplete data”, in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2018, pp. 3539–3543.

[9] G. Baechler, I. Dokmanić, L. Baboulaz, and M. Vetterli, “Accurate recovery of a specularity from a few samples of the reflectance function”, in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2016, pp. 1596–1600.

[8] A. Nair, P. Marziliano, R. F. Quick, R. E. Crochiere, and G. Baechler, “Multichannel ECG analysis using VPW-FRI”, in *International Conference on Sampling Theory and Applications*, 2013, pp. 125–128.

[7] G. Baechler, N. Freris, R. F. Quick, and R. E. Crochiere, “Finite rate of innovation based modeling and compression of ECG signals”, in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 1252–1256.

[6] R. F. Quick, R. E. Crochiere, J. H. Hong, A. Hormati, and G. Baechler, “Extension of FRI for modeling of electrocardiogram signals”, in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012, pp. 2909–2912.

Patents

[5] M. Pacholska, G. Baechler, A. Latty, A. J. Scholefield, M. Vetterli, and Y. Bellouard, *Printing multi-spectral images*, Provisional Patent, 2017.

[4] A. Latty, G. Baechler, M. Pacholska, A. J. Scholefield, and M. Vetterli, *Interferometric spectral camera*, Provisional Patent, 2017.

[3] G. Baechler, R. Parhizkar, and M. Vetterli, *Indoor localization method and system*, US Patent, US 9164164B2, 2013.

[2] G. Baechler and A. Hormati, *System and method for analysis and reconstruction of variable pulse-width signals having low sampling rates*, US Patent, US 9275013B2, 2012.

[1] G. Baechler and V. Y. Kolesnikov, *Securing two-party computation against malicious adversaries*, US Patent, US 20120070000A1, 2012.